Architectural Based Data Aggregation Techniques in Wireless Sensor Network: A Comparative Study

Kiran Maraiya[#], Kamal Kant[*], Nitin Gupta[#]

[#] Department of Computer Science and Engineering
National Institute of Technology Hamirpur, Hamirpur (H.P.) INDIA

[*] Department of Computer Science and Engineering
ASET, Amity University, Noida (U.P) ,India

***Abstract*— Data aggregation is very crucial techniques in wireless sensor network. Because with the help of data aggregation we reduce the energy consumption by eliminating redundancy .when wireless sensor network deployed in remote area or hostile environment. In the wireless sensor network have most challenging task is life time so with help of data aggregation we can enhance the lifetime of the network .In this paper we discuss the architectural based efficient data aggregation. And also discuss the protocol based network architecture.**

***Keywords*— wireless sensor network, data aggregation, architecture**

I. INTRODUCTION

The wireless sensor network is ad-hoc network. It consist a small light weighted wireless nodes called sensor nodes, deployed in physical or environmental condition. And it measured physical parameters such as sound, pressure, temperature, and humidity .these sensor nodes deployed in large or thousand numbers and collaborate to form an ad hoc network capable of reporting to data collection sink (base station). Wireless sensor network have various application like habitat monitoring, building monitoring, health monitoring, military survival lance and target tracking. However wireless sensor network is resource constraint if we talk about energy, computation, memory and limited communication capabilities.

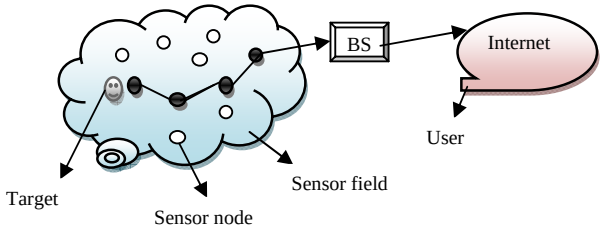

Figure 1 Architecture of the Sensor network

All sensor nodes in the wireless sensor network are interact with each other or by intermediated sensor nodes. A sensor nodes that generates data , based on its sensing mechanisms observation and transmit sensed data packet to the base station (sink). This process basically direct transmission since base station is may located very far away from sensor nodes needs
more energy to transmit data over long distances so that better techniques is to have fewer nodes sends data to the base station. These nodes called aggregator nodes and processes called data aggregation in wireless sensor network.

II. DATA AGGREGATION

A wireless sensor network is a resource constraint network, In which all sensor nodes have limited resources. In order to save resources and energy, data must be aggregated, and avoid amounts of traffic in the network. The aim of data aggregation is that eliminates redundant data transmission and enhances the life time of energy in wireless sensor network. Data aggregation is the process of one or several sensors then collects the detection result from other sensor. The collected data must be processed by sensor to reduce transmission burden before they are transmitted to the base station or sink. The wireless sensor network has consisted three types of nodes. Simple regular sensor nodes, aggregator node and querier. Regular sensor nodes sense data packet from the environment and send to the aggregator nodes basically these aggregator nodes collect data from multiple sensor nodes of the network, aggregates the data packet using a some aggregation function like sum, average, count, max min and then sends aggregates result to upper aggregator node or the querier node who generate the query.

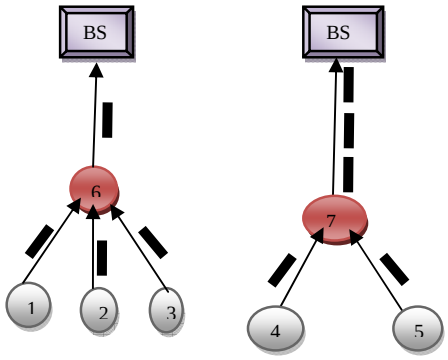

Figure 2 Data aggregation model and Non data aggregation model

It can be the base station or sometimes an external user who has permission to interact with the network. Data transmission between sensor nodes , aggregators and the querier consume lot of energy in wireless sensor network.Figure 2 contain two models one is data aggregation model and second is non data aggregation model in which sensor nodes 1,2,3,4,5 are regular nodes that collecting data packet and reporting them back to the upper nodes where sensor nodes 6,7 are aggregators that perform sensing and aggregating at the same time. In this aggregation model 4 data packet travelled within the network and only one data packet is transmitted to the base station (sink). And another non data aggregation model 3 data packet travelled within the network and all data packets are send to the base station(sink), means we can say that with the help of data aggregation process we decrease the number of data packet transmission. And also save energy of the sensor node in the wireless sensor network. With the help of data aggregation we enhance the life time of wireless sensor network. Sink have a data packet with energy efficient manner with minimum data latency. So data latency is very important in many applications of wireless sensor network such as environment monitoring, health, monitoring, where the freshness of data is also an important factor. It is critical to develop energy-efficient data-aggregation algorithms so that network lifetime is enhanced. There are several factors which determine the energy efficiency of a sensor network, such as network architecture, the data-aggregation mechanism, and the underlying routing protocol.

III. ADVANTAGE AND DISADVANTAGE OF DATA AGGREGATION IN WIRELESS SENSOR NETWORK

Advantage: With the help of data aggregation process we can enhance the robustness and accuracy of information which is obtained by entire network, certain redundancy exists in the data collected from sensor nodes thus data fusion processing is needed to reduce the redundant information. Another advantage is those reduces the traffic load and conserve energy of the sensors. **Disadvantage**: The cluster head means data aggregator nodes send fuse these data to the base station .this cluster head or aggregator node may be attacked by malicious attacker. If a cluster head is compromised, then the base station (sink) cannot be ensure the correctness of the aggregate data that has been send to it. Another drawback is existing systems are several copies of the aggregate result may be sent to the base station (sink) by uncompromised nodes .It increase the power consumed at these nodes.

IV. PERFORMANCE MEASURE OF DATA AGGREGATION

There are very important performance measures of data fusion algorithm. These performances are highly dependent on the desired application.

***Energy Efficiency*:** By the data-aggregation scheme, we can increase the functionality of the wireless sensor network. In which every sensor nodes should have spent the same amount of energy in every data gathering round. A data-aggregation scheme is energy efficient if it maximizes the functionality of the network. Network lifetime, data accuracy, and latency are some of the significant performance measures of data-aggregation algorithms. The definitions of these measures are highly dependent on the desired application.

Network lifetime:The network lifetime is defining the number of data fusion rounds. Till the specified percentage of the total nodes dies and the percentage depend on the application .If we talk about some application,simultenously working of the all the sensor nodes is crucial hence the lifetime of the network is number of round until the first nodes which improves the energy efficiency of nodes and enhance the lifetime of whole network.

Latency: Latency is evaluate data of time delay experiences by system, means data send by sensor nodes and received by base station(sink).basically delay involved in data transmission, routing and data aggregation.

***Communication overhead*:** It evaluates the communication complexity of the network fusion algorithm.

***Data accuracy*:** It is a evaluate of ratio of total number of reading received at the base station (sink) to the total number of generated. There are different types data- aggregation protocols like network architecture based data-aggregation protocols, network-flow-based data-aggregation protocols and quality of service (QOS)-aware data-aggregation protocols designed to guarantee QOS metrics. Here network architecture based protocols are described in detail.

V. IMPACT OF DATA AGGREGATION IN WIRELESS SENSOR NETWORK

In this paper we discuss the two main factors that affect the performance of data aggregation methods in wireless sensor network, Such as energy saving and delay. Data aggregation is the process, in which aggregating the data packet coming from the different sources; the number of transmission is reduced. With the help of this process we can save the energy in the network. Delay is the latency connected with aggregation data from closer sources may have to held back at intermediate nodes in order to combine them with data from source that are farther away. Basically aggregation method based on the position of the sources in the network, number of sources and the network topology. If the examine the factors, we consider the two models of the source placement. The event radius model and random source model [14]. The modelling says us that where the source are clustered near each other or located randomly, significant energy gains are possible with data aggregation. These gains are greatest when the number of sources is large, and when the sources are located relatively close to each other and far from base station. The modelling through, also seems to the suggest that aggregation latency could be non negligible.

VI. BASIC ARCHITECTURE OF DATA AGGREGATION IN WIRELESS SENSOR NETWORK

A. Centralized Architecture:

Centralized architecture is very simplest architecture of wireless sensor network. In which we can apply data fusion process. means each sensor nodes sense a data and transmit to the one central node, called central processor fusion node .this central processor fuse the reports collected by all sensor nodes. In this architecture central node have a responsibility of whole network. The basic advantage of this architecture is it can be easily detected erroneous report of information which is taken by the entire wireless sensor network. The disadvantage is that inflexible to sensor changes and the workload is concerned at a single point.

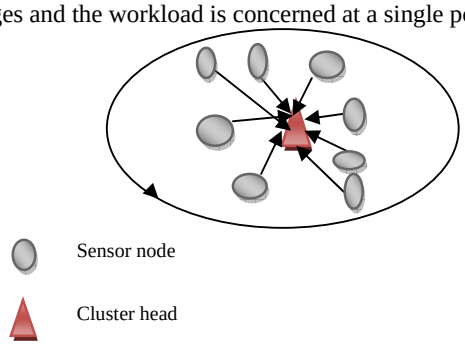

Figure 3 Centralized architecture of wireless sensor network

B. Decentralized Architecture:

The decentralized architecture of wireless sensor network, there is no single centralized node that makes decisions on behalf of all the sensor nodes. Data fusion occurs locally at each node on the basis of local observations and the information obtained from neighbouring nodes. in which all sensor nodes are connected to each other on the observation .The advantage of this architecture are scalable and tolerant to the addition or loss of sensing nodes or dynamic changes in the network.

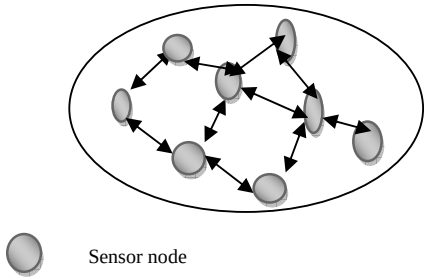

Figure 4 Dencralized architecture of wireless sensor network

C. Hierarchical Architecture:

The hierarchical architecture of wireless sensor network is very important, In this architecture all sensor node are partitioned into hierarchical level. In the practical sensor network, the level 0 may contain many normal sensors organized in a topographical area, and to minimize the transmission power, the data from individual sensor nodes will be forwarded to all the distant fusion nodes by adopting a suitable routing algorithm and minimize the transmission power, the data of sensor can be forwarded to a fusion node through the nearly sensor nodes using routing algorithm like directed diffusion or simple flooding. The advantage of this architecture is that Workload is balanced among nodes in the wireless sensor network.

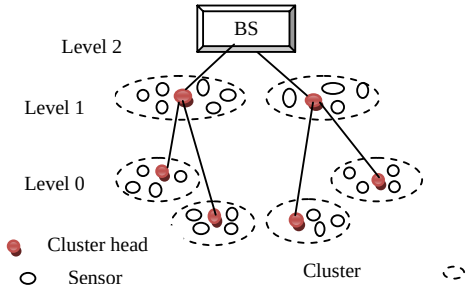

Figure 5 Hierarchical architecture of wireless sensor network

VII.DATA AGGREGATION BASED IN NETWORKS

The architecture of the sensor network plays a vital role in the performance of different data-aggregation protocols.

1.FLAT NETWORKS

Flat networks plays very important role in wireless sensor network, in which each sensor node have a equal battery power and plays the same type of role in a network. In such type of networks, data aggregation has to be done in data centric routing manner, where the sink generally sends a data packet to the sensor nodes, such as, flooding. In the flooding sensors which have data matching the data packet and transmit response data packet back to the sink.

a) Push Diffusion: Data packet flowing with the help of this diffusion techniques, In which all source node are active participants and start the diffusion while the base station respond to the source node.

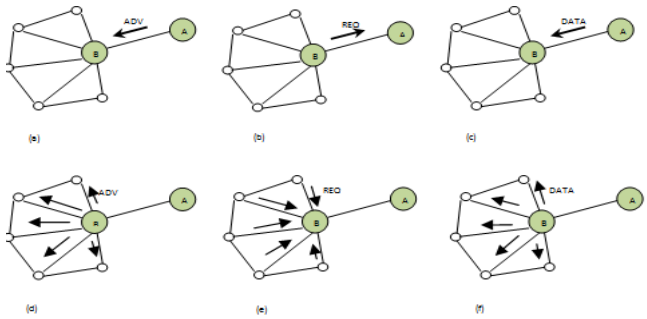

Figure 6 SPIN protocol (a) Node A starts by advertising its data to node B, (b) Node B responds by sending a request to node A. (c) After receiving the requesting data, (d) Node B then sends out advertisements to its neighbours, (e-f) who in turn send request back to B.

The sensor protocol for information via negotiation (SPIN) [3] is category of push based diffusion protocol**.** The staring node which has new data advertises the data to the neighbouring nodes in the network using the meta data. A neighbouring node which is interested in this kind of data sends a request to the initiator node for data. The initiator node responds and sends data to the sinks. Each node has a resource managing capability to keeps track of its energy usage in the sensor network. Each node polls its resources such as battery power before data transmission. SPIN is also well-suited for environments with mobile sensors, since the forwarding decisions are based on local neighbourhood information. One of the main advantages of SPIN is that topological changes are localized, since each node only requires the knowledge of its single-hop neighbours. The main disadvantage of SPIN is its inability to guarantee data delivery. For instance, in intrusion detection applications, if the nodes

interested in the data are farther away from the source node, and the intermediate nodes are not interested in the data, then the data is not delivered to the sink nodes.

b) Directed Diffusion: It is an energy-efficient data-aggregation protocol for a wireless sensor network. It is application aware paradigm; means data of the sensor is named by attribute value. With the help of directed diffusion we can enhance the life time of the network. In this scheme generally base station broadcast the message to the interested source node. After that each node receives interest. These interests define the attribute value such as name of object. Each node the get the interest can cache it for later use. As the interest is broadcasted by the network hop by hop, gradient are setups to draw data satisfying the query toward the requesting node. A gradient is a reply link to the closer from which the interest was received.

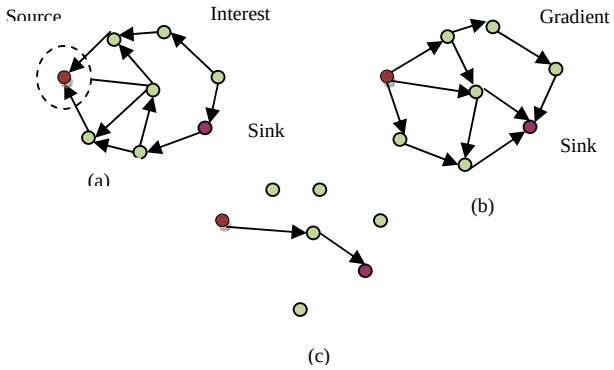

Figure 7 Directed Diffusion protocol phases. (a) Interest Propagation (b) Initial Gradient set up and (c) Data delivery along reinforcement.

2. HIERARCHICAL NETWORKS

All the communication and computation burden at the sink in flat network, that's why lot of energy is consumed. In the hierarchical network, In which data aggregation data has to be done at special nodes, with the help of these special node we can reduce the number of number of data packet transmitted to the sink. So with this network improves the energy efficiency of the whole network. Various type hierarchical data-aggregation protocols as follows.

a) Cluster-Based Networks for data aggregation: Wireless sensor network is resource constraint that's why sensor cannot directly transmit data to the base station. In which all regular sensors can send data packet to a cluster head (local aggregator) which aggregates data packet from all the regular sensors in its cluster and sends the concise digest to the base station. With the help of the scheme we save the energy of the sensors. Figure 3 shows a cluster-based sensor network. The cluster heads can communicate with the base station directly.

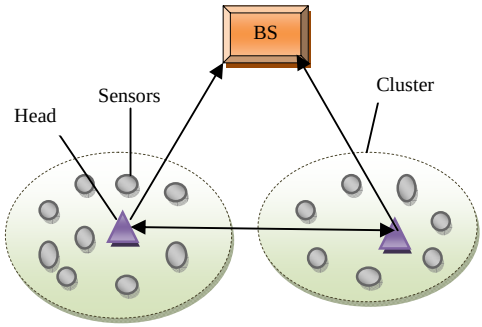

Figure 8 Cluster based sensor network. The arrows indicate wireless communication links

LEACH is the cluster-based network and data-aggregation protocol,[5] Low-Energy Adaptive Clustering Hierarchy (LEACH). It is a first energy conserving cluster formation protocol. The LEACH protocol is distributed and all sensor nodes organize into clusters for data aggregation (fusion). In which cluster head in each cluster sends the aggregated (fused) data from some sensor node in its cluster to the base station. This reduces the total number of information that is send to the base station. The data fusion is performed periodically at the cluster heads. LEACH performs two phases, the setup phase and the steady-state phase. In the setup phase networks are organise into the clusters and the selection of cluster heads. The steady-state phase

involves data aggregation process has to be done at cluster heads and data transmission to the base station (sink). LEACH has some limitations. LEACH says that all sensors have a capability to act as a cluster head and perform data aggregation process. LEACH also assumes that nodes have data packet to transmit sporadically. In LEACH, all nodes have the same amount of energy capacity in each selection round, which is based on the assumption that being a cluster head results in same energy consumption for every node. Hence, LEACH should be extended to account for node heterogeneity. In an improved version of this protocol, called LEACH-C [6], cluster formation is performed in a centralized manner by the sink. LEACH-C improves the performance of LEACH by 20 to 40 percent in terms of the number of successful data gathering rounds.

c) Data Aggregation in chain based network: In which each sensor sends data to the closer neighbour. Power-Efficient Data-Gathering Protocol for Sensor Information Systems (PEGASIS) is type of chain based data aggregation. In PEGASIS, all sensors are structured into a linear chain for data aggregation. The nodes can form a chain by employing a greedy algorithm or the sink can decide the chain in a centralized manner. In the Greedy chain formation assumes that all sensors have inclusive knowledge of the network. The farthest node from the sink initiates chain formation and, at each step, the closest neighbour of a node is selected as its successor in the chain. In each data-gathering round, a node receives data packet from one of its neighbours, aggregates the data with its own, and sends the aggregates data packet to its other neighbour along the chain. Eventually, the leader node in the are similar to cluster head sends the aggregated data to the base station. Figure 9 shows the chain-based data-aggregation procedure in PEGASIS.

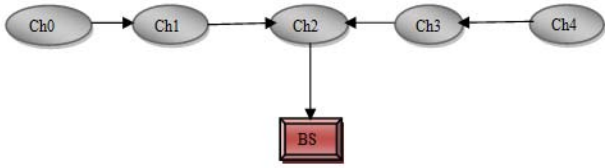

Figure 9 Chain in PEGASIS

d) Tree-Based Data Aggregation In which all node are organized in form of tree means hierarchical, with the help of intermediate node we can perform data aggregation process and data transmit leaf node root node. Tree-based data aggregation is suitable for applications which involve in-network data aggregation. An example application is radiation-level monitoring in a nuclear plant where the maximum value provides the most useful information for the safety of the plant. One of the main aspects of tree-based networks is the construction of an energy efficient data-aggregation tree.

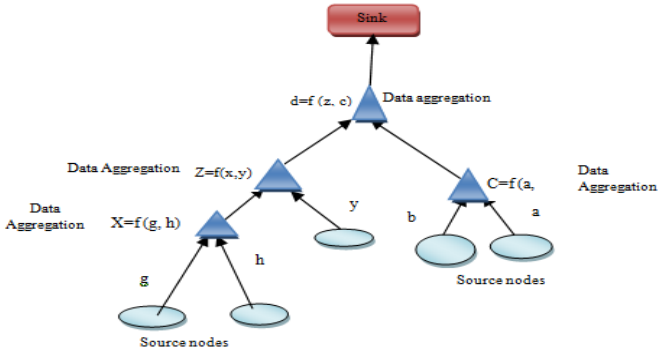

Figure 10 Minimum spanning tree – based routing protocol in a sensor network. Arrows indicate the routing path and f(.., ..) is the data aggregation function.

Ding *et al.* [8] have proposed an energy-aware distributed heuristic (EADAT) to construct and preserve a data-aggregation tree in sensor networks. S. Madden et al. in [15] proposed a Tiny Aggregation (TAG) approach [15]. TAG works in two phases: distribution phase and collection phase. If we talk about distribution phase, TAG organizes nodes in to a routing tree rooted at sink. The tree arrangement starts with broadcasting a data packet from base station specify level or distance from root. When a sensors gets this data packet it sets its own level to be the level of message plus one and select parent as sensor node from which it receives the data packet. After that, sensor node rebroadcast this message with its own level. This process continues until all nodes select their parent. After tree arrangement, sink send queries along structure to all nodes in the network. The tree can be re-constructed periodically from the sink. A sleeping sensor periodically wakes up and broadcasts hello message, which contains its path length to the sink. Each intermediate sensor receiving the interest must broadcast it at least once to setup the reverse path to the sink. The target sensor sends back the data along several paths.

e) Grid-Based Data Aggregation: In which a set of sensors is assigned as data aggregators in fixed regions of the sensor network. The sensors in a grid send the data packet directly to the aggregator of that grid. Hence, the sensors within a grid do not communicate with each other. In-network aggregation is similar to grid-based data aggregation with two major differences; each sensor within a grid communicates with its neighbouring node. Any node within a grid can assume the role of aggregator node in terms of rounds until the last node dies. This is similar to cluster-based data aggregation in which the cluster heads are fixed. In in-network aggregation, the sensor with the most critical information aggregates the data packets and sends the fused data to the sink. Each sensor transmits its signal strength to its neighbours. If the neighbour has higher signal strength, the sender stops transmitting packets. After getting data packets from all the neighbours, the node that has the maximum signal strength becomes the data aggregator. The in-network aggregation scheme is best suitable for environments where events are highly localized.

VIII. COMPARISION BETWEEN HIERARCHICAL AND FLAT NETWORK

Hierarchical network	Flat Network
Data aggregation performed by cluster heads or leader node	Data aggregation is performed by different nodes along the multi-hop path
Overhead involved in cluster or chain formation throughout the network	Data aggregation routes are formed only in regions that have data for transmission
Even if one cluster head fails, the network may still be operation	The failure of sink node may result in the breakdown of entire network
Lower latency in involved since sensor nodes perform short rang transmission to the cluster head	Higher latency is involved in the data transmission to the sink via multihop path.
Routing structure is simple but not necessarily optimal	Optimal routing can guaranteed with additional overhead
Node heterogeneity can exploited by assigning high energy nodes as cluster heads	Does not utilize node heterogeneity for improving energy efficiency.

Table 1 Comparison between hierarchical and flat networks

Protocol Name	Flat Network	Hierarchical Network
SPIN	✓	-
Directed Diffusion	✓	-
Energy Aware Routing	✓	-
Roumar Routing	✓	-
Gradient Based Routing	✓	-
CADR (Constrained Anisotropic Diffusion Routing	✓	-
ACQIRE (Active Query Forwarding in Sensor Network)	✓	-
LEACH	-	✓
PEGASIS	-	✓
Hierarchical PEGASIS	-	✓
TEEN	-	✓
APTEEN	-	✓
Energy Aware Routing For Cluster Based Sensor Network	-	✓

Table 2 Routing protocols for flat and hierarchical network

XI. CONCLUSION

In this paper we present wireless sensor network is consist a sensor node. And these nodes are resource constraint. That's why lifetime of the network is limited so the various techniques or protocol has been proposed for increasing the lifetime of the wireless sensor network. In this paper data aggregation is one of the important techniques for enhancing the life time of the network. In this paper we discuss Compressive survey or architectural based efficient data aggregation techniques, or also discuss the advantage and disadvantages and various performance measures of the data aggregation.

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey On Sensor Networks", *IEEE Communications Magazine*, Volume 40, Number 8, pp.102-114, 2002.

[2] T. Arampatzis , J. Lygeros, and S. Manesis, "A Survey Of Applications Of Wireless Sensors And Wireless Sensor Networks", *In Mediterranean Conference On Control And Automation MED05*, Nicosia, Cyprus, 2005.

[3] J. Kulik, W. R. Heinzelman, and H. Balakrishnan, "Negotiationbased Protocols for Disseminating Information in Wireless Sensor Networks," Wireless *Networks*, vol. 8, Mar. 2002, pp. 169–85.

[4] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *Proc. 6th Annual Int'l. Conf. Mobile Comp. and Net. (MobiCOM '00)*, Aug. 2000.

[5] W. R. Heinzelman, "Application-Specific Protocol Architectures for Wireless Networks," Ph.D. thesis, Massachusetts Institute of Technology, June 2000.

[6] W. R. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Trans Wireless Commun.*, Oct. 2002, pp. 660–70.

[7] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy metrics," *IEEE Trans. Parallel and Distributed Systems*, vol. 13, no. 9, Sept. 2002, pp. 924–35.

[8] M. Ding, X. Cheng and G. Xue, "Aggregation Tree Construction in Sensor Networks," *2003 IEEE 58th Vehic. Tech. Conf.*, vol. 4, no. 4, Oct. 2003, pp. 2168–72.

[9] K. Vaidhyanathan *et al.*, "Data Aggregation Techniques in Sensor Networks," Technical Report, OSU-CISRC-11/04-TR60, Ohio State University, 2004

[10] V. Raghunathan, C. Schugers, S Park, and M.B. Srivastava, "Energy-Aware Wireless Microsensor Networks", *IEEE Signal Processing Magazine*, Volume 19, Number 2, pp. 40-50, 2002.

[11] Lin F.Y.S., Yen H.H., Lin S.P., Wen Y.F.: 'MAC aware energyefficient data-centric routing in wireless sensor networks'. Proc. IEEE Int. Conf. Commun. (ICC), 2006, vol. 8, pp. 3491–3496

[12] Krishnamachari B., Estrin D., Wicker S.: 'Modeling datacentric routing in wireless sensor networks', USC Computer Engineering Technical Report, CENG 02-14, 2002.

[13] H.H. Yen, C.L. Lin: 'Integrated channel assignment and data aggregation routing problem in wireless sensor networks', IET Communications, 2009, Vol. 3, Iss. 5, pp. 784–793.

[14] Ramesh Rajagopalan and Pramod K. Varshney, Syracure University "Data aggregation techniques in sensor network : A survey" "IEEE communication survey & tutorial.4th quarter 2006

[15] A. Chandrakasan W.R. Heinzelman and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor network. IEEE Proceedings of the Hawaii International Conference on System Sciences, pages 1{10, January 2000.

[16] Ismail H. Kasimoglui Ian .F. Akyildiz. Wireless sensor and actor :research challenges. *(Elsevier) Journal*, 2(38):351{367, 2004.