Analysis of Various Deterioration Factors of Data Aggregation in Wireless Sensor Networks

N.Sugandhi[1], D.Manivannan[2]
School of Computing, SASTRA University, Tirumalaisamudram,
Thanjavur, Tamilnadu, India - 613 401.

1sugandhinatarajan@gmail.com
2dmv@cse.sastra.edu

Abstract

In embedded systems, wireless sensor network is the interesting field which provides various applications such as battlefield surveillance, disaster management, habitat monitoring, home automation, health care systems etc. Wireless sensor network (WSN) handles a huge number of tiny, cost effective, low powered, autonomous devices called sensor nodes which can perform sensing, processing and communicating operations. WSN has many constraints like energy utilization, limited memory and power consumption. Data aggregation is an intelligent technique which accumulates data from disparate sources by using various aggregation techniques. One of the crucial techniques in WSN is data aggregation which greatly increases the energy efficiency and network lifetime by reducing the number of data transmissions. To perform data aggregation in an efficient manner, there are lot of issues are to be considered. In this paper several issues related with data aggregation are discussed and comparisons of various design issues of various data aggregation protocols are also addressed.

Keywords: WSN, Aggregation, Security, Flat networks, Hierarchical networks.

I. Introduction

The WSN is formed by various sensor nodes. It may be homogeneous or heterogeneous. The nodes can sense the temperature, pressure, vibration, motion, humidity, sound etc. After sensing, processing can be taken place. This can be achieved by using various processors with implementing different techniques. The aggregation plays a vital role here for improving the efficiency and network performance. The aggregation function is nothing but combining and summarizing multiple data into single data. This greatly reduces redundancy and conserves the node energy. The aggregation function may be of selecting MAX, MIN, AVE, COUNT, SUM, MEDIAN etc. from gathering values [1]. Depends upon the applications the aggregation function can be chosen. The main consideration while performing data aggregation is providing security because there is a plenty of chances for security attacks in WSN. Due to the hostile environment of WSN, providing security in data aggregation is really a challenging task [2].

In the cluster type of networks, cluster head will be acting as aggregator node otherwise intermediate nodes will act as the aggregator nodes. Depends upon the number of sensor nodes and applications, the count of aggregators can be fixed. It can perform all aggregation functions and transmit to the BS. The aggregator node should be more secure. If it is compromised, the entire network will be affected [3]. The nearby nodes in WSN can gather the same data. So it is not meaningful for collecting redundant information. In this sort of cases, data aggregation can be effectively used It is used to reduce the number of transmitting packets and bandwidth utilization. The aggregation technique is simply known as the intelligent way of data combining and compression technique. Each technique has its own advantages as well as disadvantages. The main advantages of data aggregation are improving network performance, conserving energy, eliminating redundancy, increasing network lifetime, efficient utilization of bandwidth, enhancing accuracy, providing robustness, reducing traffic etc. The disadvantages are if the aggregator node is compromised, the entire network will be affected. The data accuracy is also to be spoiled [4].

In this paper, section 2 describes aggregation approaches, section 3 gives various performance measures, section 4 explains about architectural and design issues, section 5 deals about WSN attacks and its solutions and finally section 6 gives comparison of available secure aggregation protocols.

II. Basic Approaches of Data Aggregation

There are three general approaches used to perform data aggregation. They are 1) Centralized 2) Decentralized 3) In-network aggregation [5].

A. Centralized Approach

In this approach only one central node (A-Aggregator) is used for performing the aggregation function. All other nodes (s1, s2, s3, s4 and s5) are just sensing data and transmit to the central aggregator node. Every node in the network is connected to the central aggregator. It has full responsibility for data aggregation. It can easily identify the node which sends erroneous data but more workload for a central aggregator node. This can be known as a single data aggregator model. This method described in fig 1.

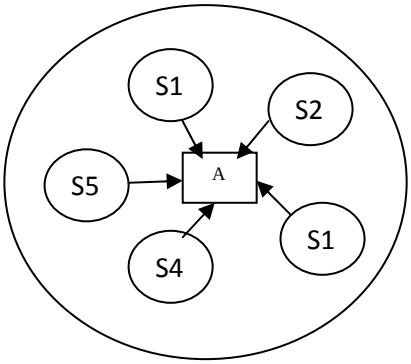


Fig 1. Centralized approach

B. Decentralized Approach

No single centralized node used for this approach. All nodes are connected to its neighbor node. Each node performs the aggregation function locally. All gets equal priority to perform the aggregation function. This approach is more scalable and tolerant of dynamic changes and node failures. It is also called as multi data aggregator model which is given in fig 2.

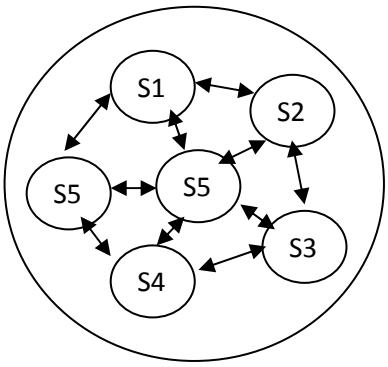


Fig 2. Decentralized Approach

C. In-network Aggregation

This approach is very efficient to improve the lifetime of the network and reduce the amount of data which needs to be transmitted. In this, the aggregator node collects the sensed data from various nodes, process the data based on aggregation function and convert the multiple data into single data, transmit the final data to the base station (BS) given in fig 3. It greatly reduces the energy consumption by sending multiple data. The count of aggregator node may be one or more depends upon the structure of the network. This approach can be classified as with size reduction and without size reduction methods. With size reduction method refers that reducing the packet length by compressing multiple data into single data. Without size reduction refers that merging the data without processing it.

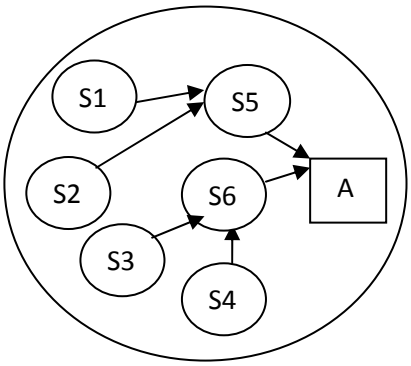


Fig 3. In-network aggregation

III. Performance Measures

The performance of data aggregation can be analyzed by different performance metrics such as energy efficiency, network lifetime, data accuracy, latency etc. These metrics are dependent on the type of WSN applications and architecture of WSN used.

A. Energy Efficiency

Without aggregation, all the nodes should send the data to the BS individually. If it is far from the nodes, it will be very difficult for communication. It is a waste of power only. It degrades the lifetime of the network also because the WSN has a battery constrained nodes only. So the aggregation effectively uses the energy.

B. Network Lifetime

It can be calculated by using first node dies. If the energy is effectively used, the lifetime of the network will also be improved. It depends on the count of performing data aggregation also. Aggregation is the main parameter to increase the network lifetime.

C. Data Accuracy

It refers that the correctness of the transmitted data. Depends upon the localization of the destination node, routing path, accuracy can be varied. It is one of the essential metrics while evaluating data integrity.

D. Latency

It is nothing but the time delay between the data transmission by source and data reception by destination. To perform aggregation, some delay is introduced. But it should be less while comparing with individual data transmission and reception.

IV. Design Issues of Data Aggregation

Issues can be mostly based on the network architecture and security point of views. Network architecture provides the descriptive view for connection and communication between nodes. Depends upon the architecture, aggregation methods can be classified as structured aggregation and structure less aggregation. In this structured aggregation uses specific architecture for performing data aggregation. The architecture can be majorly classified as flat and hierarchical network topology. Hierarchical topology can be divided into tree, cluster, grid and chain architectures. The structure less data aggregation does not use any kind of specific architecture in fig 4. So the communication takes place at any node to node in the network [6].

A. Architectural Issues

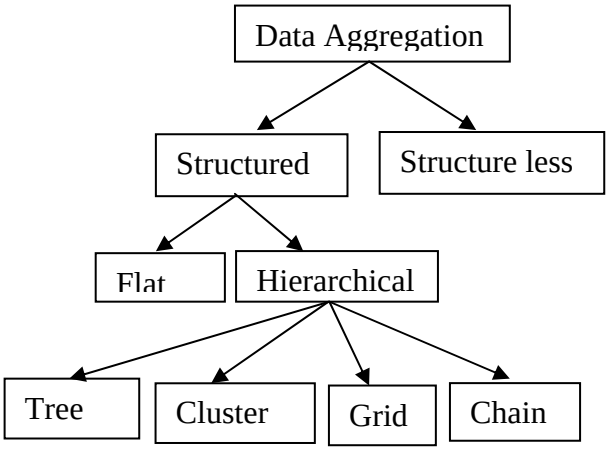
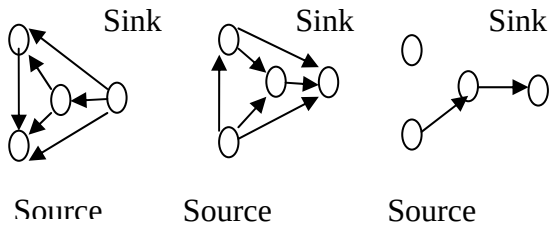


Fig 4.Classification of data aggregation methods

1) Data aggregation in flat network: In flat networks, all the sensor nodes have equal properties and perform equal role. Data aggregation can be done by using data centric routing. In this method, base station sends a query by means of flooding to all the sensor nodes. If any node has the data matches with the query transmit the response back to the BS. The multi hop path can be used to perform data aggregation. So that latency can be increased. flooding and gossiping, SPIN, directed diffusion, rumor routing, gradient based routing are some of the examples of routing protocols performing data aggregation in flat networks.

1) Directed Diffusion: It is one of the well known data aggregation routing protocols for WSN. It follows the data centric routing algorithm. On that, routing is based on the content of the data packets not based on nodes. The concept of data centric routing is to aggregate the data from various sources by reducing the redundant data given in fig 5. This routing finds routes from multiple sources to a single destination. Initially, all the sensor nodes measure the events and create gradients of information in their neighborhood. If BS wants to collect the information about the event, it will broadcast the interests of the network by hop to hop communication [8]. Intermediate nodes propagate these interests. Interests are nothing but the query tasks which required by the network. After getting the interests, the node sends the gradient about that to the respective node. The gradient refers to an attribute value and a direction. If the gradient satisfies the interests, that path is reinforced to prevent further flooding. If the BS receives data from the sources, it will refresh and re-sends the interests periodically.



a) Interests propagation b) Gradients setup c) Reinforcement delivery

Fig 5. Directed Diffusion

Drawbacks of flat networks

In flat networks, the routes for data aggregation are established only to the specific regions which have the transmitting data. Scheduling is based on contention only. Latency is increased due to the usage of multi hop path. If the sink node fails, the entire network will be collapsed. It increases the overhead also.

2) Data Aggregation in Hierarchical Networks: Hierarchical networks overcome the drawbacks of flat networks. It uses clustering, node heterogeneity and reservation based scheduling. In this type of networks, data aggregation is performed at cluster heads. So overhead is increased at CH. Even if one CH fails, aggregation is possible in other clusters. Sensor nodes can perform a short range communication to CH.

2) Tree based data aggregation: Nodes are organized in the form of a tree. It mainly consists of one root node (BS), intermediate nodes and leaf nodes which is given in fig 6. Leaf nodes are used to sense the data and the intermediate nodes are used to perform aggregation and transmit to the root node. In network aggregation can be achieved by this method. The main goal is nothing but to build energy efficient tree which performing data aggregation. EADAT (Energy-Aware Data Aggregation Tree), E-SPAN (energy-aware spanning tree) and TAG (Tiny Aggregation) are some of the examples for tree based data aggregation. In EADAT, root node broadcasts a periodic control message. Each node has the timer settings and it gets started after receiving the message from root node at first time. In this, the expiration time of the timer and node's residual energy inversely proportional to each other. In this each parent node should perform the aggregation function so the computation overhead will be increased.

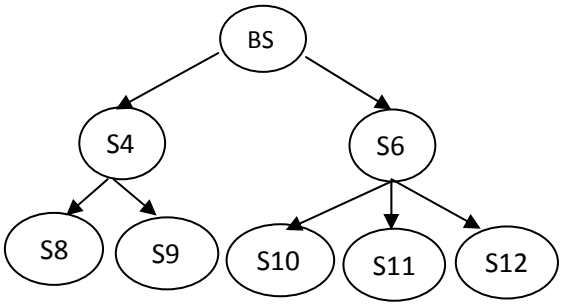


Fig 6. Tree based data aggregation

2) Cluster based data aggregation: To handle huge number of nodes and to reduce overhead, cluster based data aggregation is preferred [12]. Nodes are connected to form a cluster. Each cluster has one leader node that performs data aggregation, called cluster head given in fig 7. Normally sensor nodes are energy constrained, it does not have energy to transmit data to the BS directly. So this sort of aggregation is preferred most for energy saving and improves the lifetime of the network. LEACH (Low Energy Adaptive Clustering Hierarchy) is one of the well-known techniques for clustering mechanism. It has two phases called as setup phase and steady state phase. CH selection and formation are done in setup phase. Data collection and aggregation are done in steady state phase. Finally the aggregated data are transmitted to the BS.It makes use of TDMA for scheduling and CDMA for minimizing collisions. There are many extended versions of protocols are introduced such as E-LEACH (Energy LEACH), M-LEACH (Multi-hop LEACH) etc.

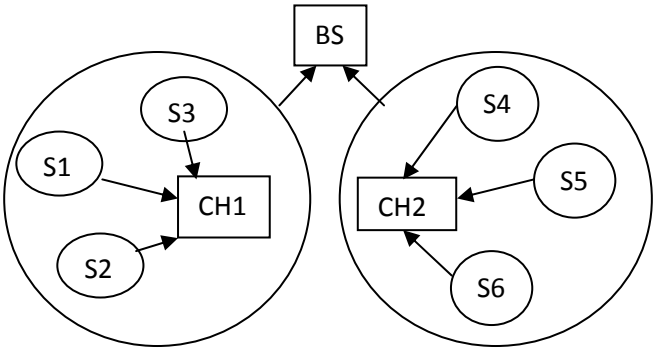
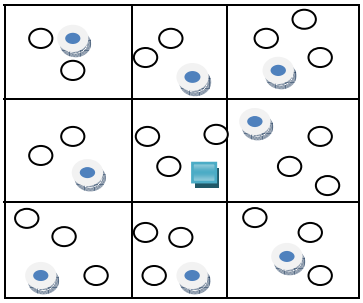


Fig 7. Cluster based data aggregation

2) Grid based data aggregation: The sensor nodes are placed in the form of a grid. It makes use of fixed topology. It performs two levels of data aggregation. local and global aggregations are performed by local and master aggregators respectively. In each and every zone, one CH is used to perform local aggregation, a subset of CHs known as master nodes selected for global aggregation. It is suitable for handling extremely low mobility nodes given in fig 8.



- Sensor node
- Local aggregator
- Global aggregator

Fig 8. Grid based data aggregation

2) Chain based data aggregation: It is one of the hierarchical methods of aggregation which forms chain architecture. The well-known protocol is nothing but PEGASIS (Power-Efficient Gathering in Sensor Information Systems). It greatly overcomes the drawbacks of LEACH by eliminating the overhead of dynamic cluster formation and minimizing the count of transmissions and receptions. By this method, energy can be evenly distributed. In chain architecture, each sensor node can make a communication with its neighbors and each get turn to be the leader for transmitting data to the BS. It uses a token passing approach. After getting the token, each node transmits the data to the aggregator node. Finally it reaches the BS given in fig 9.

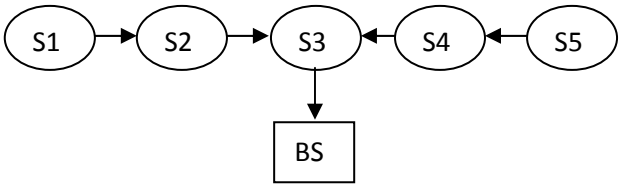


Fig 9. Chain based data aggregation

Structure Less Data Aggregation: It avoided the reconstruction of network architecture at the time of node failures due to energy depletion. This can be efficiently used for event based applications. The main challenge of this approach is making routing decisions and performing aggregation. Data-Aware Anycast (DAA) is the mechanism available for structure less data aggregation. It follows a randomized waiting approach. Initially a source node sends RTS and the type of data to all of its neighbors. If the data match with any of the nodes, it sends back CTS. The source node selects any one of the nodes which sending the CTS. It improves the data aggregation and network performance. Table I provides comparison of flat and hierarchical aggregation methods with its performance metrics.

Aggregation Methods	Energy efficiency	Network Lifetime	Latency	Accuracy
Flat	High	High	High	High
Tree(H)	Low	High	Low	High
Cluster(H)	High	High	Low	Low
Grid(H)	High	Low	Low	Low
Chain(H)	High	High	Low	High

H- Hierarchical
Table I. Comparison of Aggregation methods

B. Security Issues

Data aggregation protocols can be designed by using various design approaches but security is not achieved. So that security issues should be considered. For military and life critical applications, data transmissions, receptions and aggregation should be handled in a secured way [7].To implement this, many concepts should be considered. They are data confidentiality, data integrity, data freshness, source authentication, node availability and localization. [8]

1) Data Confidentiality: It ensures that the data should be accessed only by authorized parties. It should not be revealed to others at any cost. The sensor data and routing information both should be maintained in a secret way. It can be attained by using efficient cryptographic techniques. To achieve a secure data aggregation, it is classified as hop by hop and end to end mechanisms. In structured data aggregation method, sensor nodes transmit the encrypted data to the aggregator node. Then it decrypts that data, perform aggregation and again encrypt the data, send to BS. This process leads to more energy consumption and produces delay. In the second method, the aggregator node performs aggregation over encrypted data by using various techniques for example Privacy Homomorphism. The BS can only have the decryption key to decrypt the aggregated data. It provides end to end confidentiality and improves network performance.

2) Data Integrity: The transferred data should not be corrupted before reaching the intended destination node. The malicious node can inject false data to alter his aggregation values. Data confidentiality protects the data from reaching unintended parties, but it does not guarantee unaltered and uncorrupted data. This factor is also very essential for mission critical applications. So the source node, aggregator node should not be compromised. It's the aggregation value is changed, there is no use of achieving confidentiality also. Message authentication code and digital signatures can be efficiently used for attaining data integrity. Due to unreliable communication paths, the data can be changed. So the paths are also to be handled effectively [11].

3) Data Freshness And Node Localization: The freshness of data greatly prevents the aggregated data from a lot of replay attacks and duplicate messages. Because there is a chance for the replay of old messages. It is a waste of energy only. So the transmitted data should be resent. By achieving data freshness, network performance and energy are effectively used. The location of the destination node is very important for the source node. The

routing can be achieved by means of localization information only. The location should not be revealed to malicious nodes. It should be kept secret.

4) Authentication: Sensor nodes are transmitting and receiving the data only in wireless medium. It is vulnerable in nature because of various intruders and attacks. Node authentication ensures that the source and destination node should not be compromised. The data should the intended receiver node. Data authentication refers that the transferred data should be the same as the original data sent by the source node. For that, MAC (Message Authentication Code) computation is used. It provides shared secret key between the source and destination.

5) Availability

The availability check is also one of the important factors for energy constrained wireless sensor nodes. Due to the energy depletion, the node accessibility is reduced and it can be a chance of more intruding actions. It ensures the network survivability and prevents the attacks.

V. Various Attacks in WSN

There is a chance of various attacks due to the vulnerable nature of the network such as wireless transmission medium, resource constrained nodes, critical deployment areas etc. [3]. Table II. provides the detailed view of various deterioration factors, its causes and solutions.

Types of attacks	Causes	Solutions
Denial of service attacks (jamming)	By making interference with radio frequencies	By use of MAC and spread spectrum techniques
Replay attacks	Because of transmitting same data without data freshness	prevent this attack by time stamping all data packets
Malleability, Forge packets	Due to the injection of malicious nodes	Use of HMAC (Hash based MAC)
Physical attacks	Due to less security of using symmetric key approach	Use of Asymmetric public key approach
Energy drain attacks	Because of energy depletion	Make use of various energy harvesting techniques e.g. solar power
Sybil attacks	Making multiple fake identies	Use of authentication
Sinkhole attack	Attracts traffic to the specific compromised node	Preventing by proper routing and localization information
Sniffing attack	Capturing data by using malicious nodes	Protocols with data confidentiality
Data Integrity Attack	By injecting false data	By using digital signatures

Table II. Comparison of various attacks in WSN

VI. Comparison of Secure Aggregation Protocols

By considering architecture and security concepts, various data aggregation protocols are introduced. Table III gives the comparison about various standard data aggregation protocols details with its architecture, security requirements and performance metrics.

Protocol	Ar	DC	DA	A	NA	EE	NL	R
SIA [9]	T	Y	Y	Y	N	L	L	L
Secure DAV [10]	T	N	Y	Y	N	L	L	M
ESPDA [11]	C	Y	Y	Y	N	M	L	L
SRDA [12]	G	Y	Y	Y	N	M	L	L
SDAP [13]	T	Y	Y	Y	N	L	L	M
SELDA [14]	SL	N	Y	Y	Y	L	L	M
EIRDA [15]	C	N	N	N	N	M	L	M
CDA [16]	SL	Y	N	N	N	L	L	M
EECDA [17]	C	N	N	N	N	M	M	L
DyDAP [18]	SL	N	Y	N	N	L	L	L

Table III. Comparison of various secure aggregation protocols

Explanation of symbols in Table III

Ar- Architecture, DC - Data Confidentiality, DA- Data Authentication, A Authentication, NA-Node Availability, EE-Energy Efficiency, NL-Network Lifetime, R-Reliability, T-Tree, C-Cluster, G-Grid, SL-Structure Less, Y-Yes, N-No, M-More, Less.

VII. Conclusion

This paper provides the importance of data aggregation in WSN, different approaches to implement efficient aggregation, performance measures, architectural and security issues, various deterioration factors affecting data aggregation and its solutions. Additional with this, a brief comparison of available secure aggregation protocols is also given. By using these kinds of data, required data aggregation protocols for various WSN applications can be easily chosen.

References

[1] Vaibhav Pandey, AmarjeetKaur and Narottam Chand, "A review on data aggregation techniques in wireless sensor Network", Journal of Electronic and Electrical Engineering Vol. 1, Issue 2, pp. 01-08, 2010

[2] SuatOzdemir, Yang Xiao, "Secure data aggregation in wireless sensor networks: A comprehensive overview", Elseviar Computer Networks 53, pp. 2022–2037, 2009.

[3] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, "A survey on sensor networks", IEEE Commun. Mag. 40 (8) (2002)102-114.

[4] Priyanka K. Shah and Kajal V. Shukla, "Secure Data aggregation Issues in Wireless Sensor Network: A Survey", Journal of Information and Communication Technologies, Vol. 2, Issue 1, January 2012

[5] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Comput. Networks 52 (12) (2008) 2292-2330.

[6] K. Akkaya, M. Demirbas, R.S. Aygun," The Impact of Data Aggregation on the Performance of Wireless Sensor Networks",,Wiley Wireless Commun. Mobile Comput. (WCMC) J. 8 (2008)171–193.

[7] L. Hu, D. Evans, "Secure aggregation for wireless networks",in: Proceedings of the Workshop on Security and Assurance inAd Hoc Networks, Orlando, FL, 28 January 2003.

[8] Walters, J. P., Liang, Z., Shi, W., and Chaudhary, V., "Wireless sensor network security - A survey", Security in Distributed, Grid, Mobile, and Pervasive Computing, Auerbach Publications, CRCPres, 2007.

[9] B. Przydatek, D. Song, A. Perrig, SIA : secure information aggregation in sensor networks, in: Proceedings of SenSys'03, pp. 255–265, 2003.

[10] A. Mahimkar, T.S. Rappaport, "SecureDAV: a secure data aggregation and verification protocol for wireless sensor networks", in:Proceedings of the 47th IEEE Global TelecommunicationsConference (Globecom), November 29–December 3, Dallas, TX,2004.

[11] H. Çam, S. Ozdemir, P. Nair, D. Muthuavinashiappan, H.O. Sanli,"Energy-efficient and secure pattern based data aggregation for wireless sensor networks", Comput. Commun., Elsevier 29 (4) pp. 446–455, 2006.

[12] H.O. Sanli, S. Ozdemir, H. Çam, "SRDA: secure reference-based data aggregation protocol for wireless sensor networks", in: Proceedings of the IEEE VTC Fall Conference, Los Angeles, CA, 26–29, pp. 4650–4654, September 2004.

[13] Y. Yang, X. Wang, S. Zhu, G. Cao, "SDAP: a secure hop-by-hop data aggregation protocol for sensor networks", in: Proceedings of theACM MOBIHOC'06, 2006.

[14] S. Ozdemir, "Secure and reliable data aggregation for wireless sensor networks", in: H. Ichikawa et al. (Eds.), LNCS 4836, pp. 102–109, 2007.

[15] HemantSethi, Devendra Prasad, "EIRDA: An Energy Efficient Interest based Reliable Data Aggregation Protocol for Wireless Sensor Networks" International Journal of Computer Applications (0975 – 8887)Vol.22– No.7, pp.20-25, May 2011

[16] D. Westhoff, J. Girao, M. Acharya, “Concealed data aggregation for reverse multicast traffic in sensor networks: encryption key distribution and routing adaptation”, IEEE Trans. Mobile Comput. 5 (10) pp.1417–1431, 2006.
[17] D. Kumar, T.C. Aseri, R.B. Patel, “EECDA: Energy Efficient Clustering and Data Aggregation Protocol for Heterogeneous Wireless Sensor Networks”, in Int. J. of Computers, Communications & Control, Vol. VI, No. 1 (March), pp. 113-124, 2011.
[18] Luigi Alfredo Grieco, GennaroBoggia Alberto Coen-Porisini, “DyDAP: A Dynamic Data Aggregation Scheme for Privacy Aware Wireless Sensor Networks”, Journal of Systems and Software, Vol.85, Issue 1, pp.152-166. January 2012.