Secure Data Aggregation Techniques to Enhance Capabilities of Sensor Network

[1]Seema Dahiya, [2]Pawan Kumar Singh
[1]ECE,Phd Scholar,SRM University,[2]ECE, SRM University, Delhi-NCR

Abstract**—***Data Aggregation is a suitable technique to design an energy efficient Wireless Sensor Networks.The amount of data to be transmitted can be reduced by applying the Data aggregation algorithm. Moreover, it helps in reduction of the chance of data redundancy in design of Wireless Sensor computational Networks installed in hostile environment. In last decade,number of aggregation algorithms based on various techniques has been proposed. The most efficient data aggregation algorithms developed and published recently have been reviewedin this paper. The different categories of the data aggregation algorithms are structured, unstructured and hybrid. Since, the energy consumption of network is minimum with structured protocol, therefore, it can be used in static environments efficiently. Whereas, the Hybrid and Unstructured protocols can be used in dynamic approach. The paper also provides an introduction to in-network aggregation which help in reducing complexities of Wireless Sensor Networks.*

Keywords**—***Wireless Sensor Network, Data Aggregation,Structured Protocol, In-network Aggregation, Unstructured Protocol*

I. INTRODUCTION

A set of computational sensors in physical surroundings is known as Wireless Sensor Networks (WSN). These computational Sensors are having some limitations in terms of energy consumption, computing, Communication, storage capabilities etc. [6].The communication is considered to be most complicated and energy consuming process for any sensor node in WSN.A truthful solution for the aforesaid issue is less amount of data needed to be communicated. Commonly, a large number of extra computational nodes are installed to reduce the chances of packet loss during node failure [8]. A number of computational nodes among the extra nodes contain some redundant data which needed to be reduced using the aggregation techniques. These aggregation techniques are defined as Data Aggregation in which the internetwork processing is performed to collect data from various nodes [9]. The data is aggregated and forwarded to the designated station after collection from the various computational nodes.

Now a day, the Wireless Sensor Networks are being installed in various environments, where confidentiality is a major concern. The aforesaid areas are Telemedicine, border surveillance, emergency alarms, patient monitoring system and many more.

The existing computational nodes receive the packets and create a fixed size packet of its own. Aggregatequeries raise by the nodes for this purpose and the arithmetic operations (like SUM, MAX, etc.) perform by these Aggregate queries. The above mentioned solution is suitable for the small network only due to its energy consumption issues. Therefore, to optimize the energy consumption, the inter-network aggregation is applied to WSN. Then it is required to implement a better and efficient cosmopolitan algorithm in WSN for data aggregation. The algorithm need to fulfill the bellow mentioned requirement.

1. An optimal estimation of the information is expected from the algorithm. If any kind of the noise having the Gaussian distribution with zero mean, the variance must be closed to the CRLB (Cramer Rao Lower Bound).
2. Trustworthiness and the reliability in estimation of data received from the computational nodes, is required from the algorithm.

To evaluate reliability of computational nodes in wireless sensor networks trust & reputation systems play an important role. Evaluation of reliability at any point represents the aggregated behavior of computational nodes. Aggregation algorithms are targeted by attacker nodes to breach the security. One of the most effective security algorithms for wireless sensor networks is Trust & Reputation algorithms. Computational nodes present in hostile environments are prone to compromising attacks which leads to false injection of data.

Security and Data aggregation cannot treat together because they both have opposite goals. Actually for data aggregation we have to reduce the amount of data whereas while dealing with security concern we have to add some bits or hash function to provide security to data. In order to achieve both good aggregation & security in wireless sensor networks aggregation & security algorithms must be combined. Taking this idea in mind researchers proposed some algorithms which provides end to end security without using any cryptographic function like homomorphism algorithms.

The contributions of paper are listed below:-

1. Security Requirements of Wireless Sensor Networks
2. Overview of Internetwork Aggregation
3. Review and comparison of various data aggregation protocols

II. Security Requirement of Wireless Sensor Networks

1. *Data Confidentiality*

Confidentiality refers to privacy, steps are taken to prevent raw data from unidentified user. For this purpose various authentication algorithms are used for example biometric verification, key fobs, security tokens etc. The encryption is performed using various keying method. Homomorphic encryption is one of the most important encryption technique used in Wireless Sensors Networks[13] An Encryption algorithm is called homomorphic if it satisfies the following equation:-

$$D\left(E(x)\,\Delta_c\,E(y)\right) = D\left(E(x\Delta_m y)\right) \quad (1)$$

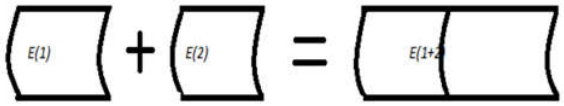

Fig. 1

To protect proprietary information and personal privacy, information access & disclosure rights are preserved for authorized nodes in a wireless sensor networks. Loss of confidentiality leads to unauthorized revelation of secret information. Data Confidentiality refers to secure access of data. Connection confidentiality in Wireless Sensor Network refers to confidentiality between nodes where a communication link is established between them. Whereas connectionless confidentiality in Wireless sensor networks refers to secure transmission between a small set of computational nodes. Selective-Field Confidentiality in wireless sensor networks to transmission of secured data within the range or reach of cluster. Traffic-Flow Confidentiality in wireless sensor networks refers to confidentiality between the path opted for transmission of message from source node to destination node through intermediate nodes. The purpose of confidentiality is to protect data from passive attack. The passive attacks used to scan open ports and vulnerabilities in a wireless sensor networks. Passive attacks are further classified two categories active & passive reconnaissance. In active reconnaissance the attacker node actively participates in network to steal or modify the information whereas passive reconnaissance here intruder monitors system without interacting with network. Several layers of protection can be deployed depending upon importance of data. To protect data to be transmitted over a connection Transmission Control Protocol (TCP) is implemented. Another aspect of confidentiality is to protect the data path from analyzing. So the attacker node will not able to identify source & destination node. Maintaining the Integrity of the Specifications.

2. *Data Integrity*

The stage helps in maintaining accuracy trustworthiness, accuracy and consistency of data throughout his entire life. The measures must be taken to ensure that the data is not altered by any of unauthorized node present. These steps are access control method like checksum etc.
These algorithms are used to protect packets from modification. Integrity is further classified into two parts Data & System Integrity. Data Integrity refers to the change in information is performed by an authorized node or not whereas system integrity refers to protect the system from unauthorized modifications. The main objective of integrity is to provide guarding against improper modification in a system or data. Loss of integrity refers to unauthorized modifications or loss of information. Connection Integrity with Recovery in wireless sensor network helps in protecting information or any modification between two nodes by an attacker node here if it detects that the data has been corrupted then the data can be recovered whereas in Connection integrity without recovery it is not possible to retrieve the data back if once modified. Selective-Field Connection integrity this provides integrity within the selected field of wireless sensor network. In connection less integrity with respect to wireless sensor networks the integrity is provided within small range of network. With confidentiality, integrity can be applied to a message or a queue of message or to a stream of bits. The most useful and easy approach to achieve integrity in the network is to apply integrity to all messages together. In order to remove redundancy of data connection oriented integrity technique is used. This techniques deals with both DoS and data modification.

3. *Data Authentication*

These algorithms are used to produce cryptographic Message Authentication Codes. Cryptographic hash functions are used to encrypt the data and provide security from attacks like collision attack, Birthday attack, Brute force Attack, etc.

These are based on cryptographic algorithms used to authenticate computational sensors. This service guarantees authentication service. The authentication service guarantees the receiver the message is from authenticated sender. To build a connection between two nodes in a wireless sensor networks both sender & receiver node must be authenticated. Secondly authentication algorithm must ensure that there is no third party attacker node between the paths. There are two authentication services generally used are Peer entity authentication & data origin authentication. Peer authentication services ensure that both the sender & receiver consist of same protocol whereas data origin authentication supports applications like email etc.

4. ***Data Availability***

This provides data continuity at a required level to improve the performance of Wireless Sensor Network. The data can be recovered during failure of nodes in mean time which could minutes, hours, days or months. Services ensure that any data or service is not denied to authorized users. This service ensures reliable access to use information. The more the system is complex the higher the level of availability required.

5. ***Data Freshness***

The resulted computation must be performed on most recent instance of the system. These ensures that these algorithms ensures that no old message can be retransmitted. These are of two types weak and strong freshness the weak freshness carries no delay whereas strong freshness carries some delay.

III. DATA AGGREGATION PROTOCOLS

Data aggregation protocols are divided into three categories Structured, Unstructured and Hybrid.

1. ***Structure data Aggregation Protocol***

Leach was the first structured data aggregation protocol where Time Division Multiple Access helps cluster heads to communicate with cluster members and Code Division Multiple Access is used to establish communication between cluster head and base station. These aggregation protocol do not consume any power while election cluster head and cluster members. Some of the structured algorithms are reviewed in this section.

1.1. Energy Efficient Clustering and Data aggregation

Energy efficiency and data aggregation are major concerns of this protocol. In this protocol single hop communication is performed within the cluster. The selection of cluster head is done by optimal probability algorithm. To elect cluster head energy of nodes are considered as election agent the node with higher energy is elected as cluster head. After election cluster head the algorithm elects the aggregated path fro transmission of data using MSRE (Maximum sum residual energy) Algorithm.

1.2. Dynamic and scalable tree Aware of spatial correlation

The algorithm is based on the concept of correlation where computational nodes are used to detect similar occurrence. For the purpose of aggregation the algorithm selects an representative computational node by applying spatial correlation procedure. The decision to elect number of representative computational nodes is based on unused energy. The purpose to implement this protocol is to reduce the cost of communication in a wireless sensor network. However it becomes very complex and costly while implementing it in a large network.

1.3. Energy efficient and balanced cluster-based data aggregation algorithm

The protocol deals with unbalanced dissipation of energy. For this reason the network is divided into unequal grid and the elected cluster head revolves in all grids of the network. The plan of dividing network into grids is based upon energy consumption. The protocol is also useful to extend the lifetime of computational node which solves a complex and major problem in wireless sensor network. The challenge for the protocol is to select the grid of appropriate size.

1.4. Delay Aware Network Structure for Wireless Sensor Network

The algorithm aware the wireless sensor network about delay in all cases. To perform communication with sink node clusters of different sizes are installed. The algorithm uses tree based network approach which guarantees minimal delay during aggregation. To communicate with fusion centers which are assumed to have infinite energy many single layer clusters are installed in the network.

1.5. Dynamical Message List based aggregation

The protocol works on the principal of dynamical message queue based on real time and cluster based aggregation. To protocol performs three major operations which includes activation of nodes, collecting the nodes to form a cluster and filtration of message. Each cluster head which is also known as filter node in this protocol consists of a dynamic list which keeps record of transmitted message and also verifies that the current message is transmitted earlier or not this reduces the redundancy and delay in packet delivering. Memory requirement is one of the major drawbacks of this protocol.

2. ***Unstructured Data Aggregation Protocol***

The purpose to propose this protocol is to increase efficiency of Wireless Sensor Network by minimizing communication cost, maintenance and queuing delay. DAA+RW (data-aware unicast and randomized waiting) is the first protocol introduced under this category. The node containing information sends an request to send (RTS) to identify next route. The priority of sending Clear to Send (CTS) is decided by shortest path algorithms. To increase the efficiency of wireless sensor networks some modifications are proposed explained below:-

2.1 Structure-free-Real time Data Aggregation Protocol

This protocol uses two procedures named as temporal and spatial occurrence which leads to data aggregation. this protocol is used to reduce the cost of maintenance for real time event. The detection of TTD of every packet helps in electing routing path. Based on real-time parameters like transmission, packet disputes, etc. time based stamping is used to calculate approximated EED and EHD. The selection of next hop in this protocol is based upon both data aggregation and real time.

2.2 Ant colony Algorithm

This algorithm is implemented at routing layer of wireless sensor network. The phases of algorithm are transmission, operations and initialization. The computational nodes are assumed to be artificial intelligent agents. To detect next hop energy estimation and shortest path algorithms are used. This algorithm, helps in increasing performance of Wireless Sensor Network.

2.3 Attribute Aware Data Aggregation Scheme

Using static routing it is not possible to aggregate assorted nodes. The aim of this scheme is to aggregate data transmitted by assorted nodes. This algorithm is advancement of ant colony algorithm. The packets are transferred on that path which have more similar type of packets. The major aim of the algorithm is to aggregate data transmitted by assorted nodes.

3. ***Hybrid Data Aggregation Protocols***

The main focus of these protocol is to aggregate data in an huge network with power efficiency and low delay. The hybrid algorithm is a combination of both structured and unstructured protocol.

3.1 Hybrid Energy Efficient Protocol

To increase efficiency of Wireless Sensor Network the protocol use both static and dynamic methods for aggregation of information. To perform aggregation the protocol uses both temporal and spatial based on the environment. In applications like event detection computational nodes are used o detect the events in environment. Depending upon the life time of packet the computational nodes wait for a random time. Due to bursty traffic because it depends upon waiting time.

3.2 Spatial correlation aware data aggregation protocol for data aggregation of Moving object in Wireless Sensor Networks

For applications of wireless sensor networks in observation field this protocol uses concept of spatial correlation for aggregation of data. This protocol is based on cluster communication and rate distortion. The aim of the protocol is used the traffic of network. The cluster head performs various complex computations to achieve its aim. To reduce number of computations the network is divided into grids.

Protocol	Published in	category	Aggregation	Design	Advantages	Limitations
EECDA	2011	structured	Electing cluster head	To minimize energy consumption	Increases lifespan of sensor networks	Not suitable for huge networks
YEAST	2011	structured	Scalable & dynamic routing	Reduces transmission of redundant data	Lesser energy consumption, accuracy	Memory requirments with higher complexities
EEBCDA	2012	structured	Grid formation and clusterhead rotates	Reduces amount of unbalanced energy dissipation	Increase in lifespan of networks with efficiency	Identification of cluster size is difficult
Delay aware network structure	2013	structured	Inter node communication	Reduce delay	Reduces delay	Static in nature
DMLDA	2015	structured	Dynamical messages lists	Reduces transmission delay	Efficient performance	Extensive memory
RAG	2011	unstructured	Temporal and satatical Convergence	On-time dilivery	Used t manage delivery time line	Large waiting
DAACA	2012	unstructur	Ant colony	Reduces consumption	Optimal path	High

		ed	optimization	of energy	election	complexities
ADA	2012	unstructured	adaptive and dynamic routing	Data aggregation	Energy efficient	High complexities
SFEB	2014	unstructured	Two phase aggregation	Reduces consumption of energy	Energy efficient	High delay
HEAP	2013	hybrid	Temporal &spatial cnvergence	Energy efficient aggregation	Energy effivient	Not suitable for brusty information
Spatial correlation aware protocol	2013	hybrid	Rate distortion based	Aggregation for moving objects	Energy efficiency	High computation overhead

IV. In-Network Aggragation

The in-network aggregation model is based on query model and system architecture. Brief outlines of these techniques are as follow:-

1. *System Architecture*

A set of computational nodes forming a wireless sensor network with multi-hop routing protocol used to provide communication link between two computational nodes. The system uses tree topology as a design of network. The sensor can be used as a node or aggregator depending upon need. To remove ambiguity we generally assumes that leaves are the source node and internal nodes are aggregators. Source node is used to monitor environmental conditions. Process of applying queries on the network is based on sink.

While performing computation in a Wireless sensor network the battery is depleted so these networks are resource constraints. To overcome this problem each sensor is synchronized loosely so they can mutually decide when to perform computation which leads to lower consumption of energy.

2. *Query Model*

In this mode a continuous query is registered at the source when the network is installed in setup phase or by broadcasting. A query may have following form

QUERY TEMPLATE
SELECT AGG (attr) FROM Sensors
WHERE pred EPOCH DURATION

The aggregate function AGG is processed on attr attribute of computational sensors, pred is predicate andT is the length of period. At every instance pushing query result to querior is responsibility of wireless sensor network. Whereas to collect results when required pull based approach is used. The aggregation functions are further divided into two parts distributive and non-distributive aggregation.

2.1 Distributive Aggregates. As we have earlier explained aggregation using SUM. Here each source generates the value and transmits it to the parent node in tree like structure and the partial sum is called PSR (Partial State Record).

2.2 Non Distributive Aggregates. The lack of distributive property in above method the nondistributive aggregate is not able to solve in a single route. The nondistributive aggregation takes multiple routes to perform computations. Based on formula given below:

$$O(\log |D|)$$

Here |D| is the domain size.

V. Conclusion

The main aim behind using data aggregation scheme is to reduce energy consumption and increasing life span of wireless sensor network. Because of trade of relationship of data aggregation schemes with delay, it is an important task to make an efficient Wireless Sensor network keeping energy consumption and delay minimum. The paper above reviewed some techniques to improve efficiency of Wireless Sensor Networks. However all the aggregation algorithms are considered with the aim to reduce energy consumption, Size of data, Communication Distance etc.

VI. References

[1] H. Hayouni, and M. Hamdi, "Secure Data Aggregation with Homomorphic Primitives in Wireless Sensor Networks: A critical Survey and open Research Issues" Proceedings of 2016 IEEE 13th International Conference on Networking, Sensing and control Mexico City, Mexico, April 28-30, 2016 .

[2] B. Bhushan, K.Kaushik,G.sahoo, "Concealed Data Aggregation with dynamic intrusion detection system to remove vulnerabilities in Wireless Sensor Networks" CNDC-2016 pp 81-96,2016 CS&IT-CSCP 2016 DOI-10.5121/csit.2016.60908.

[3] Shafi Patel et.al "Exploring Alternative Topologies for Network-on-Chip Architectures" BIJIT, 2011

[4] S. Siddiqui, A.A. Khan, and S. Ghani,"A survey on Data Aggregation Mechanisms in Wireless Sensor Networks".

[5] S.Verma, P. Pillai and Y.F. Hu"Performance Analysis of Data Aggregation and security in WSN-Satellite Integrated Networks ", 2013 IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Mobile and Wireless Networks.

[6] Dilip Kumar et. Al "EECHDA: Energy Efficient Clustering Hierarchy and Data Accumulation For Sensor" BIJIT

[7] S. Papadopoulos, A. Kiayias, and D.Papadias "Exact In-Network Aggregation with Integrity and Confidentiality", IEEE transactions on Knowledge and Data Engineering, Vol 24 No10 Oct 2012.

[8] L.F. Akyilidiz, W.su, Y. Sankarsubramaniam and E. Cayirci, "Wireless Sensor Networks: A survey", Elsevier Computer Networks, Vol.38, Issue 4,393-422,2002.

[9] D.Wanger,"Resilient aggregation in sensor networks", ACM Worksop on security of adhoc and sensor Networks 2004.

[10] M. & M.M. R. Esnaashari, "Data Aggregation in sensor networks using learning automata," Wireless Networks, vol. 16 no. 3, pp. 687-699, 2010

[11] R. Rajagopalan and P.K. Varshney,"Data Aggregation techniques in sensor networks: a Survey," 2006.

[12] S.Madden, M.J. Franklin, J.Hellerstein, and W. Hong,"TAG: A Tiny Aggregation Service for Ad-Hoc Sensor Networks," Proc. Fifth symp. Operating Systems Design and Implementation(OSDI) 2002.

[13] Y. Yao and J. Gehrke,"The COUGAR Approach to In-network Query Processing in Sensor Networks," ACM SIGMOD Record, vol. 31 no. 3 pp9-18, 2002.

[14] L.Hu and D. Evans ,"Secure aggregation for wireless Networks,"Proc. Symp. Applications and the internet Workshops (SAINT-W),2003.

[15] B.Patel and D.Jinwala, "Exploring Homomorphic Encryption in Wireless Sensor Networks," , Informatics Engineering and Information Science, Vol.251, 400-40, 2011

[16] L. Wassetrman, "All of Statics: A Concise Course in Statistical Interference", New York, NY, USA: Springer

[17] Mathioudakis,N.WhiteandN. Harris:"Wireless Sensor Networks: applications utilizing sattellite links", 18th IEEE International synopsim on Personal Indoor and Mocile Radio Communications Athenns Greece, 2007 .

[18] J. Yick , B. Mukherjee, and D. Ghoshal,"Wireless Sensor Network Survey", Computer Networks, vol.52, no. 12, pp. 2292-2330, August 2008.

[19] C.M.Chen, Y.H.Lin, Y.C.Lin, and H.M.Sun,"RCDA: Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks"IEEE Transactions on Parallel and Distributed Systems, Vol. 23, 2012