Secure Data Aggregation in Wireless Sensor Networks: A Survey

Charu Sharma
PhD Research Scholar
M.M. Engineering College
M.M. University ,
Mullana,Ambala, Haryana, India-133207
er.charusharma@mmumullana.org

Dr. Rohit Vaid
CSE Department
M.M. Engineering College
M.M. University ,
Mullana,Ambala, Haryana, India-133207
rohitvaid@mmumullana.org

***Abstract*—**Wireless Sensor Networks (WSNs) consists of millions of nodes which have partial battery life, computation and low storage. Interaction among nodes during data transfer consumes a large part of energy consumption. So data aggregation techniques are required, to minimize communication cost and energy consumption during data transfer. Various securities like freshness, data integrity, and confidentiality become compulsory in data aggregation when sensor node is organized in hostile environment. In this paper different secure data aggregation schemes for WSNs will be discussed.

Keywords*-WSNs, data aggregation, security.*

*****

I. INTRODUCTION

WSNs are resource constraints in terms of limited storage, memory, energy etc. Communication among nodes during data transfer consumes a large part of energy consumption.The performance of WSNs can not only be affected by resources but the deployment nature does also. As the sensor nodes are organized in hostile environments the nodes get more prone to physical attacks. So to minimize these problems data aggregation is used as it eliminates redundant data. Figure1 show how data aggregation works. The far away nodes send their data to aggregator nodes. The aggregator nodes collect data and combine its own data and send the data to next aggregator in the direction of BS. But when an aggregator node is compromised, it is very easy for enemy to change the result or insert false data into the network.

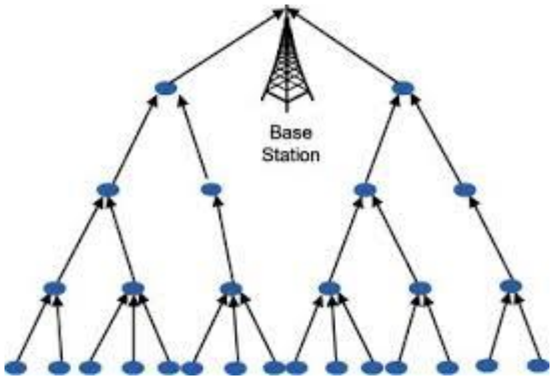

Figure1. Data Aggregation

Different challenges in WSNs are:

- *Resources constraints*: The sensor nodes have very low bandwidth, storage, processing capabilities etc. So these nodes should cooperate with each another to increase the network life time.
- *Unreliable Communication:* Unreliable transfer of packets in high density sensor network results in damage or loss of packets which is another risk to a sensor security.
- *Unattended operation:* As sensor nodes are managed remotely and have high exposure to physical attacks. As a result there is more probability that opponent has compromised the node if the node is left unattended for longer period in the network.
- *Scalable to large number of sensor nodes*: WSNs should be scalable to adapt addition of new nodes in the network to avoid frequent node failures.
- *To minimize reassemble cost:* It is very important to minimize the individual cost of each node to validate the overall cost of WSNs.
- *Latency:* Latency is defined as the delay time when the sender send the data until the receiver will successfully received that packet. As the nature of environment constantly changes in WSNs, it is necessary to receive the packet within a valid time period.
- *Lack of global identification:* The global positioning information of each node is not possible in large wireless sensor networks as it would increase high communication overhead of the entire network.

The remaining section of the paper is structured as follows: In II section, related works are explained. In Section III, classification of secure aggregation techniques is presented. In section IV, comparison of different aggregation schemes is shown and in Section V conclusions are made.

II. LITERATURE SURVEY

A comprehensive approach for research has been carried out in WSNs in recent past years. Zhang, and K. Leung(May 2010) surveyed that no schemes were present for secure data transfer in early years that could enable a better battery power and storage efficient framework .

WSNs produce a large quantity of data. So this data need to be aggregated in different levels in order to remove redundant data. For this different data aggregation approaches is considered in which node parameters such as bandwidth, time, power, signal strength should be considered.

After continuous research and development in this field, new solutions such as clustering can be used for proficient outputs M. Demirbas, and R. S. Aygun(2008). Due to high flexibility, low cost, Fault tolerance properties of sensor nodes generate many new application areas in remote sensing but some factors such as scalability, dynamic network topology, energy consumption should also be considered seriously during aggregation.

In order to minimize energy consumption, Zhang, and K. Leung(May 2010) proposed a data aggregation scheme which is extensively used. In this model the network is grouped into clusters and each cluster should have one representative node known as cluster head (CH) which combines all the data of the nodes within the cluster and send that data to the base station. In this approach only CH requires long distance transmission, so the energy of the remaining nodes in the clusters is conserved. The efficiency can be increased by minimizing cluster costs.

M P Singh, D K Singh(2010) proposed an effective mechanism for effective and efficient data aggregation by using the concepts like global weight calculation of nodes, using data cube aggregation techniques etc.

III. CLASIFICATION OF EXISTING SECURE DATA AGGREGATION SCHEMES

As shown in figure2 the data aggregation scheme is divided into two models-a) Single aggregator model b) multi-aggregator model.

a) One aggregator model

In this type of model data aggregation take place between the sensing nodes and the base station. The data from all nodes is collected on single node which is known as one aggregator point in the network. This aggregator should have the potential to perform high computation and long distance communications. The drawback of this model is that the redundant data still move in the network before reaching to the aggregator.Figure2-a represent this model. However this model is not suited for larger networks.

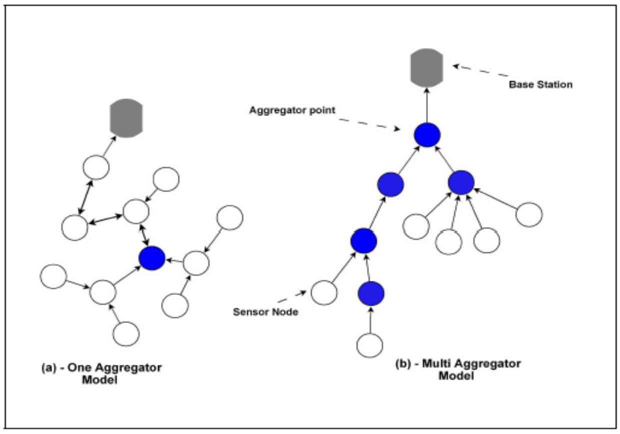

Figure 2: Single and Multi-aggregator model

The data aggregation schemes under this model are further categorize as-

With verification: It improves the aggregator ability to differentiate between the valid and invalid readings which are aggregated from different nodes.

- *Without Verification:* This is considered only when the scheme designers is not considered data integrity requirement.

b) Multi-aggregator model

As shown in Figure2-b in this model the data collected are aggregated more than one time before reaching to the base station. So this model is best suited when the network size is very large and when data redundancy is more in lowest level. This model is also further categorized as-

- *With verification:* It improves the aggregator ability to differentiate between the valid and invalid readings which are aggregated from different nodes. But this phase is more difficult as in this model the data is aggregated more than one time so the aggregator duty is to check whether the final aggregated results readings should not have very high variations when compared with intermediate results.
- *Without Verification:* This is considered only when the scheme designers is not considered data integrity requirement.

Figure3 shows the classification of existing secure data aggregation schemes based on one aggregator and multi aggregator model.

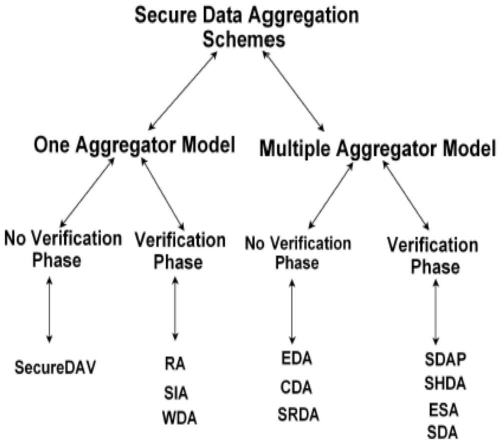

Figure3.Existing secure data aggregation schemes

IV. COMPRISION OF SECURE DATA AGGREGATION SCHEMES

In this segment different secure data aggregation schemes are compared on different factors such as cryptographic techniques used, security service provided, number of data transmitted to the aggregator etc.

A. EXISTING SCHEMES DESCRIPTION

SecureDAV Mahimkar & Rappaport(2004)scheme overcome the weakness present in ESA and SDA. This scheme provides data integrity. In this scheme one sensor node is elected as a aggregator node. All nodes within cluster send their data to the aggregator node. The aggregator node received the data from all the nodes within cluster, aggregate it and broadcast the average values of the reading to the base station if the average readings lie above the threshold value. This scheme has two disadvantages. Firstly, it requires high communication cost during data validation and secondly this scheme uses only AVG aggregation functions.

Yang et al(2006) presented a new protocol known as secure hop-by-hop data aggregation protocol(SDAP) which can bear the affect of more than one compromising node. This model is based on two principles-a) divide and conquer b) commit and attest. In order to minimise the harm caused by the compromise aggregator SDAP uses divide and conquer rule which divide the hierarchical network into multiple sub tree structure and hence increases the number of aggregators in the network and decreases the nodes in each subtree. But this scheme is only suited for hierarchical networks.

Chan et al.(2006) presented a new scheme in SIA which uses fully distributed networks in place of single aggregator model and apply aggregate -commit-prove framework. In this scheme each parent performs an aggregation functions when they get all the data from its chid nodes. The aggregation is done using Merkle hash tree. When base station receives data, it broadcasts it into the rest of the network so that every node in the network can check whether its contribution was aggregated or not. This scheme results in lot of communication over head.

A new Secure reference based data aggregation scheme (SRDA) is developed by Sanli et al.(2004) which sends only variation of sensed value and reference value in spite of sending raw data. In this scheme every sensor node calculate differential data which is calculated as sensed data value – reference value, encrypt it and send it to CH. To achieve high security during data transmission authors propose to use cryptographic algorithms with different adjustable parameters such as number of rounds, distance between CHs, amount of data etc. Any increase or decrease in number of rounds changes the security strength.

Domingo-Ferrer (2002) used privacy homomorphic (PH) encryption schemes which failed to provide reasonable security. Westhoff et al. (2006) highlighted the problem faced during the encryption of aggregated data. Authors in this paper proposed a new protocol known as Concealed data Aggregation (CDA), which uses additive homomorphic encryption scheme which allows the aggregator node to combine the data securely. But the drawback of this scheme is that it is very expensive and adds up to 22% communication overhead when compared with RC5. This scheme ensures only data confidentiality.

Castelluccia et al. (2005) presented a new scheme based on homomorphic encryption (EDA). This scheme allows an aggregator node to carry out aggregation functions and combine the encrypted data which it received from its child nodes. This scheme uses modular addition instead of xor. This scheme is highly secure because in this if aggregator is compromised, original data cannot be disclosed by the attacker. The drawback of this scheme is that it generates communication overhead and ensures only data confidentiality.

B. *SECURITY SERVICE PROVIDED*

As discussed in section A, it is concluded that each proposed scheme provides different security requirements. Table1 shows the comparison of different aggregation schemes based on security requirements.

Table1.Comparision between secure data aggregation schemes

Schemes	Integrity	Authentication	Freshness	Confidentiality
CDA				✗
SDA	✗	✗	✗	
SIA	✗	✗	✗	✗
SHDA	✗	✗	✗	
WDA	✗	✗		
SecureDAV	✗	✗		✗
SRDA			✗	✗
SDAP	✗	✗	✗	✗
ESA	✗	✗	✗	✗
EDA				✗

FromTable1. It is concluded that SRDA, CDA, EDA met the minimum requirements when the adversary have limited access to the network.SIA and SHDA meet the minimum security requirements when the adversary has low computational strength to begin an attack against the secure system. However secureDAV and WDA failed to meet the security requirements because they do not provide data freshness. SIA and SHDA have met all the requirements even when the adversary is very strong and has the capability to compromise any number of nodes in the network.

C. *Cryptographic primitives used in different schemes*

After comparing existing schemes it has been analyzed that different authors used different cryptographic primitives in their proposed schemes.Table2 shows how these Cryptographic primitives varies from model to model.

Table2: Cryptographic primitives used

SCHEMES	SYMMETRIC KEY	PUBLIC KEY	READINGS COMMITENT	INTERACTIVEPROTOCOL
CDA	✗			
SDA	✗			
SIA	✗		✗	✗
SHDA	✗		✗	✗
WDA	✗			

SecureDAV		✗	✗	
SRDA	✗			
SDAP	✗		✗	✗
ESA	✗			
EDA	✗			

V. *CONCLUSION ANF FUTUR SCOPE*

In this paper different data aggregation schemes have been reviewed on different classification along with their advantages and disadvantages. In future, it is planned to assess more data aggregation schemes on some other basis such as based on attacks against existing schemes, adversary model etc.

References

[1] Hu, L. & Evans, D. (2003), Secure aggregation for wireless network., in ‘SAINT Workshops’, IEEE Computer Society, pp. 384–394.

[2] Jadia, P. & Mathuria, A. (2004), Efficient secure aggregation in sensor networks., in L. Boug´e & V. K.Prasanna, eds, ‘HiPC’, Vol. 3296 of Lecture Notes in Computer Science, Springer, pp. 40–49.

[3] Przydatek, B., Song, D. X. & Perrig, A. (2003), SIA:Secure Information Aggregation in Sensor Networks., in I. F. Akyildiz, D. Estrin, D. E. Culler & M. B. Srivastava, eds, ‘SenSys’, ACM, pp. 255–265.

[4] Du, W., Deng, J., Han, Y. S. & Varshney, P. (2003), A witness-based approach for data fusion assurance in wireless sensor networks, in ‘IEEE Global Communications Conference (GLOBECOM)’, Vol. 3,pp. 1435– 1439.

[5] Mahimkar, A. & Rappaport, T. S. (2004), Secure-DAV: A secure data aggregation and verification protocol for sensor networks., in ‘Global Telecommunications Conference’, Vol. 4, pp. 2175– 2179.

[6] Yang, Y.,Wang, X., Zhu, S. & Cao, G. (2006), SDAP: a secure hop-by-hop data aggregation protocol for sensor networks., in ‘Proceedings of the 7th ACM Interational Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2006, Florence,Italy, May 22-25, 2006’, ACM, pp. 356–367.

[7] Chan, H., Perrig, A. & Song, D. (2006), Secure hierarchical in-network aggregation in sensor networks., in A. Juels, R. N. Wright & S. D. C. di Vimercatieds, ‘ACM Conference on Computer and Communications Security’, ACM, pp. 278–287.

[8] Sanli, H. O., Ozdemir, S. & Cam, H. (2004), SRDA: Secure reference-based data aggregation protocol for wireless sensor networks, in ‘Vehicular Technology Conference’, pp. 4650– 4654.

[9] Domingo-Ferrer, J. (2002), A provably secure additive and multiplicative privacy homomorphism., inH. Chan & V. D. Gligor, eds, ‘ISC’, Vol. 2433 of Lecture Notes in Computer Science, Springer,pp. 471–483.

[10] Wagner, D. (2003), Cryptanalysis of an algebraic privacy homomorphism., in C. Boyd & W. Mao, eds,‘ISC’, Vol. 2851 of Lecture Notes in Computer Science, Springer, pp. 234–239

[11] Castelluccia, C., Mykletun, E. & Tsudik, G. (2005), Efficient Aggregation of Encrypted Data in Wireless Sensor Networks., in ‘MobiQuitous’, IEEE Computer Society, pp. 109–117.

[12] E. Liu, Q. Zhang, and K. Leung(May 2010), “Residual Energy-Aware Cooperative Transmission (React) in Wireless Networks,” In Wireless and Optical Communications Conference (Wocc), pp. 1–6.

[13] K. Akkaya, M. Demirbas, and R. S. Aygun(2008), “The Impact of Data Aggregation on the Performance of Wireless Sensor Networks”, Wiley Wireless Communications and Mobile Computing (Wcmc) Journal,Vol. 8, pp. 171-193

[14] Shio Kumar Singh, M P Singh, D K Singh, “A Survey of Energy-Efficient Hierarchical Cluster-Based Routing in Wireless Sensor Network’s”, 570 Volume: 02, Issue: 02, pp: 570-580 (2010).