ISSN No. 0976-5697

Volume 3, No. 5, Sept-Oct 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Data Aggregation Security in Wireless Sensor Network

Pawan Kumar Goel*
Research Scholar, Department of CSE
Mewar University Rajasthan, India
erpawangoel@rediffmail.com

Dr. Vinit Kumar Sharma
Associate Professor Department of Mathematics & CSE,
SIET, Shamli, India
vksharmaxyz_1@rediffmail.com

***Abstract*:** Wireless Sensor Networks (WSNs) have been an attractive paradigm for pervasive computing oriented applications. Data aggregation is a widely used technique in wireless sensor networks. Data aggregation is the process of summarizing and combining sensor data in order to reduce the amount of data transmission in the network. The security issues, data confidentiality and integrity, in data aggregation become vital when the sensor network is deployed in a hostile environment. So data aggregation process is required which combines the data coming from various sensors, remove the redundancies in those data and then enroot them. But in hostile environment these aggregated data should be protected from several forms of attacks to achieve the security needs (like data confidentiality, data integrity and source authentication). The paper investigates the relationship between security and data aggregation process. In this paper general security issues in WSNs have been explored and we present an extensive study to provide a comprehensive review of the existing secure aggregation schemes for in-network aggregation in wireless sensor networks and analyze possible security threats on them.

***Keywords*:** Data aggregation, WSNs.

I. INTRODUCTION

Wireless sensor networks (WSN) consist of a great deal of sensor nodes with limited power, computation, storage, sensing and communication capabilities. [1] Sensors are becoming more and more inexpensive due to the advancement of the relevant technologies, so WSN will have broad applications in either controlled environments (such as home, office, warehouse, etc) or uncontrolled environments (such as hostile or disaster areas, toxic regions, etc). In these applications, the data collected by sensor nodes from their physical environment need to be assembled at a host computer or data sink for further analysis. Typically, an aggregate (or summarized) value is computed at the data sink by applying the corresponding aggregate function, e.g., MAX, COUNT, AVERAGE or MEDIAN to the collected data. In large sensor networks, computing aggregates in-network, i.e., combining partial results at intermediate nodes during message routing, significantly reduces the amount of communication and hence the energy consumed. In wireless sensor networks, the benefit of data aggregation increases if the intermediate sensor nodes perform data aggregation incrementally when data are being forwarded to the base station. However, while this continuous data aggregation operation improves the bandwidth and energy utilization, it may negatively affect other performance metrics such as delay, accuracy, fault-tolerance, and security.

II. DATA AGGREGATION

Data aggregation protocols aim to combine and summarize data packets of several sensor nodes so that amount of data transmission is reduced.[2] An example data aggregation scheme is presented in Fig. 1 where a group of sensor nodes collect information from a target region. When the base station queries the network, instead of sending each sensor node's data to base station, one of the sensor nodes, called data aggregator, collects the information from its neighboring nodes, aggregates them (e.g., computes the average), and sends the aggregated data to the base station over a multi-hop path.

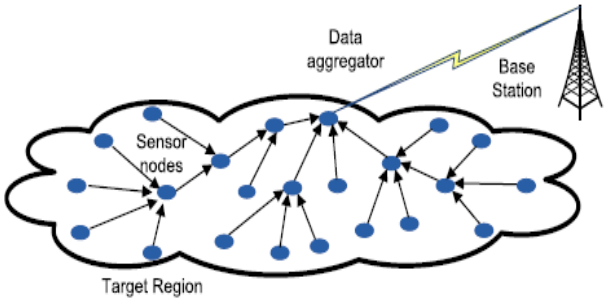

Figure. 1: Data aggregation in a wireless sensor network.

In wireless sensor networks, the benefit of data aggregation increases if the intermediate sensor nodes perform data aggregation incrementally when data are being forwarded to the base station. However, while this continuous data aggregation operation improves the bandwidth and energy utilization, it may negatively affect other performance metrics such as delay, accuracy, fault-tolerance, and security. [3][4]As the majority of wireless sensor network applications require a certain level of security, it is not possible to sacrifice security for data aggregation.

III. SECURITY ISSUES IN WSNS

A. Requirement of Data Aggregation Security:

The security requirements of a wireless sensor network can be classified as follows:[5]

a. ***Data Confidentiality****:* Ensures that information content is never revealed to anyone who is not authorized to receive it. It can be divided (in secure data aggregation schemes) into a hop by hop basis and an end-to-end basis. In the hop-by-hop basis, any aggregator point needs to decrypt the received encrypted data, apply some sort of aggregation function, encrypt the

aggregated data, and send it to the upper aggregator point. This kind of confidentiality implementation is not practical for the WSN since it requires extra computation.[5] On the other basis, the aggregator does not need to decrypt and encrypt data and instead of this, it needs to apply the aggregation functions directly on the encrypted data by using homomorphic encryption.

b. Data Integrity: ensures that the content of a message has not been altered, either maliciously or by accident, during transmission process. Confidentiality itself is not enough since an adversary is still able to change the data although it knows nothing about it. Suppose a secure data aggregation scheme focuses only on data confidentiality. An adversary near the aggregator point will be able to change the aggregated result sent to the base station by adding some fragments or manipulating the packet's content without detection. Moreover, even without the existence of an adversary, data might be damaged or lost due to the wireless environment.

c. Data Freshness: ensures that the data are recent and that no old messages have been replayed to protect data aggregation schemes against replay attacks. In this kind of attack, it is not enough that these schemes only focus on data confidentiality and integrity because a passive adversary is able to listen to even encrypted messages transmitted between sensor nodes can replay them later on and disrupt the data aggregation results. More importantly when the adversary can replay the distributed shared key and mislead the sensor about the current key.

d. Data Availability: ensures that the network is alive and that data are accessible. It is highly recommended in the presence of compromised nodes to achieve network degradation by eliminating these bad nodes. Once an attacker gets into the WSN by compromising a node, the attack will affect the network services and data availability especially in those parts of the network where the attack has been launched. Moreover, the data aggregation security requirements should be carefully implemented to avoid extra energy consumption. If no more energy is left, the data will no longer be available. When the adversary is getting stronger, it is necessary that a secure data aggregation scheme contains some of the following mechanisms to ensure reasonable level of data availability in the network:

a) Self-healing: that can diagnose, and react to the attacker's activities especially when he gets into the network and then start corrective actions based on defined policies to recover the network or a node.

b) Aggregator Rotation: that rotates the aggregation duties between honest nodes to balance the energy consumption in WSN.

***e.* Authentication:** There are two types of authentication; entity authentication, and data authentication.

Entity authentication allows the receiver to verify if the message is sent by the claimed sender or not. Therefore, by applying authentication in the WSNs, an adversary will not be able to participate and inject data into the network unless it has valid authentication keys. On the other hand, data authentication guarantees that the reported data is the same as the original one. In a secure data aggregation, both entity and data authentication are important since entity authentication ensures that some exchanged data between sensors. For instance, electing an aggregator point or reporting invalid aggregated results are authenticated using their identity while data authentication ensures that raw data are received at the aggregators at the same time as they are being sensed.

f. Non-repudiation: ensures that a transferred packet has been sent and received by the person claiming to have sent and received the packet. In secure aggregation schemes, once the aggregator sends the aggregation results, it should not be able to deny sending them. This gives the base station the opportunity to determine what causes the changes in the aggregation results.

g. Data Accuracy: One major outcome of any aggregation scheme is to provide an aggregated data as accurately as possible since it is worth nothing to reduce the number of bits in the aggregated data but with very low data accuracy. A trade-off between data accuracy and aggregated data size should be considered at the design stage because higher accuracy requires sending more bits and thus needs more power.

h. Secure Localization*:* The sensor network often needs location information accurately and automatically. However, an attacker can easily manipulate non secured location information by reporting false signal strengths and replaying signals, etc.

B. Classes of Security Attacks:

Attacks on the computer system or network can be broadly classified as interruption, interception, modification and fabrication.

a. Interruption is an attack on network availability, for example physical capturing of nodes, insertion of malicious nodes.

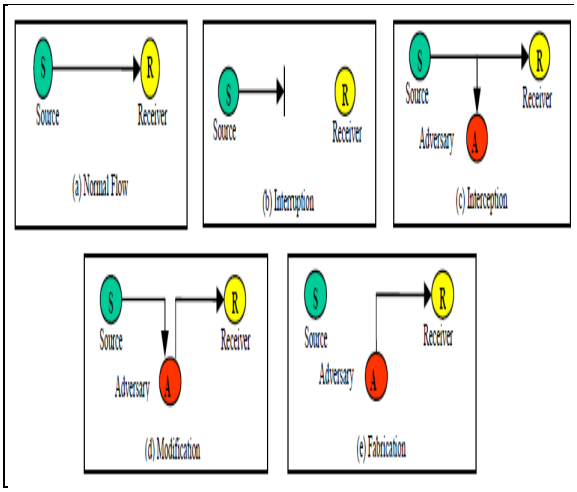

Figure.2 Security Attack Classes in WSN

b. Interception is an attack on confidentiality. Compromised node can gain unauthorized access to sensor nodes data.

c. Modification is an attack on integrity. Modification means an unauthorized party not only accesses the data but tampers it.

d. Fabrication is an attack on authentication. In fabrication, an adversary injects false data and compromises the trustworthiness of the information relayed.

IV. ATTACKS ON WSN AGGREGATION

A. Types of Attacks on WSN Aggregation:

WSNs are vulnerable to different types of attacks due to the nature of the transmission medium (broadcast), remote and hostile deployment location, and the lack of physical security in each node. However, the damage caused by these attacks varies from scheme to scheme according to the assumed adversarial model. In this section, these attacks that might affect the aggregation in the WSN are discussed.[5][6]

a. ***Denial of Service Attack(DoS):*** is a standard attack on the WSN by transmitting radio signals that interfere with the radio frequencies used by the WSN and is sometimes called jamming. As the adversary capability increases, it can affect larger portions of the network. In the aggregation context, an example of the DoS can be an aggregator that refuses to aggregate and prevents data from traveling into the higher levels.

b. ***Node Compromise:*** is where the adversary is able to reach any deployed sensor and extract the information stored on it which is sometimes called supervision attack. Considering the data aggregation scenario, once a node has been taken over, all the secret information stored on it can be extracted.

c. ***Sybil Attack:*** is where the attacker is able to present more than one identity within the network. It affects aggregation schemes in different ways. Firstly, an adversary may create multiple identities to generate additional votes in the aggregator election phase and select a malicious node to be the aggregator. Secondly, the aggregated result may be affected if the adversary is able to generate multiple entries with different readings. Thirdly, some schemes use witnesses to validate the aggregated data and the data is only valid if n out of m witnesses agreed on the aggregation results. However, an adversary can launch a Sybil attack and generate n or more witness identities to make the base station accept the aggregation results.

d. ***Selective Forwarding Attack:*** With no consideration about security, it is assumed in the WSN that each node will accurately forward received messages. However, a compromised node may refuse to do so. It is up to the adversary that is controlling the compromised node to either forward the received messages or not. In the aggregation context, any compromised intermediate nodes have the ability to launch the selective forwarding attack and this subsequently affects the aggregation results.

e. ***Replay Attack:*** In this case an attacker records some traffic from the network without even understanding its content and replays them later on to mislead the aggregator and consequently the aggregation results will be affected.

V. SECURE AGGREGATION SCHEMES

The resource constrained sensor nodes and necessity of plain data for aggregation process pose great challenges when implementing security and data aggregation together. This section attempts to describe the secure data aggregation schemes.[4]

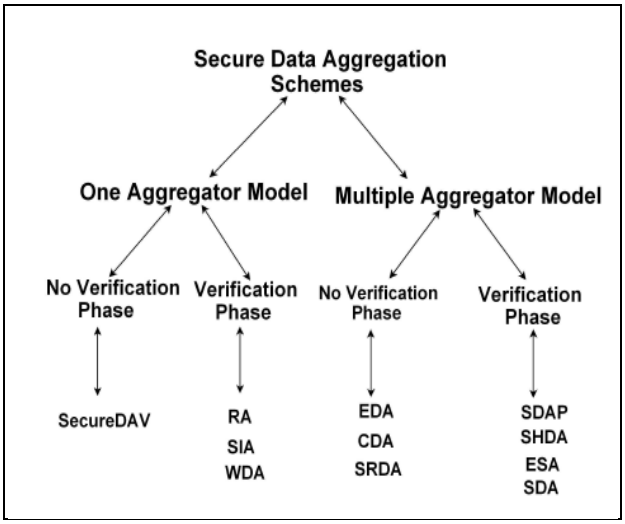

Figure 3: Classification of Existing Secure Data Aggregation Schemes.

a. The first secure data aggregation (SDA) was proposed by Hu & Evans (2003) who studied the problem of data aggregation once one node is compromised. This protocol achieves resilience against a node compromise by delaying the aggregation and authentication at the upper levels. Therefore, sensors measurements are forwarded unchanged and then aggregated at the second hop instead of aggregating them at the immediate next hop. Thus, the sensor needs to buffer the data to authenticate it once the shared key is revealed by the base station. Moreover, the proposed scheme only offers data integrity, freshness and authentication. Even though it increases the confidence in the sensor readings integrity the data can be altered once a parent and child in the hierarchy are compromised. Once a compromised node is detected, no practical action is taken to reduce the damage caused by this compromise which affects the data availability in the network. Much worse, once a grandfather node detects a node compromise, it could not decide whether the cheating node is the child or the grandchild. [7][8][9][10][11][12].

b. SDA scheme is improved in ESA by Jadia & Mathuria (2004). Instead of using µTESLA to authenticate the base station's broadcast in the validation process to reveal the shared key with sensors, the authors used one-hop pair-wise keys (to encrypt data between a node and its parent) and two-hop pair-wise keys (to encrypt data between a node and its grandparent). This will improve the secure aggregation scheme by adding data confidentiality and reducing the memory overhead since data does not need to be stored until the key is revealed.

c. Przydatek et al. (2003) proposed a secure information aggregation (SIA) framework for WSNs called aggregate-commit-prove. This framework provides resistance against a special type of attack called stealthy attacks aggregate manipulation where the attacker's goal is to make the user accept false aggregation results without revealing its presence to the user. It consists of three node categories: a home server, a base station, and sensor nodes. SIA assumes that each sensor has a unique identifier and shares a separate secret cryptographic key with both the home server and the aggregator. The keys enable message authentication and encryption if data confidentiality is required.SIA consists of three parts: collecting data from sensors and

locally computing the aggregation result, committing to the collected data, and reporting the aggregation result while proving the correctness of the result. SIA offers data integrity, authentication, data freshness, and confidentiality (if required). A witness based data aggregation (WDA) scheme for the WSN is being proposed by Du et al. (2003) to assure the validation of the data sent from an aggregator node to the base station. In order to prove the validity of the aggregated result, the aggregator node has to provide proofs from several witnesses. A witness is one who also performs data aggregation like the aggregator node, but does not forward its result to the base station. Instead, each witness computes the message authentication code (MAC) of the result and then sends it to the aggregator node which must forward the proofs to the base station.

d. Moreover, Secure DAV (Mahimkar & Rappaport 2004) improved the data integrity vulnerability in SDA and ESA by signing the aggregated data. In Secure DAV, each sensor within a cluster will have its share of its secret cluster key and then it will be able to generate a partial signature on the aggregated data. Once an aggregator receives sensor readings in the same cluster, it aggregates them and broadcasts the average value of the readings. Each sensor in the cluster compares its reading with the average value received from the aggregator. Then, it partially signs the average value only and only if the difference between the received average value and its reading is less than a certain value (threshold). Then, the aggregator (cluster-head) combines partial signatures to form a full signature of the aggregated results and sends it to the base station.

e. Yang et al. (2006) proposed a secure hop-by hopdata aggregation protocol (SDAP) that can tolerate more than one compromised node. SDAP is based on two principles: divide-and conquer and commit-and-attest. In order to reduce the damage caused by compromising an aggregator at a high level in the per-hop aggregation scheme, SDAP uses the divide-and conquer principle to divide the network tree into multiple logical sub-trees which increases the number of aggregators and reduces the number of nodes in each subtree. Consequently, the damage caused by compromising an aggregator of a subtree is reduced. The other principle, that is commit-and-attest, enhances the ordinary hop-by hop aggregation scheme by adding a commitment property, and helps the base station to prove the correctness of the aggregated data.

f. Furthermore, Chan et al. (2006) extended the work in SIA by applying the aggregate-commitprove framework in fully a distributed network instead of single aggregator model. In general, this scheme (SHDA) offers exactly what the SIA does data integrity, authentication, and confidentiality. Each parent sensor performs an aggregation function whenever it has heard from its child nodes. In addition, it has to create a commitment to the set of the input used to compute the aggregated result by using a merkle hash tree. Then, it forwards the aggregated data and the commitment to its parent until it reaches the base station. Once the base station received the final commitment values, it rebroadcasts them into the rest of the network in an authenticated broadcast. Each node is responsible for checking whether its contribution was added to the aggregated data or not.

g. Sanli et al. (2004) developed a new data aggregation technique called the Secure Reference-Based Data Aggregation scheme (SRDA) that sends only the difference between sensed data and the reference value (called differential value) instead of raw data. Deference value is taken as the average value of previous sensor readings. In SRDA scheme, each sensor computes the differential data (sensed data - reference value), encrypts it, and then sends it to the cluster-head.

h. a new algorithm using homomorphic encryption and additive digital signatures to achieve confidentiality, integrity and availability for in network aggregation in wireless sensor networks proposed by (Julia Albath and Sanjay Madria in 2008) they prove that prove that our digital signature algorithm which is based on the Elliptic Curve Digital Signature Algorithm (ECDSA).

i. Moreover, the problem of aggregating encrypted data in the WSN is being addressed in (Westhoff et al. 2006). The proposed protocol, called Concealed Data Aggregation (CDA), uses an additive and multiplicative homomorphic encryption scheme that allows the aggregator to aggregate encrypted data.

j. Furthermore, a new secure data aggregation scheme based on homomorphic encryption (EDA) is proposed by (Castelluccia et al. 2005) This allows an aggregator to execute the aggregation function and aggregate the encrypted data that are received from its children with no need for decryption and to recover the original messages. It uses a modular addition instead of the xor (Exclusive-OR) operation that is found in the stream ciphers.

VI. CONCLUSION

This paper provides a detailed review of secure data aggregation concept in wireless sensor networks. To give the motivation behind secure data aggregation, first, the security requirements and security classes of wireless sensor networks, security attacks on WSN aggregation are presented second, an extensive literature survey is presented by summarizing the data aggregation schemes. There are still open issues with WSN security requirements which enforce security for duplicate sensitive aggregation functions during data aggregation process.

VII. REFERENCES

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

[2] B.Krishnamachari, D. Estrin, S. Wicker, The impact of data aggregation in wireless sensor network, in: Proceeding of the 22nd International Conference on Distributed Computing Systems Workshops, 2002, pp.575-578.

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

[4] Hani Alzaid, Ernest Foo, Juan Gonzalez Nieto,Secure Data Aggregation in Wireless Sensor Network: A Survey,in: Proceeding of the 6th Australasian Information Security Conference (AISC 2008), Wollongong, Australia, 2008

[5] P.Mohanty,S.Panigrahi, N. Sharma, Security Issues in WSN: A survey, Journal of Theoretical and Applied Information Technology , pp. 14-27.

[6] Adrian Perrig and John Stankovic and David Wagner, Security in wireless sensor networks," Communications of the ACM, vol. 47, pp. 53-57, 2004.

[7] Widmer, M. Zorzi, In-network aggregation techniques for wireless sensor networks: a survey, IEEE Wireless Commun. 14 (2) (2007) 70–87.

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

[9] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, Directed diffusion for wireless sensor networking, in: IEEE/ACM Transactions on Networking, vol. 11, 2003, pp. 2–16.

[10] B. Krishnamachari, D. Estrin, S. Wicker, The impact of data aggregation in wireless sensor networks, in: Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops, 2002, pp. 575–578.

[11] 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 Vimercati,eds, 'ACM Conference on Computer and Communications Security', ACM, pp. 278–287.

[12] YI YANG, XINRAN WANG, SENCUN ZHU, and GUOHONG CAO, SDAP: A Secure Hop-by-Hop Data Aggregation Protocol for Sensor Networks, ACM Transactions on Information and Systems Security, Vol. 11, No. 4, Article 18, Pub. date: July 2008, pp: 18: 1 – 18: 42.