Volume 5, Issue 1, January 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

An Overview of Secure Data Aggregation in Wireless Sensor Network

Kaushal J. Patel*
M.E.Student
Information Technology Department
GCET, V V Nagar. Affiliated to G.T.U
Gujarat, India

Nirav M. Raja
Assistant Professor
Information Technology Department
GCET, V V Nagar. Affiliated to G.T.U
Gujarat, India

Abstract— Data Aggregation is widely used technique in wireless sensor networks (WSNs). The Security become a major issue, data confidentiality and data integrity in data aggregation. data aggregation should be incorporated in order to save energy. In this paper, the secure data aggregation schemes are categorized into hop by hop aggregation and end to end aggregation, and then the secure data aggregation schemes are reviewed and analyzed based on four phases: bootstrapping, data aggregation, verification and remedy. Data aggregation could save energy and bandwidth of the networks. end to end secure data aggregation can guarantee end-to-end confidentiality and privacy.

Keywords— Data aggregation, wireless sensor networks, data confidentiality, data integrity, Homomorphic Encryption.

I. INTRODUCTION

A wireless sensor network (WSN) consists of a large number of tiny sensor nodes deployed over a geographical area also referred as sensing field.[1] Each node is a low-power device that integrates computing, wireless communication and sensing abilities. These sensors have limited lifetime since they are battery-powered. They also have short communication range. Examples of WSN-based applications include environmental monitoring, flood detection, and Army Surveillance.[1]The main idea behind developing WSNs is to continuously collect data through the sensors.[1] The collected data is then sent to base stations to get analyzed, and accordingly, application based decisions are taken. Due to the limited lifetime of sensors and their short communication range, developing WSNs faces a number of challenges.[2]

Data aggregation algorithms for WSN define the techniques for gathering data from sensors, and deciding which data that need to be sent to the base station and when is the best time to send this data.[2] The need for good data aggregation algorithms is required to reduce the number and size of packet transmission in order to save the sensors energy, and to reduce the transmission of redundant data.[1][2]

In this paper, we will consider the security issues in the data aggregation of WSN. Specifically, the fundamental security issue is *data confidentiality* which protects the sensitive transmitted data from passive attacks, such as eavesdropping.[5] Data confidentiality is especially vital in a hostile environment, where the wireless channel is vulnerable to eavesdropping. The other security issue is *data integrity*, which prevents the compromised source nodes or aggregator nodes from significantly altering the final aggregation value.[5][8] A compromised node can modify, forge or discard messages. We categorize the secure data aggregation schemes into hop by hop aggregation and end to end aggregation, and then review and analyse the secure data aggregation schemes based on four phases: bootstrapping, data aggregation, verification and remedy. The fourth phase, remedy phase, is important and necessary.[4] [5]

In hop by hop data aggregation sensor nodes sensed data and encrypted this data and send to aggregator node. Aggregator node decrypt the data and use aggregation function and aggregate data from various node then encrypt it again and send to Base station. In end to end data aggregation sensor node sensed data and encrypted this data then send to the aggregation node. The aggregator directly aggregated the encrypted sensor data without the encryption keys of each sensor nodes.[6] sum, average, median, min, max are generally used as a aggregation function but in end to end secure data aggregation Homographic Encryption apply on aggregator node. This scheme used additive or multiplicative privacy homographic aggregation function and maintains end to end data confidentiality and data integrity.[9]

This paper is as follows, In Section II, the types of attacks in wireless sensor networks are presented. The secure data aggregation requirements are described in Section III. In the Section IV, we briefly address the four phases of secure data aggregation protocols. Then we categorize the secure data aggregation schemes into hop by hop aggregation and end to end aggregation.

II. TYPES OF ATTACKS IN WSN

A. *Denial of Service*

Denial of Service (DoS) is produced by the unintentional failure of nodes or malicious action. In wireless sensor networks, several types of DoS attacks in different layers might be performed.

B. Stealthy Attack

It is an intelligent attack to disinform a sensor network in a manner that to escape from attack discovery; that is, the BS accepts false sensors aggregation data. There are mainly two kinds of schemes to prevent stealthy attack. [5] One is passive, which protects all the sensor data and guarantee data confidentiality. The other one is active, which actively detect all the sensor data and make sure all the sensor data falls into certain range.[5]

C. Sybil Attack

The attacker disguises itself as a valid sensor node in the network, and normally the single invalid node could play more than one role during data aggregation process, so that it could attack the network from a large extend without being detected.[4][5]

D. Replay Attack

Replay attack means the attacker repeatedly send the previous sensing information, which affects the freshness of sensor data. Also the BS cannot get the most recent information from each sensor node.[5]

III. SECURE DATA AGGREGATION REQUIREMENTS

Data Confidentiality: Confidentiality is the ability to conceal messages from a passive attacker so that any message communicated via the sensor network remains confidential. This is the most important issue in network security. A sensor node should not reveal its data to the neighbours.

Data Authentication: Authentication ensures the reliability of the message by identifying its origin.

Data Integrity: Data Integrity in sensor networks is needed to ensure the reliability of data and refers to the ability to confirm that the message has not been tempered with, altered or changed. Even if the network has confidentiality measures there is still a possibility that the data integrity has been compromised by alterations. The integrity of the network will be in trouble when : **(1)** A malicious node present in the network injects false data. **(2)** Unstable conditions due to wireless channel cause damage or loss of data.

Data Availability: Availability determines whether a node has the ability to use the resources and whether the network is available for the messages to communicate. However, failure of the base station or cluster leader's availability will eventually threaten the entire sensor network. Thus availability is of primary importance for maintaining an operational network.

Data Freshness: Even if Data Confidentiality and Data Integrity are assured, there is a need to ensure the freshness of each message. Informally, data freshness suggests that the data is recent, and it ensures that no old messages have been replayed. To solve this problem a nonce or another time-related counter, can be added into the packet to ensure data.

IV. SECURE DATA AGGREGATION MECHANISM

1) key-based mechanism: The key-based mechanism in secure data aggregation is mainly used to guarantee data confidentiality. In most of hop by hop secure data aggregation protocols, pair-wise key is largely used due to the necessary of data encryption and decryption between each pair of communication nodes.[6] Also randomly generated key and key chain mechanism are used in some protocol design. It is more secure compared with the fixed key mechanism since it is difficult for the attacker to detect the regularity of the randomly generated key or the function which generates the key chain.[6] Privacy homomorphism (PH) is a special key based scheme. Most of the end to end secure data aggregation protocols are based on PH.[7][9][10][12] As it does not require intermediate node decrypts the encrypted data, it could directly aggregates encrypted data and recover the plain data when the BS get the final result.[6]

2) Signature based mechanism: The signature mechanism combined with key based mechanism is used to check the valid sensor nodes.[6]

3) Merkle hash tree based mechanism: Merkle Hash tree is used as a very important tool for data verification. [6]

4) MAC based mechanism: Message Authentication Code (MAC) is used as a tool of checking sensor node authentication. MAC combined with some sensor node features.[6]

V. SECURE AGGREGATION SCHEMES

In this section, all the protocols are analysed based on four phases which consist of a complete secure data aggregation protocol.

1) Bootstrapping phase: In wireless sensor networks, the bootstrapping phase is to set up the network and to distribute the keys for conducting intended tasks. There are mainly two kinds of structures for WSNs. The bootstrapping of encryption can be realized by two methods: 1) pair-wise key distribution among each pair of sensor nodes; 2) group-wise key distribution.[5][6]

2) Data aggregation phase: In this phase is to choose the proper routing and aggregate data securely. Data aggregation schemes can be classified into two groups based on data use in the aggregation process, one is hop by hop; the other is end to end. In the hop by hop schemes, the aggregator aggregates data after decryption, and then encrypted the aggregated data.[5][6] So normally the aggregator shares a key with the sensor nodes from which the data is transmitted. In the end to end schemes, the aggregator directly aggregated the encrypted sensor data without the encryption keys of each sensor nodes.[6]

3) Verification phase: In this Phase, both sensor data and sensor nodes validate. For sensor data validation, Merkle Hash tree and MAC base mechanism used and signature based is also best phase in secure data aggregation.[5][6]

4) Remedy/recovery phase: The remedy phase is another initialization phase. So most of the factors considered in the very first phase should also be considered here, such as re-build of sensor network structure; re-distribution of sharing keys and also totally clean out all the invalid information related to the compromised nodes.[6]

A. Hop by Hop Secure Data Aggregation

Hop by hop data aggregation protocol requires data encryption and decryption. Because of , most hop by hop data aggregation protocols have a relatively fixed network structures and aggregator nodes.[6] In addition, the multiple data decryption during data aggregation process which offers the attackers more chances to get the row sensor data and attack network.[6]

B. End to End Secure Data Aggregation

Compared to hop by hop data aggregation, end to end data aggregation relatively flexible structure and routing protocol. The aggregation functions are limited, as all the computation is done to the encrypted sensor data. [6] Most of the end to end data aggregation protocols do not provide verification and remedy phase, because the sensor data provide end to end confidentiality and integrity. This two security requirements provide end to end data privacy. [5][6]

C. Homomorphic Encryption

A homomorphic encryption scheme allows arithmetic operations on cipher texts. One example is a multiplicatively homomorphic scheme, where the decryption of the efficient manipulation of two cipher texts yields the multiplication of the two corresponding plaintexts. Homomorphic encryption Schemes are especially useful whenever some party not having the decryption key(s) needs to perform arithmetic operations on a set of cipher texts.[9][12] A more formal description of homomorphic encryptions schemes is as follows. Enc () denote a probabilistic encryption scheme and let M and C be its plaintext and cipher text spaces, respectively. If M is a group under operation $\oplus$, we say that Enc() is a $\oplus$-homomorphic encryption scheme, if, for any instance Enc() of the encryption scheme, given c1 = Enck1 (m1) and c2 = Enck2 (m2) for some m1, m2 $\in$ M, there exists an efficient algorithm that can generate—from c1 and c2—a valid cipher text c3 $\in$ C for some key k3 such that:

$$\mathbf{c3 = Enck3\ (m1 \oplus m2)}$$

In other words, decrypting c3 with k3 yields m1 $\oplus$m2. In this article, we mainly consider additive homomorphisms: $\oplus$ is the + operation.[6][9][12] we do not require k1, K2, k3, to be the same, although they need to be equal in most homomorphic encryption schemes.[9] Since k3 can be distinct from k1 and K2, some identifying information, (needs to be attached to the aggregated cipher text to identify the keys required for decryption.[9][12]

A homomorphic encryption scheme allows arithmetic operations on cipher texts.[6][9][12] Homomorphic encryption schemes allow aggregation on cipher text. One example is a multiplicative homomorphic scheme, where the decryption of the efficient manipulation of two cipher texts yields the multiplication of the two corresponding plaintexts [9]. In additive homomorphic encryption, we encrypt by adding a key to the data value, and we decrypt by subtracting a key from the aggregated value. An important property of the encryption and decryption functions is that they are commutative.[12] Homomorphic encryption schemes are especially
Useful whenever some party not having the decryption key(s) needs to perform arithmetic operations on a set of cipher texts.

Advantages of Homomorphic Encryption

1) Homomorphic Encryption makes it possible to give user a way to perform some operation on encrypted data without decryption key. So we are using the same property in this paper, the cluster head is doing aggregation operation on encrypted data.
2) Secret data is shared between several parties in such a way that no single party can retrieve the secret data.

Table 1 presents encryption policies, possible attacks, and vulnerabilities in data aggregation schemes.
First, we provide a lightweight data aggregation mechanism which protects data when data are processed in aggregators. Aggregators can help to eliminate redundant data without decrypting data. Thus, aggregators do not need to spend extra power in data decryption, and more network lifetime can be guaranteed. Second, our proposed scheme is resilient to known-plaintext attacks, chose plaintext attacks, cipher text-only attacks, and man-in the middle attacks.
In secure data aggregation require encryption technology in which directly performs data aggregation on encrypted data. CH directly perform data aggregation on received data and transform result to the sink so there is no need to decrypt data to perform aggregation again and encrypt and forward result to the sink.

TABLE I ENCRYPTION POLICIES, ATTACKS IN DATA AGGREGATION SCHEMES

Encryption Policy	Possible Attacks	Data Confidentiality	Data Integrity	Data Aggregation
Sensors transmit reading without encryption[15]	Eavesdropping Man-in-the-Middle	No	No	Generating wrong aggregated results

Sensors transmit encrypted readings with permanent keys[15]	Known-plaintext attack Chosen-plaintext attack Man-in-the-middle	Yes	No	Data Aggregation cannot be achieved
Sensors transmit encrypted readings with Privacy Homomorphism Scheme	None of Above	Yes	Yes	End to End Secure data Aggregation

VI. CONCLUSIONS

In this paper, we analysed the secure data aggregation schemes based on the four phases. We survey the related work for secure data aggregation in WSN and classify them into two cases: hop by hop and end to end encrypted data aggregation and also summarize the proposed techniques for protecting data confidentiality and data integrity. We also reviewed homomorphic encryption scheme. This scheme maintains this two security requirements and provide end to end privacy.

REFERENCES

[1] I. Akyildiz, W. Su, M. Vuran, and E. Cayirci, "A Survey On Sensor Networks", *IEEECommunications Magazine, Volume 40, Number, 2002.*

[2] P.N.Renjith, E. Baburaj, "An Analysis on Data Aggregation in wireless sensor networks", ICRCC © 2012.

[3] Chanjuan Wei, Yanjie Gao, J. Yang, "Cluster-Based Routing protocols in Wireless Sensor Network: A Survey", IEEE © 2011.

[4] Yan-Xiao Li, Lian-Qin, Qian-Liang, "Research On Wireless Sensor Network Security", IEEE © 2010.

[5] Yingpeng Sang, Hong Shen, "Secure Data Aggregation in Wireless Sensor Networks: A Survey", IEEE © 2006.

[6] Jia Guo, Jian'an Fang, Xuemin Chen, "Survey on Secure Data Aggregation for Wireless Sensor Networks", IEEE©2011.

[7] A.S.Poornima, B.B.Amberker, "SEEDA : Secure End-to-End Data Aggregation in Wireless Sensor Networks", IEEE © 2010.

[8] Hung-Min Sun, Chiung-Hsun Chen, and Po-Chi Li, "A Lightweight Secure Data Aggregation Protocol for Wireless Sensor Networks", IEEE © 2011.

[9] Juan Wei, Shanqing Guo, Qiuliang Xu, "Secure Homomorphic Aggregation Algorithm of Mixed Operations in Wireless Sensor Networks", IEEE © 2011.

[10] Suat Ozdemir, Yang Xiao, "Secure data aggregation in wireless sensor networks: A Comprehensive overview", ELSEVIER © 2009

[11] Shih-I Huang, Shiuhpyng Shieh, J. D. Tygar." Secure encrypted-data aggregation for wireless sensor networks".SpringerScience 2009.

[12] Neha Chhabra, Parikshit Singla "A Security Enhancing Homomorphic Encryption" Volume 2 Issue 7,July 2013.

[13] Sanjeev Setia, Sankardas roy and Sushil jajodia "Secure Data Aggregation in Wireless Sensor Networks" IEEE.

[14] Wenbo He, Hoang Nguyen, Xue Liu, Klara Nahrstedt, Tarek Abdelzaher. "SPDA: Secure and Privacy-preserving Data Aggregation in Wireless Sensor Networks".

[15] Jacques Bahi, Christophe Guyeux, Abdallah Makhoul. Secure Data Aggregation in Wireless Sensor Networks.

[16] Homomorphism versus Watermarking Approach. ADHOCNETS 2010, 2nd Int. Conf. on Ad Hoc Networks, Dec 2009, Canada.

