ISSN:2320-0790

An International Journal of Advanced Computer Technology

CDAMA: CONCEALED DATA AGGREGATION SCHEME FOR MULTIPLE APPLICATIONS IN WIRELESS SENSOR NETWORKS

[1]Lini Johnson, [2]Dr. S. Uma, [3]Mr. M. Mathan Kumar

[1]PG Scholar, Dept of CSE, [2]Prof. & Head, [3]Asst. Prof.
PG Department of Computer Science & Engineering
Hindusthan Institute of Technology, Coimbatore

Abstract: Generally, Data Mining is the process of analysing data from different perspectives and summarizing the information. Wireless sensor networks (WSNs) consist of thousands of sensor nodes (SN) that gather data from deployed environments. Efficient data aggregation scheme with security is needed to provide security in the wireless sensor networks as they may suffer unauthorized aggregation attacks. In previous homomorphic encryptions have been applied to conceal communication during aggregation such that enciphered data can be aggregated algebraically without decryption. It reduces a large amount of transmission and it is the most practical technique. Although data aggregation could significantly reduce transmission, it is vulnerable to some attacks. The objective of the proposed system is to design a scheme to detect these attacks or mitigate their impact. It can change the path in the intermediate level. So this will not get any aggregation problem and packet delivery issue.

Keywords: Wireless sensor networks, Sensor node, Concealed data aggregation for multiple applications, Cluster head, Aggregators, Elliptic curve method, Privacy homomorphic encryption.

1. INTRODUCTION

WIRELESS sensor networks (WSNs) consist of thousands of sensor nodes (SN) that gather data from deployed environments. Currently, there are plenty of rich applications proposed for WSNs, such as environment monitoring, accident reporting, and military investigation. Depending on the purpose of each application, SN is customized to read different kinds of data (e.g., temperature, light, or smoke). Typically, SN is restricted by the resources due to limited computational power and low battery supply; thus, energy saving technologies must be considered when we design the protocols. For better energy utilization, cluster-based WSNs have been proposed. In cluster-based WSNs, SN resident in nearby area would form a cluster and select one among them to be their cluster head (CH). The CH organizes data pieces received from SN into an aggregated result, and then forwards the result to the base station based on regular routing paths. Generally, aggregative operations are algebraic, such as the addition or multiplication of received data, or statistical operation, such as a median, a minimum, or a maximum of a data set. Although data aggregation could significantly reduce transmission, it is vulnerable to some attacks. For instance, compromising a CH will allow adversaries to forge aggregated results as similar as compromising all its cluster members. To solve this problem, several studies, such as the delay aggregation SIA, ESPDA, and SRDA, have been proposed. An alternative approach for this problem is to aggregate encrypted messages directly from SN, thereby avoiding the forgery of aggregated result. Since CHs are not capable of encrypting messages, compromising a CH earns nothing in aggregated results. Based on this concept, Wu et al.gave the proposal to allow CHs to classify encrypteddata without decrypting them. Following this concept, Westhoff et al. and Girao et al. proposed concealed data aggregation (CDA) supporting richer operations on aggregation. Unlike Wu et al.'s work, CDA utilizes the privacy homomorphism encryption (PH) to facilitate aggregation in encrypted data. By leveraging the additive and multiplicative homomorphism properties, CHs are able to execute algebraic operations on encrypted numeric data. Basically, CDAMA is a modification from Boneh et al.'s PH scheme. Here, it also supposes three practical application scenarios for CDAMA, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs. In practice, SN having different purposes, e.g., smoke alarms

and thermometer sensors may be deployed in the same environment. If we apply conventional concealed data aggregation schemes the ciphertexts of different applications cannot be aggregated together; otherwise, the decrypted aggregated result will be incorrect. The only solution is to aggregate the ciphertexts of different applications separately. As a result, the transmission cost grows as the number of the applications increases. By CDAMA, the ciphertexts from different applications can be encapsulated into "only" one ciphertext. Conversely, the base station can extract application-specific plaintexts via the corresponding secret keys. The second scenario is designed for single application WSNs. Compared with conventional schemes; CDAMA mitigates the impact of compromising SN through the construction of multiple groups. An adversary can forge data only in the compromised groups, not the whole system. The last scenario is designed for secure counting capability. In previous schemes, the base station does not know how many messages are aggregated from the decrypted aggregated result; leaking count knowledge will suffer maliciously selective aggregation and repeated aggregation. In CDAMA, the base station exactly knows the number of messages aggregated to avoid above attacks.

2. SYSTEM MODEL

Here, we state two models for further uses, aggregation model and attack model. The aggregation model defines how aggregation works; the attack model defines what kinds of attacks a secure data aggregation scheme should protect from.

2.1 Aggregation Model

In WSNs, SN collect information from deployed environments and forward the information back to base station (BS) via multihop transmission based on a tree or a cluster topology. The accumulated transmission carries large energy cost for intermediate nodes. To increase the life time, tree-based or cluster networks force the intermediate nodes (a subtree node or a cluster head) to perform aggregation, i.e., to be aggregators (AG). After aggregation done, AGs would forward the results to the next hop. In general, the data can be aggregated via algebraic operations (e.g., addition or multiplication) or statistical operations (e.g., median, minimum, maximum, or mean). For example, an AG can simply forward the sum of numerical datareceived instead of forwarding all data to the next hop.

2.2 Attack Model

First of all, we categorize the adversary's abilities as follows:

1. Adversaries can eavesdrop on transmission data in a WSN.
2. Adversaries can send forged data to any entities in a WSN (e.g., SN, AG, or BS).

3. Adversaries can compromise secrets in SNs or AGs through capturing them. Second, we define the following attacks to qualify the security strength of a CDA scheme. Part of these attacks refer to Peter et al.'s analysis Based on adversary's abilities and purposes, we further classify these attacks into three categories. In the first category A, an adversary wants to deduce the secret key (i.e., decrypting arbitrary cipher texts). Category A consists of four attacks that are commonly used in qualifying an encryption scheme. In practice, the first two attacks are feasible in WSNs. Here, we use them to qualify the underlying homomorphic encryption schemes. In categoryB, an adversary wants to send the forged messages to cheat the BS even though she does not know the secret key. This category consists of two attacking scenarios based on specific features deriving from PH schemes. The last category C consists of three attacks and considers the impact of node compromising attacks. The first attack is the case of compromising an AG, and the last two attacks are cases of compromising an SN. We discuss them separately because they store different secrets in the PH schemes.

A1. Ciphertext only attack. An adversary can deduce the key from only the encrypted messages.

A2. Known plaintext attack. Given some samples of plaintext and their ciphertext, an adversary can deducethe key or decrypt any ciphertext.

A3. Chosen plaintext attack. Given some samples of chosenplaintext and their ciphertext, an adversary can deduce thekey or decrypt any ciphertext.

A4. Chosen ciphertext attack. Given some samples of chosen ciphertext and their plaintext, an adversary can deduce the key or decrypt any ciphertext she has not chosen before. The model is CCA1, also called lunchtime attacks.

B1. Unauthorized aggregation. An adversary can aggregatesniffed ciphertexts into forged but format-valid ciphertexts.

B2. Malleability. An adversary can alter the content of aciphertext.

C1. B1/B2 under compromised AG. When an adversarycaptures an AG and compromises its secret, she can use it tolaunch B2/B3 attacks with higher probability of success.

C2. Unauthorized decryption under compromised SN. When an adversary captures an SN and compromises its secret, she can decrypt not only the ciphertexts from that SN but also the ciphertexts from the other remaining SNs.Asymmetric schemes can defend against unauthorizeddecryption under

compromised secrets because knowingthe public key is useless for decryption.

C3. Unauthorized encryption under compromised SN. Whenan adversary captures an SN and compromises its secret,she can impersonate not only that SN but also the otherremaining SNs to generate legal ciphertexts.

3. SYSTEM ARCHITECTURE DESIGN

Conventional hop-by-hop aggregation schemes are insecure because an adversary is able to forge aggregated results such as compromising all the AG's child nodes when he compromises the secret of an AG.CDAMA is designed by using multiple points, each of which has different order. It can obtain one scalar of the specific point through removing the effects of remaining points (i.e., multiplying the aggregated cipher text with the product of the orders of the remaining points). The security of CDAMA and BGN are based on the hardness assumption of subgroup decision problem, whereas CDAMA requires more precise secure analysis for parameter selections. In WSNs, SN collect information from deployed environments and forward the information back to base station (BS) via multihop transmission based on a tree or a cluster topology. The accumulated transmission carries large energy cost for intermediate nodes. To increase the lifetime, tree-based or cluster networks force the intermediate nodes (a subtree node or a cluster head) to perform aggregation, i.e., to be aggregators (AG). After aggregation done, AGs would forward the results to the next hop. In general, the data can be aggregated via algebraic operations (e.g., addition or multiplication) or statistical operations (e.g., median, minimum, maximum, or mean). For example, an AG can simply forward the sum of numerical data received instead of forwarding all data to the next hop. The AG aggregates those ciphertexts through modular addition. And the BS decrypts the cipher text received by modular subtraction with all the temporal keys. If an adversary tries to forge aggregated results, he must compromise all SNs. However, their scheme cannot prevent the adversary from injecting forged data packets into the legitimate data flow. In addition, key synchronization must be guaranteed because each SN must rekey after each encryption.

Each intermediate node can modify, forge or discard messages, or simply transmit false aggregation values, so one compromised node is able to significantly alter the final aggregation value. Further, aggregation interferes with message encryption. Encrypting messages using a unique key shared between each device and the base station since each intermediate node needs to understand the received messages to perform aggregation. Storing the same key on every device to enable encryption or authentication, since an adversary who recovers the key from a single device would then be able to control the entire network. The design is aiming at providing lightweight security mechanisms to effectively detect node misbehaviour (dropping, modifying or forging messages, transmitting false aggregate value). Thus enable a base station to trust results from a sensor network, even if an adversary may be able to deploy intruder nodes inside the network and recover the key material from a single node. By using this cluster head method, efficiently reduce the transmission. But it is cluster head method, compromising a cluster head will allow adversaries to forge aggregated results. Since CHs are not capable of encrypting messages, compromising a CH earns nothing in forging aggregated results. In addition to that to provide data confidentiality to the clients the database-as-a-Service model is used for the client has to secure their database through privacy homomorphism encryption (PH) schemes because PH schemes keep utilizable properties than standard ciphers.

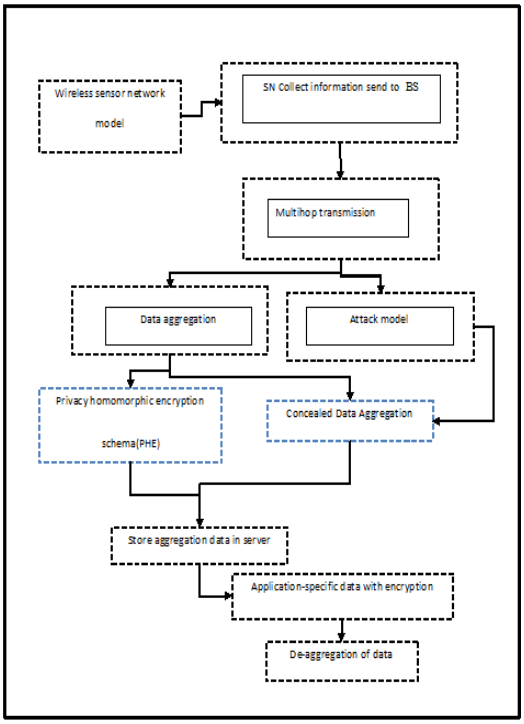

Figure3.1 Data Aggregation

3.1Generalization of CDAMA

CDAMA (k ¼ 2) can be generalized to CDAMA (k > 2). The paradigm of generalization uses different generators to construct different key pairs for groups. For security reason, the order of E should be

large enough. Therefore, when k becomes large, the length of cipher text will also expand.

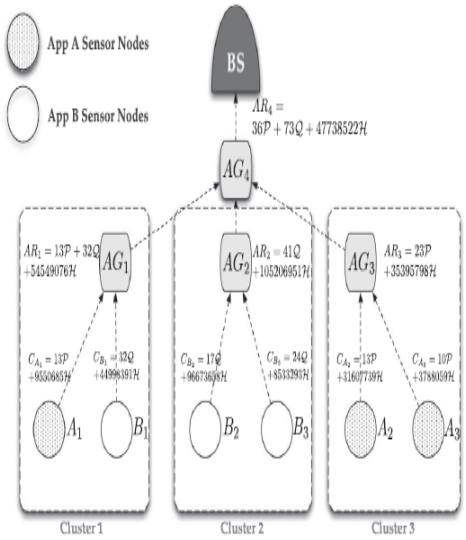

Figure3.2 CDAMA Generalization

Key pre distribution.

If we know the locations of deployedSNs, and can preload necessary keys and functions into SNsand AGs so that they can work correctly after being spread out over a geographical region.

Key post distribution.

Before SNs are deployed to their geographical region, they are capable of nothing about CDAMA keys. These SNs only load the key shared with the BS prior to their deployment, such as the individual key in LEAP and the master secret key in SPINS once these SNs are deployed, they can run the LEACH protocol to elect the AGs and construct clusters. After that, the BS sends the corresponding CDAMA keys, encrypted by the pre shared key, to SNs and AGs.

To maintain data privacy and reduce the communication overhead, sensed reading should been encrypted by SNs and the corresponding ciphertexts must be aggregated. The solution satisfying this requirement has already been proposed, called CDA. Even if aggregation on cipher texts is possible aggregation of multi-application is still hard because the decryption cannot extract application-specific aggregated result from a mixed cipher text.

REFERENCES

[1]. A. Perrig, J. Stankovic, and D. Wagner, "Security in Wireless Sensor Networks," Comm. ACM, vol. 47, no. 6, pp. 53-57, June 2004.

[2]. B. Przydatek, D. Song, and A. Perrig, "SIA: Secure Information Aggregation in Sensor Networks," Proc. First Int'l Conf. Embedded Networked Sensor Systems, pp. 255-265, 2003.

[3]. D. Westhoff, J. Girao, and M. Acharya, "Concealed Data Aggregation for Reverse Multicast Traffic in Sensor Networks: Encryption, Key Distribution, and Routing Adaptation," IEEE Trans. Mobile Computing, vol. 5, no. 10, pp. 1417-1431, Oct. 2006.

[4]. H. Sanli, S. Ozdemir, and H. Cam, "SRDA: Secure Reference-based Data Aggregation Protocol for Wireless Sensor Networks," Proc. IEEE 60th Vehicular Technology Conf. (VTC '04-fall), vol. 7, 2004.

[5]. I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," IEEE Comm. Magazine, vol. 40, no. 8, pp. 102-114, Aug. 2002.

[6]. L. Hu and D. Evans, "Secure Aggregation for Wireless Networks," Proc. Symp. Applications and the Internet Workshops, pp. 384-391, 2003.

[7]. R. Min and A. Chandrakasan, "Energy-Efficient Communication for Ad-Hoc Wireless Sensor Networks," Proc. Conf. Record of the 35th Asilomar Conf. Signals, Systems and Computers, vol. 1, 2001.

[8]. S. Peter, D. Westhoff, and C. Castelluccia, "A Survey on theEncryption of Converge cast-Traffic with In-Network Processing,"IEEE Trans. Dependable and Secure Computing, vol. 7, no. 1, pp. 20-34, Jan.-Mar. 2010.

[9]. S. Zhu, S. Setia, and S. Jajodia, "LEAP+: Efficient Security Mechanisms for Large-Scale Distributed Sensor Networks,"ACM Trans. Sensor Networks, vol. 2, no. 4, pp. 500-528, 2006.

[10]. Y. Wu, D. Ma, T. Li, and R.H. Deng, "Classify Encrypted Data in Wireless Sensor Networks," Proc. IEEE 60th Vehicular Technology Conf., pp. 3236-3239, 2004.