MULTIPLE SENSOR APPLICATIONS USING SECURE DATA AGGREGATION

Mr. S.M. Bharathi, Mr. D. Durai Kumar,
Department of Information Technology,
Ganadipathy Tulsi’s Jain Engineering College,
Kaniyambadi – 632 102.
bharathism31@gmail.com

ABSTRACT

A wireless sensor network, data aggregation scheme that reduces a large amount of transmission is the most practical technique. In earlier studies, Homomorphic encryptions have been applied to conceal communication during aggregation such that enciphered data can be aggregated algebraically without decryption. Since aggregators collect data without decryption, adversaries are not able to forge aggregated results by compromising them. However, these schemes are not satisfy multi-application environments. Second, these schemes become insecure in case some sensor nodes are compromised. Third, these schemes do not provide secure counting; thus, they may suffer unauthorized aggregation attacks. Therefore, I use a new concealed data aggregation scheme extended from Boneh et al.’s Homomorphic public encryption system. The proposed scheme has three contributions. First, it is designed for a multi-application environment. The base station extracts application-specific data from aggregated cipher texts. Next, it mitigates the impact of compromising attacks in single application environments. Finally, it degrades the damage from unauthorized aggregations. To prove the proposed scheme’s robustness and efficiency, I also include the timer which is used to give alert message to base station prevent from attackers.

Index Terms— Base stations, Concealed data aggregation, homomorphic encryption, wireless sensor networks

1. INTRODUCTION

Wireless sensor networks (WSNs) consist of thousands of sensor nodes (SN) that gather data from deployed environments. There are plenty of rich applications proposed for WSNs, such as environment monitoring, accident reporting, and military investigation [1]. Depending on the purpose of each application, SN are customized to read different kinds of data (e.g., temperature, light, or smoke). Typically, SN are 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 [2], 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 or statistical operation. 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 [4] as similar as compromising all its cluster members. To solve this problem, several studies, such as the delay aggregation have been proposed.

In this proposed scheme, called CDAMA, provides CDA between multiple groups. Basically, CDAMA is a modification from Boneh et al.’s [13] PH scheme Here, I also suppose three practical application scenarios for CDAMA, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA. The first scenario is designed for multi-application WSNs, all of which can be realized by only CDAMA.
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. In CDAMA, the base station exactly knows the number of messages aggregated to avoid above attacks.

2 SYSTEM MODEL

Here, I 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.

WSN Setup

In this phase I have to set up the WSN environment by designing base station, Aggregators and multiple sensors. The communications between them also have to establish. Group Public and Private Key established that keys are known by Application sensors and Base station.

Key Distribution

It briefly explains how to deliver the group public keys to SNs securely. There are two main approaches.

Key pre-distribution

If I know the locations of deployed Sensor nodes (SN), I can preload necessary keys and functions into SNs and Aggregators (AG) 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 and the master secret key. 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.

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 lifetime, tree-based or cluster networks force the intermediate nodes (a sub tree 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.

Attack Model

First of all, I 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, I define the following attacks to qualify the security strength of a CDA scheme. Part of these attacks refers to Peter et al.’s analysis [15]. Based on adversary’s abilities and purposes, I further classify these attacks into three categories. In the first category A, an adversary wants to deduce the secret key (i.e., decrypting arbitrary ciphertexts). 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 [15]. Here, I use them to qualify the underlying homomorphic encryption schemes. In category B, 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. I 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 deduce the key or decrypt any ciphertext.
A3. Chosen plaintext attack. Given some samples of chosen plaintext and their ciphertext, an adversary can deduce the key or decrypt any ciphertext.
A4. Chosen cipher text 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 aggregate sniffed ciphertexts into forged but format-valid ciphertexts.
C1. B1/B2 under compromised AG. When an
adversary captures an AG and compromises its secret, she can use it to launch 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 unauthorized decryption under compromised secrets because knowing the public key is useless for decryption.

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

3 PRELIMNARIES

Privacy Homomorphic Cryptosystem

Privacy homomorphic encryption (PH) is an encryption scheme with homomorphic property. The homomorphic property implies that algebraic operations on plaintexts can be executed by manipulating the corresponding ciphertexts; for instance, \( D_K(E_K(m_1) \odot E_K(m_2)) = m_1 \oplus m_2 \)

where \( E_K(.) \) is the encryption with key K, \( D_K(.) \) is the decryption with key K, and \( \odot \) and \( \oplus \) denote operations on ciphertexts and plaintexts, respectively. In general, operations and can be addition, multiplication, and so on. Similar to conventional encryption schemes, PH schemes are classified to symmetric cryptosystems, when the encryption and decryption keys are identical, or asymmetric cryptosystem when the two keys are different.

Symmetric PH schemes, Castelluccia et al.’s scheme, usually are more competitive in terms of efficiency than asymmetric schemes. The most notable asymmetric PH schemes are based on elliptic curve cryptography (ECC). Compared with RSA cryptosystems, ECC provides the same security with a shorter key size and shorter ciphertexts. A 160-bit ECC cryptosystem provides the same security as a 1,024-bit RSA cryptosystem. In energy-constraint WSNs, constructing PH via ECC is more efficient.

CDA Based on PH

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. To diminish this impact, PH schemes have been applied to WSNs [9], [10], [11], [12], [14]. By PH schemes, SNs encrypt their sensed readings and allows AGs to homomorphically aggregate their ciphertexts without decryption. Therefore, compromising AGs earns no advantage of forging aggregated results.

Westhoff et al. [9] and Girao et al. [10] proposed CDA based on symmetric PH to facilitate the aggregation of encrypted data. In contrast to symmetric PH construction, Myklebost et al. [11] adopted public-key-based PH to construct their systems, and Girao et al. [12] extended the ElGamal PH encryption to construct an aggregation scheme. In these schemes, because all SN in a network only share a common key for encryption [9], [10], [11], [12], an adversary can forge the aggregated results by simply compromising one SN.

To solve this problem, Castelluccia et al. [14] proposed an encryption scheme similar to one-time pad. In each transmission, individual SN generates temporary key from a pseudo random number generator (PRNG) and adds its messages with the key under modulation. The AG aggregates those ciphertexts through modular addition. And the BS decrypts the ciphertext 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.

BGN Scheme

In 2006, Boneh et al. [13] proposed a public-key PH scheme, which integrates the Paillier with the Okamoto-Uchiyama encryption schemes. I call it, BGN for simplicity. BGN provides additive and multiplicative homomorphism. Since the multiplicative property, based on the bilinear pairing [13], is much expensive and inefficient for SNs, we only utilize the additive homomorphism of BGN. It provides a possible application for BGN, data aggregation. Furthermore, I modify BGN to fit multigroup construction or stronger security and better applicability.
4 SYSTEM ARCHITECTURE

Bs -> Base Station
AG1, AG2, AG3 -> Sub Aggregator or Cluster Head (CH)
AG4 -> Main Aggregator

When the scenario of multiple applications working concurrently is more realistic in most cases. Study indicates that deploying multiple applications in a shared WSN can reduce the system cost and improve system flexibility. The reason is because an SN supports multiple applications and can be assigned to different applications dynamically. For instance, three different kinds of SNs, smoke detectors, temperature collectors, and light detectors, are deployed in the same building. Each room contains an AG and some SNs. A big challenge for the AGs, AG1 to AG4, is to aggregate the sensed readings from the different applications to a mixed aggregated result. Unfortunately, two limitations make the aggregation more difficult:

1. To maintain data privacy and reduce the communication overhead, sensed reading should be encrypted by SNs and the corresponding cipher texts must be aggregated. The solution satisfying this requirement has already been proposed, called CDA.
2. 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.

5 CDAMA

BGN is implemented by using two points of different orders so that the effect of one point can be removed by multiplying the aggregated ciphertext with the order of the point, and then the scalar of the other point can be obtained. Based on the same logic of BGN, CDAMA is designed by using multiple points, each of which has different order. I can obtain one scalar of the specific point through removing the effects of remaining points (i.e., multiplying the aggregated ciphertext 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, discussing. I use CDAMA (k = 2) to explain how it works in multiple groups.

CDAMA (k=2) Construction

Assume that all SNs are divided into two groups, GA and GB. CDAMA contains four procedures: Key generation, encryption, aggregation, and decryption, listing in Fig. 2. As we can see, CDAMA (k = 2) is implemented by using three points P, Q, and H whose orders are q1, q2, and q3, respectively. The scalars of the first two points carry the aggregated messages in GA and GB, respectively, and the scalar of the third point carries randomness for security. As shown in the DEC functions, by multiplying the aggregated ciphertext with q2 q3 (i.e., SK in GA), the scalar of the point P carrying the aggregated message in GA can be obtained. Similarly, by multiplying the aggregated cipher-text with q1 q3 (i.e., the SK in GB), the scalar of the point Q carrying the aggregated message in GB can be obtained. In this way, the encryptions of messages of two groups can be aggregated to a single ciphertext, but the aggregated message of each group can be obtained by decrypting the ciphertext with the corresponding SK.

Considering deployment, the private keys should be kept secret and only known by the BS. SNs in the same group shares the same public key and no other entities outside the group know the group public key. Another major change is the decryption procedure. By performing individual decryption, the BS extracts individual aggregated results of different groups from an aggregated ciphertext.
Multiple Sensor Applications Using Secure Data Aggregation

**Figure-2. Procedures of CDAMA (k =2).**

**A Concrete Example**

Now, I use an instance to describe how CDAMA (k =2) works. In WSN consists of six SNs and four AGs. After deployments, they form three clusters. Each SN belongs to either application A or B. Without loss of generality, sensors A1, A2, and A3 perform application A and keep the public key PKA = (n, E, P, H, and T). The others, B1, B2, and B3 keep PKB = (n, E, Q, H, T). Four aggregators, AG1 to AG4 are deployed to gather messages from their child nodes. To simplify the example, we set the order of P, Q, and H to small numbers. We assume that |q1| = |q2| = |q3| = 10, e.g., ord(P) = q1 = 521, ord(Q) = q2 = 523, ord(H) = q3 = 541. And we assume T = 128 and x = 3 such that the maximal sensed value in both applications is at most 42 (i.e., T_A = T_B = 42).

We assume the messages of these sensors are MA1 = 13, MA2 = 21, MA3 = 10, MB1 = 32, MB2 = 17, and MB3 = 24. They are encrypted to the corresponding ciphertexts. After the aggregation by the AGs, the BS receives the final aggregated result AR4 whose value is 36P + 73Q + 477, 385, 22H. The aggregated result in application A, MA = MA1 + MA2 + MA3 = 36 can be obtained by decrypting AR4 using SK_A in the following steps:

1. Compute q2q3 * AR4 = 282943 * (36P + 73Q + 477, 385, 22H) = 101, 859, 48P = 398P, where 521P = 523Q = 541H = \infty.
2. MA = log_p (q2q3 * AR4) = log_398P, where p = q2q3 (mod 521) and 521p = \infty. Since MA = log_398P, we infer that MA = 36 (mod 521).
3. Finally, through Pollard’s _ method, MA = 36 can be obtained by the BS.

Similarly, the BS can extract the aggregated result MB in application B by computing the discrete logarithm of q1q3 * AR4 to the base point −Q = q1q3 * Q.
6 APPLICATIONS

In this section, I propose three applications that are realized by only CDAMA multi group construction.

Multi-Application WSNs

Compared with the multi-application WSNs, the scenario of a single application is more commonly discussed in WSNs. However, the scenario of multiple applications working concurrently is more realistic in most cases. It indicates that deploying multiple applications in a shared WSN can reduce the system cost and improve system flexibility. The reason is because an SN supports multiple applications and can be assigned to different applications dynamically.

![Figure-3. A multi-application WSN example](image)

For instance, three different kinds of SNs, smoke detectors, temperature collectors, and light detectors, are deployed in the same building. Fig. 3 shows this typical case. Each room contains an AG and some SNs. A big challenge for the AGs, AG1 to AG4, is to aggregate the sensed readings from the different applications to a mixed aggregated result.

Conventional Aggregation Model with Multiple Groups

Interestingly, applying CDAMA to the conventional aggregation model can mitigate the impact from compromising attacks. In Fig. 4, all SNs are in the same application, e.g., fire alarm, but they can be arranged into two groups through CDAMA construction. Each group could be assigned a distinct group public key. Once an adversary compromised a SN in group A; it only reveals $P_KA$, not $P_KB$. Since the adversary can only forge messages in group A, not group B, the SNs in group B can still communicate safely.

![Figure-4. Two groups for a single application](image)

The ideal case is that CDAMA assigns every node for its own group, resulting in the strongest security CDAMA ever offered. However, this is impractical because the size of ciphertext becomes extremely large when we construct groups with a huge group number. Thus, assigning a reasonable number of groups for a single application not only keeps the overhead acceptable but also mitigates the impact of compromising attacks.

Aggregation with Secure Counting

The main weakness of asymmetric CDA schemes is that an AG can manipulate aggregated results without encryption capability. An AG is able to increase the value of aggregated result by aggregating the same ciphertext of sensed reading repeatedly, or decrease the value by selective aggregation. Since the BS does not know the exact number of ciphertexts aggregated, repeated or selective aggregation may happen. To avoid this problem, we adopt CDAMA ($k = 2$) scheme to provide secure counting for single application case, i.e., the BS exactly knows how many sensed readings are aggregated while it receives the final result.

Security Analysis and Comparison

In this section, I analyze the security of CDAMA and other conventional schemes. More specifically, we compare CDAMA with four well-known CDA schemes: CDA [9], [10], Castelluccia et al.’s scheme [14], Mykletun et al.’s scheme [11], and TinyPEDS [12]. In Mykletun et al.’s scheme, the authors applied several well-known public key PH schemes to WSNs. They recommended two schemes which are suitable for WSNs, EC-OU and EC EG. Since TinyPEDS [12] is the same as the EC-EG scheme[11], we chose TinyPEDS as a candidate. In addition to these four schemes, BGN—from which our proposed CDAMA is extended—is also analyzed. Consequently, we analyze CDA, Castelluccia et al.’s scheme, TinyPEDS, EC-OU, BGN, and CDAMA based on the attack model.
7 CONCLUSIONS

In this section, I designed multi-application environment, CDAMA is the first CDA scheme. Through CDAMA, the cipher texts from distinct applications can be aggregated, but not mixed. For a single-application environment, CDAMA is still more secure than other CDA schemes. When compromising attacks occur in WSNs, CDAMA mitigates the impact and reduces the damage to an acceptable condition. In above applications, CDAMA is the first CDA scheme that supports secure counting. The base station would know the exact number of messages aggregated, making selective or repeated aggregation attacks infeasible. Finally, the performance evaluation shows that CDAMA is applicable on WSNs while the number of groups or applications is not large.

In future, we wish to apply CDAMA to realize aggregation query in Database-As-a-Service (DAS) model. In DAS model, a client stores her database on an untrusted service provider. Therefore, the client has to secure their database through PH schemes because PH schemes keep utilizable properties than standard ciphers. Based on PH schemes, the provider can conduct aggregation queries without decryption.

REFERENCES


