Efficient Data Gathering and Aggregation for Multiple Applications in Wireless Sensor Networks

Mrs. M. Fathima Begum  
Assistant Professor  
C. Abdul Hakeem College of Engg. and Tech.

Mr. M. Abdul Naseer  
Assistant Professor  
C. Abdul Hakeem College of Engg. and Tech.

Mr. S. Abdul Kani  
Sr. Assistant Professor  
Indur Institute of Engg. and Tech.

Abstract
Data aggregation in wireless sensor networks refers to acquiring the sensed data from the sensors to the gateway node. It reduces the amount of power consumed during data transmission between the sensor nodes. Generally homomorphic encryptions have been applied to conceal communication during aggregation. Since enciphered data can be aggregated algebraically without decryption, here adversaries are able to forge aggregated results by compromising them. However, these schemes are not satisfying multi-application environments, provide insecure transmission and do not provide secure counting for unauthorized aggregation attacks. In this paper, we propose a new concealed data aggregation scheme extended from homomorphic privacy encryption system. The proposed scheme designed for a multi-application environment, mitigates the impact of compromising attacks in single application environments and also it can avoid the damage from unauthorized aggregations by the privacy homomorphic encryption scheme.

Keywords: Elliptic curve cryptography, cluster head, privacy homomorphic encryption, wireless sensor networks, Concealed data aggregation

I. INTRODUCTION

Wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a main location. The more modern networks are bi-directional, also enabling control of sensor activity. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control, machine health monitoring, and so on.

Sensor node (SN) is restricted by the resources due to limited computational power and low battery supply [1]; 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 aggregator(AG) and then forwards the result to the base station based on regular routing paths. After aggregation done, AGs would forward the results to the next hop. In general, the data can be aggregated via algebraic operations (addition or multiplication) or statistical operations (median, minimum, maximum, or mean).

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 [1], and SRDA [4] 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 forging aggregated results[2].

In this paper, the proposed scheme, called CDAMA, provides CDA [3] between multiple groups. In this scenario is designed for multi-application WSNs. In data transmission, the cipher texts of different applications cannot be aggregated together. Otherwise, the decrypted aggregated result will be incorrect. The only solution is to aggregate the cipher texts of different applications separately.
As a result, the transmission cost grows as the number of the applications increases. During data transmission in WSNs, the ciphertext of different application can be aggregated together by CDAMA. Here, CDAMA mitigates the impact of compromising SN through the construction of multiple groups. In CDAMA, the base station exactly knows the number of messages aggregated to avoid above attacks by providing the secure counting.

II. **Preliminaries**

A. **Related Work**

In recent years, Adversaries can eavesdrop on transmission data, Adversaries can send forged data to any entities (e.g., SN, AG, or BS) and compromise secrets in SNs or AGs in a WSN[5],[6]. These can be solved by proposed scheme Privacy homomorphic cryptosystem, and secure count.

B. **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 [7]; PH schemes are classified to symmetric cryptosystem when the encryption and decryption keys are identical, or asymmetric cryptosystem (also called public key cryptosystem) when the two keys are different. Symmetric PH schemes [8], are more competitive in terms of efficiency than asymmetric schemes. The most notable asymmetric PH schemes are based on elliptic curve cryptography (ECC). ECC provides the security with a shorter key size and shorter ciphertexts. A 160-bit ECC cryptosystem provides the same security. In energy constraint WSNs, constructing PH via ECC is more efficient.

C. **System Architecture**

Fig 1. shows that Different sensors node send the sensed data to the cluster head. By PH scheme, cluster head encrypt the sensed data forwards to the aggregator and its homorphically aggregate their ciphertexts without decryption. Sensor nodes in the same group share the same public key and no other entities outside the group know the group public key. In data transmission, CDAMA scheme applied to the aggregator then the ciphertexts from different applications can be encapsulated into “only” one ciphertext by scalar multiplication of elliptic curve cryptosystem. Conversely, the base station can extract application-specific plaintexts via the corresponding secret keys. So, the transmission cost is reduced.

These CDAMA approach are all built on elliptic curves, encryption and aggregation are based on two kinds of operations, point addition and point scalar multiplication. These, scalar multiplication used to Base station individually decrypts the aggregated ciphertext to extract the aggregated value of each application. ECC provides the security with a shorter key size and shorter ciphertexts.

![System Architecture](image)

Fig. 1: System Architecture.
CDAMA provide data counting to the aggregator (AG). It degrades the damage from unauthorized aggregations. Here, Sensor node sends the public key with data to the aggregator. Aggregator combines the secure counts and data with public key and its sends to the base station. Base station verifies the secret key with combination of secure counts and data to user. Here, base station applies the PH scheme for securely decrypt the data to the user.

III. CDAMA

CDAMA is designed by using multiple points, each of which has different order. We 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).

A. 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. 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

![Fig. 2: Procedures of CDAMA(k=2)](image)

by decrypting the ciphertext with the corresponding Secret key (SK). Considering deployment, the private keys should be kept secret and only known by the BS. SNs in the same group share the same public key and no other entities outside the group know the group public key.

B. Generalization of CDAMA

CDAMA (k = 2) can be generalized to CDAMA (k > 2). The paradigm of generalization uses different generators to construct different key pairs for groups. The generalized CDAMA is shown in Fig. 3

Under CDAMA, the ciphertexts from different applications can be aggregated together, but they are not mixed. The ciphertexts can be integrated into a ciphertext and transmitted to the BS. The BS then individually decrypts the aggregated ciphertext to extract the aggregated value of each application.
C. Key Distribution

In this section describe how to deliver the group public keys to SNs securely. There are two main approaches.

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

Key postdistribution. 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, and the master secret key. Once these SNs are deployed, they can run the to elect the AGs and construct clusters. After that, the BS sends the corresponding CDAMA keys, encrypted by the preshared key, to SNs and AGs.

IV. APPLICATIONS

In this section, we propose three applications that are realized by only CDAMA multigroup construction.

D. Conventional Aggregation Model with Multiple Groups

CDAMA applied to the conventional aggregation model can mitigate the impact from compromising attacks. Fig 4 shows an example of this case. 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.

| KEYGEN($r$): generate public-private key pairs for group $G_i$, $\forall i = 1 \sim k$ |
| --- |
| 1. Based on security parameter $r$, compute elements $(q_i, q_{i+1}, \ldots, q_{k+1}, E)$, where $E$ is the set of elliptic curve points which form a cyclic group; $\text{ord}(E) = n$; $n$ is the product of $q_i, \ldots, q_{k+1}$ and $q_i, \ldots, q_{k+1}$ are large primes; the bit length of $q$ is the same, i.e., $|q_i| = \cdots = |q_k| = |q_{k+1}|$. |
| 2. Randomly pick up $k+1$ generators, $G_1, \cdots, G_{k+1} \in E$ where $\text{ord}(G_i) = n$, $\forall i$. |
| 3. Compute point $H = (\prod_{i=1}^{k} q_i) \ast G_{k+1}$ such that $\text{ord}(H) = q_{k+1}$. |
| 4. Let $T$ be the maximum plaintext boundry where Pollard’s $\lambda$ method is feasible, and let $T_i = \left\lfloor \frac{T}{x} \right\rfloor$, $i = 1 \sim k$, where $x = \text{the average number of sensors in an application}$. |
| 5. Compute point $P_i = (\prod_{i=1, i \neq j}^{k+1} q_i) \ast G_i$ such that $\text{ord}(P_i) = q_i$, for $i = 1, \cdots, k$. |
| 6. Output $G_i$’s group public key ($PK_i$): $PK_i = (n, E, P_i, H, T_i)$. |
| 7. Output the private key $SK_i = (q_1, q_2, \cdots, q_{k+1})$. |

ENC($PK_i$, $M$): Message encryption in $G_i$

1. Check if message $M \in \{0, \cdots, T_i\}$. |
2. Randomly select $R \in \{0, \cdots, n-1\}$. |
3. Generate the ciphertext $C$ as: $C = M \ast P_i + R \ast H$ where $P_i \in PK_i$. |
4. Return $C$. |

AGG($C_1, C_2$): Message aggregation on two ciphertexts $C_1$ and $C_2$

1. Aggregated ciphertext $C' = C_1 + C_2 = \sum_{i=1}^{k} (\sum M_i) \ast P_i + (\sum R_i) \ast H$ where $\sum M_i$ represents the aggregated result of group $G_i$, $\sum R_i$ presents the aggregated randomness of all groups. |
2. Return $C'$. |

DEC($SK_i$, $C$): Message decryption on $C$ for group $G_i$ using private key $SK_i$

1. Compute $M = \sum M_i = \log_{g_i}((\prod_{i=1, i \neq j}^{k} q_i) \ast C)$ where $g_i = (\prod_{i=1, i \neq j}^{k+1} q_i) \ast P_i$. |
2. Return $M$. |

Fig. 3: Procedure for generalization CDAMA.

Once an adversary compromised a SN in group A; it only reveals $PK_A$, not $PK_B$. Since the adversary can only forge messages in group A, not group B, the SNs in group B can still communicate safely.
E. Aggregation with Secure Counting

The BS does not know the exact number of ciphertexts aggregated (here, we call “count”), 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.

Fig.5 shows, that the BS obtains the aggregated result M and its count. If a malicious AG launches unauthorized aggregations, such as repeated or selective aggregation, count value would be changed to a bigger or smaller value than the reference count. Since the AG does not know the base points P and Q, unauthorized aggregations have to alter the values of count and M simultaneously; it is impossible to alter M without changing count. Meanwhile, the BS knows the number of deployed sensors through gathering topology information, the BS can detect unauthorized aggregation based on the value of count.

V. DISCUSSIONS

A. Efficient Scalar Multiplication

In CDAMA, the efficiency of encryption and decryption depends on the performance of scalar multiplication on elliptic curves. Decryption is not

---

Fig. 4: Two groups for a single application

Fig. 5: Secure counting procedures of CDAMA(k=2) for a single application
considered here because a BS is considering as powerful as a workstation.

Some well-known techniques are listed as follows:
1) By using endomorphism structures
2) By using a different representation of the scalar multiplicand, like nonadjacent form (NAF).
3) By using different coordinates, such as projective coordinates.
4) By choosing optimal extension fields. Through these technologies, encryption operations can be accelerated.

B. Generating Suitable Curves

The main challenge of constructing CDAMA is generating the set of elliptic curve points with a given order (generating the curves with given orders). The BGN scheme adopts pairing-friendly curves (also called super singular curves) to construct their scheme because bilinear pairing is necessary. However, these curves do not have computational efficiency because the length of the underlying field doubles; if the given order is k-bit long, the underlying prime field requires 2k bits.

In CDAMA, we select different approach because bilinear pairing is no longer required and length of the prime field doubles based on the given order in pairing-friendly curves. To find suitable curves in CDAMA, approaches to generate desired curves.

VI. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed approach.

A. Performance Gain of CDAMA

For a leaf node, the energy consumption in CDAMA is several thousand times greater than that in DFS because the encryption cost of CDAMA is significantly greater than the cost of AES.

For an AG, the energy consumption in DFS is increased tenfold whenever the AG reaches to the next layer, whereas the energy consumption in CDAMA are all kept the same. The reason for this is because the ith layer AG must forward 10 messages in DFS but only 10 messages in CDAMA. For a forwarder, the main energy consumption depends on transmission; therefore, CDAMA allows the forwarder to spend only 1 percent of transmission cost in DFS. In summary, in contrast to DFS, the closer to the BS a node is, the more energy the node can save via CDAMA. The computation cost of CDAMA is significantly large. Although data aggregation can reduce the communication effectively, sensors must pay higher computation cost for encryption and aggregation. To argue with this point, we estimate the performance gain from the whole WSN based on CDAMA.

As a result, CDAMA can extend the lifetime of higher layer AGs and forwarders effectively, but the lifetime of leaf nodes will be shortened to less than one-thousandth of original lifetime. Thus, the major challenge in CDAMA is to reduce the encryption cost.

VII. CONCLUSION

For a multi-application environment, CDAMA is the first CDA scheme. Through CDAMA, the ciphertexts 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. Besides the 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 the future, we wish to apply CDAMA to realize aggregation query in Database-As-a- In provider. Therefore, the client has to secure their database stores her database on an untrusted service through PH schemes Service (DAS) model.

DAS model, a client because PH schemes keep utilizable properties than standard ciphers. Based on PH schemes, the provider can conduct aggregation queries without decryption. The most important of all is that we do not have to consider the computation cost and the impact of compromising secret keys (i.e., compromising a client in DAS model is harder than compromising a sensor). Those drawbacks will no longer be issues in CDAMA.

REFERENCES

[1] H. Cam, S. Özdemir, P. Nair, D. Muthuvalu, and H.O. Sanli, “Energy-Efﬁcient Secure Pattern Based Data Aggregation for Wireless Sensor Networks,” Computer Comm., vol. 29, no. 4, pp. 446–455, 2006.
[2] 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.
[3] J. Girao, D. Westhoff, M. Schneider, N. Ltd, and G. Heidelberg, “CDA: Concealed Data Aggregation for Reverse Multicast Traffic in Wireless Sensor Networks,” Proc. IEEE Int’l Conf. Comm. (ICC ’05), vol. 5, 2005.
[4] H. Sanli, S. Özdemir, 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] 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] R. Cramer and V. Shoup, “A Practical Public Key Cryptosystem Provably Secure against Adaptive Chosen Ciphertext Attack,” Proc. 18th Ann. Int’l Cryptology Conf. Advances in Cryptology, pp. 13-

[7] C. Castelluccia, E. Mykletem, and G. Tsudik, “Efficient Aggregation of Encrypted Data in Wireless Sensor Networks,” Proc. Second Ann. Int’l Conf. Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous ‘05), pp. 109-117, 2005.

[8] J. Domingo-Ferrer, “A Provably Secure Additive and Multiplicative Privacy Homomorphism,” Proc. Fifth Int’l Conf. Information Security, pp. 471-483, 2002.