An Efficient Data Aggregation Scheme With Security And Cost Optimization Scheme

Sangeetha.P
PG Scholar
K.S.Rangasamy College Of Technology
Namakkal
India
sangeepunitha@gmail.com

Poornima.R
Assistant Professor
K.S.Rangasamy College Of Technology
Namakkal
India
poorniom@gmail.com

ABSTRACT--Sensor networks are constructed with the set of sensor devices. Sensor device transfers the data to other nodes. Radio frequency is used for the data transfer process. Bandwidth and Traffic factors are considered in sensor networks. Data aggregation technique is used to transfer the sensed data from the sensor node to the base station. Privacy homomorphism encryption technique is used with the data aggregation process in sensor networks. Cluster heads can directly aggregate the cipher text without decryption. The base station only retrives the aggregated result, not individual data. Recoverable property supports the sensing data extraction from the aggregated data values. Recoverable Concealed Data Aggregation(RCDA) technique integrates the data aggregation and data aggregation process. Elliptic curve Cryptography (ECC) and boneh signature scheme are used for the security process. RCDA model is adopted on homogenous and hetrogenous WSN. Recoverable Concealed Data Aggregation(RCDA) technique is integrated with sleep-wakeup scheduling schemes. Dynamic cost prediction algorithm is proposed to estimate cost factor. key distribution process is managed with network life time details. Dynamis Aggregation function selection model is supported by the system.

Keywords—Concealed Data Aggregation, Wireless Sensor Networks,Privacy Homomorphism Encryption, Eliptic Curve Cyptography,Security

I.INTRODUCTION

Wireless sensor networks (WSN) have been widely deployed in many applications, e.g., military field survellience, health care, environment monitor, accident report ,etc. A WSN is composed of a large number of sensors which collaborators with each other. Each sensor detects a target within its radio range, performs simple computations and communications with other sensors.

Generally, sensors are constrained in battery power, communication, and communication capability; therefore, reducing the power consumption is a critical concern for a WSN. The original concept is to aggregate multiple sensing data by performing algebraic or statistical operations such as addition, multiplication ,median, minimum, maximum and mean of a data set etc .Normally data aggregation is performed by cluster heads if the whole network is divided into several groups known as clusters9]. For example, in military fields, sensors are deployed to ensure radiation or chemical pollution. The base station may require the maximum value of all sensing data to trigger the immediate response; thus, each cluster head selects the maximum value of all sensing data of its cluster members and sends the result to the base station. Obviously, communication cost is reduces since only aggregated results reach the base station.

Unfortunately, an adversary has the ability to capture cluster heads. It would cause the compromise of the whole cluster; consequently, several schemes, such as ESPDA [4] and SRDA have been proposed. However , these schemes restrict the data type of aggregation or cause extra transmission overhead. Besides , an adversary can still obtain the sensing data of its cluster members after capturing a cluster head.

To solve above problems completely, two ideas are used in recent research [6][8].First, data are encrypted during transmission. Second, cluster heads directly aggregate encrypted data without decryption .A well known approach named Concealed Data Aggregation(CDA) has been proposed based on these ideas.CDA provides both end to end encryption and in-networking processing in WSN. Since CDA applies Privacy Homomorphism(PH) encryption with additive homomorphism, cluster heads are capable of executing addition operations on encrypted numerical data. Later, Several PH based data aggregation schemes have been proposed to achieve higher security levels.

II.RELATED WORKS

Numerous secure data aggregation schemes have been proposed. These schemes are designed for different security requirements. A number of schemes[1] have been proposed on the commit and attest principle. In these schemes, the base station broadcasts aggregation results to all sensors. Then ,every sensor verifies that its sensing data were indeed counted. Another work[3] can actually count and sum even if a few compromised sensors inject false values.YU[2] introduces a random sampling technique that enables aggregation queries to not only detect malicious sensors, but also to tolerate them.

On the other hand, several studies [5] attempt to provide confidentiality. That is, an aggregator can directly execute addition operations on encrypted numerical data.CDA places more emphasis on passive attack. More specifically, it considers if adversaries can eavesdrop the communications on the air. After CDA, succeeding Research has been proposed to achieve higher security levels. They consider the following scenario. If sensors within the same cluster encrypt their sensing data with a common secret key, an adversary may decrypt or fake the aggregated ciphertext by compromising only one sensor. Later, Mykletun et al. proposed a data aggregation scheme based on addition homomorphic public-key encryption. It seems more secure since every sensor stores only public key. The adversary cannot launch the same attack through compromising only one sensor. Nevertheless, the adversary can still impersonate other legel sensors to send the forged cipertexts to the cluster head with the same public key. Authenticity of data is not supported.

III. NETWORK AND ATTACK MODEL

In this section, we first describe the network models and define the attack model. Then, Mykletun et al.'s and Boneh et al.'s schemes are reviewed since they are the foundation of the proposed schemes.

A. Network Model

A WSN is controlled by a base station (BS). A BS has large bandwidth, strong computing capability, sufficient memory, and stable power to support the cryptographic and routing requirements of the whole WSN. Besides the BS, sensors (SNs) are also deployed to sense and gather responsible results for the BS. Typical SNs are small and low cost; hence, SNs are limited on computation, storage, and communication capability.

Generally, all SNs in a WSN may be divided into several clusters after being deployed. Several research have shown that a cluster-based WSN has several advantages such as efficient energy management, better scalability of MAC (medium access control) or routing, etc. Each cluster has a cluster head (CH) responsible for collecting and aggregating sensing data from SNs within the same cluster. A CH then sends the aggregation results to the BS. In a homogeneous WSN, cluster heads act as normal SNs. On the other hand, cluster heads act as by powerful high-end sensors (H-Sensors), in a heterogeneous WSN which incorporates different types of SNs with different capabilities.

B. Attack Model

The attack model is defined based on the ability of adversaries. Here, we consider the following three cases: 1. Without compromising any SN or CH. An adversary can only eavesdrop on packets in the air, so he can modify or inject the forged messages with this public information. 2. Compromising SNs. After compromising a SN, an adversary can obtain secrets such as encryption/ decryption keys. Then, an adversary can obtain sensing data and packets passed through the captured SN or impersonate this compromised sensor to forge malicious data. 3. Compromising CHs. After compromising a CH, an adversary can obtain the secrets and perform the following attacks. First, an adversary can decrypt the ciphertext of sensing data sent by its cluster members. Second, an adversary can generate forged aggregation results.

C. Mykletun et al.'s Encryption Scheme

Mykletun et al. proposed a concealed data aggregation scheme based on the elliptic curve ELGamal (EC-EG) cryptosystem. It consists of four procedures: key generation (KeyGen), encryption (Enc), aggregation (Agg), and decryption (Dec). Symbol + and * denote addition and scalar multiplication on elliptic curve points, respectively.

D. Boneh et al.'s Signature Scheme

Boneh et al. proposed an aggregate signature scheme which merges a set of distinct signatures into one aggregated signature. This scheme consists of five procedures: key generation (KeyGen), signing (Sign), verifying, aggregation and verifying aggregated signature (Agg-Verify). Boneh et al.'s scheme is based on bilinear map en which is defined as $e_n = G_1 * G_2 \rightarrow G_r$, where groups G_1, g_2, and GT are cyclic groups of prime order n. G_1 and G_2 are ntorsion point groups on an elliptic curve € under a finite field F_p, i.e., n * p = n * Q = ∞, where Λp € G_1 and ΛQ € G_2. GT is the group of nth root of uity in an extension field Fp^k , i.e., G_T = {X € Fp^k |$x^n = 1_T$ }. The group operation in G_1 and G_2 is point addition and one in G_T is multiplication over a finite field.

IV. A RCDA SCHEME FOR HOMOGENEOUS WSN (RCDA-HOMO)

In this section, we propose a recoverable concealed data aggregation scheme named RCDA-HOMO for homogeneous WSN. RCDA-HOMO is composed of four procedures: Setup, Encrypt-Sign, Aggregate, and Verify. The Setup procedure is to prepare and install necessary secrets for the BS and each sensor. When a sensor decides to send sensing data to its CH, it performs Encrypt-Sign and sends the result to the CH. Once the CH receives all results from its members, it activates Aggregate to aggregate what it received, and then sends the final results to the BS. The last procedure is Verify. The BS first extracts individual sensing data by decrypting the aggregated ciphertext. Afterward, the BS verifies the authenticity and integrity of the decrypted data based on the corresponding aggregated signature.

To present RCDA-HOMO in a simple way, we choose Cluster 1(see Fig. 1) as an example. SNω is selected as CH of Cluster 1 which contains the remaindering sensors, {SN_1,……,$SN_{\omega-1}$ }. The detailed procedures are listed as follows:

Encrypt-Sign: This procedure is triggered while a sensor decides to send its sensing data to the cluster head (CH_1 in Fig. 1). Detailed steps are listed as follows:

1.Encoding d_i :$m_i=d_i||0^\beta$, where β=1.(i-1).
2.After Encoding SN_i computers:
a. signature:$\sigma_i=x_i*h_i$, where $h_i=H(d_i)$.
b.ciphertext:$c_i=(r_i,s_i)=(k_i*G,M_i+k_i*Y)$, where k_i is randomly selected from {0….n-1}, M_i = map(m_i) = m_i_ G, and n, G, Y € P_{BS}.

Aggregate: The Aggregate procedure is launched after the CH has gathered all ciphertext-signature pairs, i.e., CH_1 gathered ϖ-1 pairs ((c_1, σ_1),…,($c_{\omega-1}$, $\sigma_{\omega-1}$)) over a period of time. Aggregation operations are given as follows:

1. Aggregated ciphertext
   $C^\wedge=\Sigma^{n-1}_{i=1} C_i=(\Sigma^{n-1}_{i=1} r_i, \Sigma^{n-1}_{i=1} s)$
2. Aggregate signature
   $\sigma^\wedge=\Sigma^{n-1}_{i=1}\sigma_i.$
3. Send the aggregate result ($C^\wedge$,$\sigma^\wedge$) to the BS.

Verify: While receiving($C^\wedge$,$\sigma^\wedge$) from CH_1, BS can recover and verify each sensing data via the following steps:

1. BS obtains M' by decrypting with R_{BS} M'=-t*r^+S^=M_1+…+$M_{\omega-1}$.
2. BS obtains m' from M' through the reserve function rmap(): m'=rmap(M')=m_1+….+$m_{\omega-1}$.
3. BS obtains each sensing data from m' by Decode function: Decode (m', ω-1, 1):d_i=m'[(i-1). 1,i.1-1], where i=1,….,ω-1.
4. BS verifies each d_i via checking whether the equation e_n holds or not. Each element hi is derived from hashing $h_i=H(d_i)$. Note that en is the bilinear map. for all d_i, if the equation holds, BS accepts ; otherwise , BS rejects.

Similarly, the BS may receive other ciphertext and signature pairs form other clusters. The BS can recover all sensing data within the whole WSN. After confirming if it wants since all individual are reverted.

V.RCDA SCHEME FOR HETROGENOUS WSN

Here, we consider another environment, hetrogenous WSN. A Concealed data aggregation scheme for hetrogenous WSN has been proposed[7]; however, their scheme does not provide data integrity and recovery. we first propose naïve RCDA-HETE scheme. Later, we will propose another scheme named RCDA-HETE scheme if H-sensors are designed to tamper resistant.

A. *Naïve RCDA-HETE Scheme*

Actually, RCDA-HOMO can be applied to heterogeneous WSN without modification. We call this approach naïve RCDA-HETE. Since H-Sensors ara capable of stronger computation ability and stable power supply, they can perform more complex tasks than L-Sensors. Thus, H sensors can act as cluster heads. Obviously, naïve RCDAAHETE also achieve the Recovery property.

B. *RCDA-HETE Scheme*

RCDA-HETE is composed of five procedures: Setup, Intra cluster Encrypt, and Inter cluster Encrypt, Aggregate, and Verify. In the Setup procedure, necessary secrets are loaded to each H-Sensor and L-Sensor. Intra cluster Encrypt procedure involves when L-Sensors desire to desire to send their sensing data to the corresponding h-sensor. In the Inter cluster Encrypt procedure, each H-Sensor aggregates the received data and then encrypts and signs aggregated result. In addition if an H-Sensor receives ciphertexts and signatures from other H-Sensors on its routing path, it activates the aggregate Procedure. Finally, the verify procedure ensures the authenticity and integrity of each aggregated result.

In our design each L-Sensor is required to share a pairwise key with its cluster head. For example L-sensor L_j I would share a key k^j_i with the corresponding cluster head H_j. If the BS knows the cluster information before deployment the pairwise keys can be preloaded to all L-sensors and H-Sensors. However in most WSN environment sensors are randomly deployed. Thus, we propose a simple key exchange scheme. Intracluster Encrypt: This procedure ensures the establishment of a secure channel between L-Sensors and their H-Sensor. L^1_1 encrypts d^1_i with K^1_i and sends $E_{ki1}(d^1_i)$ to H_1. After receiving E_{kil} (d^1_i), H_1 decrypts the ciphertexts to obtain the plaintext d^1_i. Intercluster Encrypt: After collecting all sensing data from all cluster members, an H-Sensors performs the preferred aggregation function on these data as it result. For example, H_1 select d^1_i as the aggregated result by predefined property, such as maximum or minimum. Then, H_1 performs the following steps:

1. Encoding as m: $m=\delta_1 \| 0^{\beta}$, where β=1. (i-1).
2. After encoding, H1 computers:
   a. Signature:_1=x1_h1,where X1 is RH1 and h1=H(i).
   b. ciphertext:c_1=(r1,s1)=(k1*G,M_1+k_1*y), where k is randomly selected from {1…n},M1=map(m_1)=m_1*G, and G,Y€P_{BS}.

3.H_1 sends the pair (C_1,σ_1) to H_3.Similarly each H_j also calculates (C_j,σ_j) from δ_j in other clusters.

Aggregate:If H_3 receives (C_1,σ_1) from H1 and (C_2,σ_2) from H2,H3 execute this procedure to aggregate (C_1,σ_1),(C_2,σ_2) and its own (C_3,σ_3) as follows:

1. Aggregated ciphertexts: $C_3^{\wedge}=(\sum^3_{i=1}\gamma_I, \sum^3_{i=1}S_i)$.
2. Aggregated signature: $\sigma_3{}^\wedge=(\sum^3_{i=1}\sigma_i)$

   Finally, H3 sends ($C_3^{\wedge}$,$\sigma_3{}^\wedge$) to H5. Similarly, H5 can also aggregate (C_4, σ_4), (C_5,σ_5), and ($C_3^{\wedge}$, $\sigma_3^{\wedge}$) and get a new aggregate result ($C_5^{\wedge}$, $\sigma_5^{\wedge}$) to the BS.

   Verify: After receiving the end result ($C_5^{\wedge}$, $\sigma_5^{\wedge}$), BS will perform the following steps:

1. Obtain M' by decrypting $C_5^{\wedge}$:M'=-t*r+s=M_1+M_2+..+M_5.
2. Obtain m' from M' through the reserve function rmap():m'=rmap(M')=m_1+m_2+….+m_5.
3. Obtain δ_I from m' using the Decode function:Decode(m',5,1):δ_i=m$'$ [(1-1).1,1.i-1], where i=1….5
4. Check whether $e_n(\sigma_{5\wedge},g2)=\Pi^5_{i=1}\, e_n(h_i,p_i)$ holds or not. Element h_i is derived by hashing δ_I,i.e.,h_i=H(δ_i).e_n isnthe bilinear map.If the equation holds, accept all δ_I; otherwise rejectr.After checking the integrity of each δ_I the BS can further perform the aggregation function on all δ_i.

C. *Recovery Property*

The recovery property attempts to provide two functionalities. First, BS can verify the integrity and authenticity of sensing data.

VI.DATA AGGREGATION WITH COST OPTIMIZATION AND SECURITY

The sensor data capturing is tuned with scheduling schemes. The scheduling schemes are applied to manage energy levels. The redundant sensor nodes are placed to manage failures and data aggregation process. The data capture process is divided into a set of sessions. In each data session a set of sensors are assigned to the data capture process and remaining sensor nodes are sensors. The scheduling techniques are applied to improve the energy level and lifetime of the sensor network. Recoverable Concealed Data Aggregation technique is integrated with sleep wakeup scheduling schemes. The sleep-wake up scheduling algorithm is used to schedule data capturing on the sensor nodes. The scheduling process initiates the route changes and query responses update operations under the sensor network.

The cost estimation methods are used to predict the cost requirements for the data capture process. The cost factor includes the energy cost and transmission cost for the nodes. The energy consumption levels are considered for each data aggregation process. The computational cost is measured for confidentiality and integrity operations. The computational cost is also estimated for data aggregation process. The energy usage for data concealment is also considered in the cost prediction process. The system selects the feasible cost level based data aggregation and transmission process .Dynamic cost prediction algorithm is proposed to estimate cost factor in RCDA scheme. The cost estimation is dynamically initiated for all data capture sessions. The scheduling factors also considered in the cost prediction process.

The sensor network nodes are divided into two forms. They are homogenous and hetrogeneous nodes. In the homogeneous network model all the sensor nodes are assigned with the same properties such as energy level, sensing coverage, transmission coverage and buffer size factors. The scheduling is initiated with reference to the properties. The heterogeneous network is formed nodes with different properties. the data security is provided with Mykletunet al.'s Encryption scheme .The key values are distributed to the sensor nodes. key distribution process is managed with network lifetime details. The concealed data values are secured with the distributed key values. Boneh et al.'s Signature scheme is used for the data integrity verification in the sensor network query processing,

The Recoverable Concealed Data Aggregation (RCDA) scheme is enhanced to reduce cost under the homogeneous network model. The data aggregation is carried out with a set of aggregation functions such as maximum, minimum and average data values. The data aggregation scheme is build with any one of the data aggregation methods. The RCDA scheme is enhanced with a set data aggregation function mechanism. Dynamic aggregation function selection model is supported by the system. The user can select any aggregation function at runtime. The destination node can select the aggregation function. The data values are aggregated with reference to the selected aggregation function.

VII. CONCLUSION

Sensor networks are constructed to monitor the environment. Individual and aggregated data values are updated to the base station based on queries. Recoverable Concealed Data Aggregation (RCDA) technique is improved with scheduling and cost management methods. Choice based aggregation function selection model is used in the system. The sensor network is constructed with secured data transmission process. Network connectivity is managed by the system. Cryptographic overhead is controlled in the system. The system supports energy efficient transfer scheme.

REFERENCES

[1] Y.Yang,X.Wang,S.Zhu, and G.Cao, "SDAP: A Secure Hop-by-Hop Data Aggregation Protocol for Sensor Networks," ACM Trans.Information and System Security(TISSEC), vol.11,no.4,pp.1-43,2008.

[2] H.Yu, "Secure and Highly available Aggregation Queries in Large Scale Sensor Networks via set Sampling," Proc.IEEE Int'l Conf.Information Processing in Sensor Networks,pp.1-12,2009.

[3] S. Roy, S. Setia, and S. Jajodia, "Attack-Resilient Hierarchical Data Aggregation in Sensor Networks," Proc. ACM Fourth Workshop, pp. 71-82,2006.

[4] H. C, am, S.O. zdemir and H. Ozgur Sanil, “Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks,” J. Computer Comm., vol. 29, pp.44-455,2006.
[5] S. Roy, S. Setia, and S. Jajodia, “Attack-Resilient Hierarchical Data Aggregation in Sensor Networks,” Proc. ACM Fourth Workshop Security of Ad Hoc and Sensor Networks, pp. 71-82,2006.
[6] 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 Coputing, vol. 5, no. 10, pp. 1417-1431, Oct. 2006.
[7] S. Ozdemir, “Concealed Data Aggregation in Heterogeneous Sensor Networks Using Privacy Homomorphism,” Proc. IEEE Int’1 Conf. Pervasive Services, pp. 165-168, July 2007.
[8] E. Mykletun, J. Girao, and D. Westhoff, “Public Key Based Cryptoschemes for Data Concealment in Wireless Sensor Networks,” Proc. IEEE Int’1 Conf. Comm., June 2006.
[9] Chien-Ming Chen, and Hung-Min Sun, “RCDA: Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks,” IEEE Transactions on Parallel and Distributed Systems, April 2012.