An Integrity-Assured Concealed Data Aggregation Scheme for Wireless Sensor Networks

L.J. Yang, C. Ding, M. Wu
Nanjing University of Posts & Telecommunications 66 Xin-mofan Road
Nanjing, Jiangsu Province, China

***Abstract*-To address the contradiction between data aggregation and data security in wireless sensor networks, an integrity-assured concealed data aggregation scheme is proposed based on the privacy homomorphism and the aggregate message authentication code techniques. The proposed scheme provides both end-to-end privacy and data integrity in data aggregation for wireless sensor networks. Besides, it supports any type of aggregation functions since the base station can recover each sensing data collected by all sensors even if these data have been aggregated by aggregators. The performance analysis shows that the proposed scheme is efficient in computation and communication, and is feasible for resource limited sensor networks.**

Keywords- wireless sensor networks; data aggregation; privacy preserving; integrity protection

I. INTRODUCTION

Wireless sensor networks (WSNs) are comprised of a large number of small sensor nodes which are spatially distributed across the field of interest. WSNs have been widely deployed in many areas including military, healthcare and environment, etc. Sensor nodes are usually resource-limited and power-constrained. In order to conserving communication bandwidth and energy, data aggregation technology was introduced. The concept of data aggregation [1] 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.

Unfortunately, sensor nodes are vulnerable to be captured in a hostile environment. In terms of security, data aggregation is risky. A sensor node that is compromised by an adversary can either illegally disclose the data it collects from other nodes or report arbitrary values as its aggregation results. Therefore, an adversary can compromise both the confidentiality and the integrity of the data of a large portion of the WSN by capturing some data aggregators that are close to the base station (BS).

Both data aggregation and security are critical for WSNs, so achieving secure data aggregation has been an attractive goal for researchers. Some existing secure data aggregation schemes, such as delay aggregation [2], SIA [3], and SDAP [4] have been proposed to solve the data fabrication problem. However, most of these schemes cause negative effect on other performance metrics, such as delay and data confidentiality. Recently, a popular idea is to study this problem based on privacy homomorphism (PH). In PH-based secure data aggregation schemes such as CMT [5], CDA [6], and Mykletun et al's scheme [7], the aggregators directly aggregate ciphertext without decryption. However, these PH-based schemes usually limit the type of data aggregation, and could not provide data integrity.

In this paper, we propose an integrity-assured concealed data aggregation scheme which is based on the privacy homomorphism and the aggregate message authentication code (MAC) techniques [8] to solve the above problems. Our proposal provides both end-to-end privacy and data integrity in data aggregation for WSNs. Besides, in our scheme, the BS can recover each sensing data collected by all sensors even if these data have been aggregated by aggregators, thus can perform any aggregation functions on them. Performance analysis indicates the proposed scheme is efficient in terms of computation and communication.

II. THE PROPOSED SCHEME

A. System Model

We consider a cluster-based WSN, which is comprised of a BS and a number of sensor nodes (*SN*). Generally, BS which connects the system to the networks and users has large bandwidth, strong computing capability, and sufficient memory and stable power to support the cryptographic and routing requirements of the whole WSN. Typically, *SN*s deployed to sense and gather related data are tiny and low-cost devices, hence *SN*s are limited on computation, storage and communication capability. After deployment, all *SN*s are divided in several clusters. *SN*s in the same cluster select one of them as the cluster head (*CH*), which is responsible for collecting and aggregating sensor data from *SN*s within the same cluster and finally sends aggregated results to BS.

B. Detailed Procedures of the Proposed Scheme.

In this section, we proposed an integrity-assured concealed data aggregation scheme based on the privacy homomorphism and aggregate MAC techniques. The proposed scheme is composed of four polynomial algorithms: *Setup*, *Encrypt-MAC*, *Aggregate* and *Decrypt-Verify*. The *Setup* algorithm is to prepare and install necessary parameters and key materials for the BS and each sensor node before deployment. When a sensor node decides to send sensor data to its *CH*, it performs *Encrypt-MAC* and sends the results to its *CH*. Once the *CH* receives all results from its members, it performs the *Aggregate* to aggregate what it received, and then sends the final results (aggregated ciphertext and MAC) to BS. After receiving data from *CH*s, BS performs *Decrypt-Verify* to extract individual sensing data by decrypting the aggregated ciphertext, and verify the authenticity and integrity of the plaintexts based on the corresponding aggregated MAC.

To present the proposed scheme in a simple way, we consider the case that there is only one BS in the sensor network. Without loss of generality, we assume that there are η sensor nodes i.e. $SN_1, SN_2, \cdots, SN_\eta$, and SN_η is selected as *CH* of this cluster. The detailed procedures are listed in Fig.1.

III. PERFORMANCE ANALYSIS

In this section, we analyze the performance of our proposal in terms of computation and communication overhead respectively. The costs on computation and communication of each procedure are listed in Table 1.

A. Computation Overhead

We employ the EC-EG homomorphism encryption algorithm to provide data end-to-end privacy, the operation parameters are selected from the elliptic curve defined on the finite field $\mathbb{F}_p$. The decryption and integrity validation operations are run on the BS, which is a resource-rich node. Thus, we are mainly concerned with the cost on the sensor node and cluster head (aggregator). The sensor node runs the *Encrypt-MAC* function, which needs two $|p|$-bit scalar point multiplication, one $|p|$-bit point addition, and one hash operations. The cluster head runs *Aggregate* function, which needs $2(\eta\text{-}2)$ $|p|$-bit point addition, and $(\eta\text{-}2)$ XOR operations when the number of cluster member is $(\eta\text{-}1)$. In order to achieve the 1024-bit RSA equivalent security, parameter p is selected as a 160-bit large prime.

B. Communication Overhead

We choose the elliptic curve $E(\mathbb{F}_{160})$, where one point (x, y) occupied 2×160 bit. With the help of node compression techniques, the point can be compressed to 161bits (21 byte). In our scheme, the message sent by each sensor node is in the form of $(c_i \| \text{tag}_i)$, which is comprised of a ciphertext and a MAC tag. The ciphertext $c_i = (R_i, S_i)$ occupies 42 bytes. The MAC employs HMAC algorithm, thus tag_i is 128 bits (16 byte). Therefore, the message sent by each node has a length of 58 bytes. Here we do not take the extra overhead arose from TinyOS packet encapsulation into consideration. The size of packets sent is fixed even if the cluster has numerous member nodes. The message sent by each sensor node and cluster head is in constant size, including one (aggregated) ciphertext and one (aggregated) tag.

The above performance analysis indicates the proposed scheme is efficient in terms of computation and communication.

TABLE I. PERFORMANCE EVALUATION.

Role	**Sensor node (SN)**	**Cluster head (CH)**
Procedure	Encrypt-MAC	Aggregate
Computation	$2M_{160} + 1A_{160} + 1H$	$2A_{160} + \text{XOR}$
Communication	58 bytes	58 bytes

IV. CONCLUSION

In this paper, we propose an integrity-assured concealed data aggregation scheme based on the privacy homomorphism and the aggregate message authentication code techniques. The proposed scheme provides both end-to-end privacy and data integrity for WSNs. Besides, in our scheme, the BS can recover each sensing data collected by all sensors even if these data have been aggregated by aggregators, thus can perform any aggregation functions on them. Performance analysis indicates the proposed scheme is efficient in computation and communication, and is feasible for WSNs.

ACKNOWLEDGEMENTS

This work is financially supported by National Basic Research Program of China (973 Program) under Grants 2011CB302903, the National Natural Science Foundation of China (Grant No. 61100213), the Key Program of Natural Science for Universities of Jiangsu Province (Grant No.10KJA510035), the Specialized Research Fund for the Doctoral Program of Higher Education (20113223120007), the Science and Technology Program of Nanjing (201103003) and the Postgraduate Innovation Project Foundation of Jiangsu Province (Grant No. CXLX11_0411).

REFERENCES

[1] Fasolo E, Rossi M, Widmer J, et al. In-network aggregation techniques for wireless sensor networks: a survey[J]. IEEE Wireless Communications, 2007, 14(2): 70-87.

[2] Lingxuan H,Evans D. Secure aggregation for wireless networks. Proceedings of 2003 symposium on Applications and the Internet Workshops (SAINT'03). Orlando, FL, USA, 2003: 384-391.

[3] Przydatek B, Song D,Perrig A. SIA: Secure information aggregation in sensor networks[C]. Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (Sensys'03), Los Angeles, California, USA, 2003: 255-265.

[4] Yang Y, Wang X, Zhu S, et al. SDAP: A Secure Hop-by-Hop Data Aggregation Protocol for Sensor Networks[J]. ACM Transactions on Information System Security (TISSEC), 2008, 11(4): 1-43.

[5] Castelluccia C, Chan A C-F, Mykletun E, et al. Efficient and provably secure aggregation of encrypted data in wireless sensor networks[J]. ACM Transactions on Sensor Networks, 2009, 5(3): 1-36.

[6] Westhoff D, Girao J,Acharya M. Concealed Data Aggregation for Reverse Multicast Traffic in Sensor Networks: Encryption, Key Distribution, and Routing Adaptation[J]. Mobile Computing, IEEE Transactions on, 2006, 5(10): 1417-1431.

[7] Mykletun E, Girao J,Westhoff D. Public Key Based Cryptoschemes for Data Concealment in Wireless Sensor Networks. Proceedings of 6th International Conference on Communication (ICC'06). Istanbul, Turkey 2006: 2288-2295.

[8] Katz J,Lindell A. Aggregate Message Authentication Codes[C]. Proceedings of the Cryptographers' Track at the RSA Conference, 2008: 155-169.

Setup: BS generates the following parameters.

1. BS generates elliptic curve parameter $D=(p,E,P,n)$ and the key pair (x,Y) by *KeyGen*(τ) algorithm in EC-EG scheme;
2. BS preloads $\langle D,Y,sk_i\rangle$ for each sensor SN_i ;
3. BS keeps the master key *x* and sk_i (shared with each sensor SN_i)

Encrypt-MAC: when a sensor decides to send its sensing data to its *CH*, it does the following steps.

1. generates $tag_i = MAC_{sk_i}(data_i)$;
2. encodes the sensing data, encode $(data_i): m_i = data_i \| 0^{\beta}$, where $\beta = l\cdot(i-1)$;
3. maps the message m_i to a point on the curve: $M_i = map(m_i)$;
4. computes the ciphertext $c_i = (R_i, S_i) = (k_i * P, M_i + k_i * Y)$, where $k_i \in_R [1, n-1]$;
5. sends (c_i, tag_i) to the *CH*.

Aggregate: *CH* runs this procedure after it has gathered all ciphertext-tag pairs.

1. Aggregates ciphertext: $C = \sum_{i=1}^{\eta-1} c_i = \left(\sum_{i=1}^{\eta-1} R_i, \sum_{i=1}^{\eta-1} S_i\right)$;
2. Aggregates tags: $Tag = \oplus_{i=1}^{i=\eta-1} tag_i = \oplus_{i=1}^{i=\eta-1} MAC_{sk_i}(data_i)$;
3. Sends the aggregated result (C, Tag) to the BS.

Decrypt-Verify: on receiving (C, Tag) from *CH*, BS can recover and verify each sensing data

1. BS obtains the aggregated plaintext *M* by decrypting with master key *x*, $M = -x * R + S = M_1 + \cdots + M_{\eta-1}$;
2. BS computes $m = rmap(M) = m_1 + \cdots + m_{\eta-1}$;
3. BS obtains each sensing data from *m* by Decode function: $Decode(m, \eta-1, l) = \{data_i = m[(i-1)\cdot l, i\cdot l - 1]\}$, where $i = 1, \cdots, \eta-1$;
4. BS accepts the data if the equation $Tag = \oplus_{i=1}^{i=\eta-1} MAC_{sk_i}(data_i)$ holds, otherwise, rejects.

FIGURE I. THE PROPOSED SCHEME.