Secure Data Aggregation and Data Recovery in Wireless Sensor Networks

John Major. J, Shajin Prince, Akuluri Rakesh

Abstract—several data aggregation schemes based on privacy homomorphism encryption have been designed and reviewed on wireless sensor networks. Cluster heads can exactly aggregate the cipher texts without decryption; thus, transmission overhead is reduced. Though, the base station only fetches the aggregated result, which origin two problems. First, the usage of aggregation function is obliged. Second, the base station cannot confirm the data integrity and authenticity. This paper go to overcome the above two drawbacks. In the design, the base station can recover all the sensing data even the data has been aggregated. Besides, the design has been concluded and adopted on both homogeneous and heterogeneous wireless sensor networks.

Keywords- Data aggregation; Wireless sensor networks; Privacy homomorphism encryption.

I. INTRODUCTION

Wireless Sensor Networks (WSN) has been widely deployed in many applications, e.g., health care, military field surveillance, accident report, environment monitor, etc. A WSN is composed of a large number of sensors which conspires with each other. Each sensor finds a target within its radio range, achieves simple computations, and communicates with other sensors. The sensors are constrained in computation capability, communication, and battery power; accordingly, reducing the power consumption is a critical concern for a WSN. In recent times, a practical solution called data aggregation was introduced. The original idea is to aggregate multiple sensing data by performing algebraic or statistical operations such as addition, multiplication, median, maximum, minimum, and mean of a data set, etc. Data aggregation is performed by cluster heads if the whole network is divided into several groups known as clusters. The base station (sink) may require the maximum value of all sensing data to trigger the immediate response; hence, each cluster head selects the maximum value of multiple sensing data of its cluster members and sends the result to the base station. Definitely, communication cost is reduced as only aggregated results reach the base station.

Unfortunately, an adversary has the ability to capture cluster heads. It would origin the compromise of the whole cluster; consequently, several schemes, such as ESPDA and SRDA, have been proposed. However, these schemes restrict the data type of aggregation or cause extra communication overhead. Besides, an adversary can still access the sensing data of its cluster members after capturing a cluster head. To solve above problems completely, two ideas are used in recent research. First, data are encrypted during transmission. Second, cluster heads express aggregate encrypted data without decryption. A well-known entrance named Concealed Data Aggregation has been proposed based on these two ideas. CDA adds both end-to-end encryption and in-networking processing in WSN. CDA applies privacy homomorphism encryption with additive homomorphism; cluster heads are capable of executing addition operations on encrypted binary data. Next, several PH-based data aggregation schemes have been proposed to achieve higher security levels.

In the above PH-based schemes, the base station receives only the aggregated results. But, it brings two problems. First, the method of aggregation functions is strained. For example, these designs only allow cluster heads to perform additive operations on cipher texts sent by sensors; thus, they are ineffective if the base station desires to query the maximum value of all sensing data. Second, the base station cannot justify the integrity and authenticity of each sensing data. These issues seem to be solved if the base station can receive all sensing data rather than aggregated results, therefore this method is in direct contradiction to the concept of data aggregation that the base station obtains only aggregated results. Hence, we attempt to design an approach that allows the base station to receive all sensing data but still reduce the transmission overhead.

In this paper, the concept named Recoverable Concealed Data Aggregation (RCDA). In RCDA, a base station can recover each sensing data generated by all sensors even if these data have been aggregated by cluster heads (aggregators). With these own data, two functionalities are prepared. First, the base station can find out the integrity and authenticity of all sensing data. Second, the base station can execute any aggregation functions on them. Then, we recommend two RCDA schemes named RCDA-HOMO and RCDA-HETE for homogeneous and heterogeneous WSN respectively. In the security analysis, we manifest that the proposed schemes are secure under our attack model.

II. RELATED WORKS

Numerous secure data aggregation schemes have been proposed. These schemes are methods for different security requirements. [4] to prevent the redundant data transmission from sensor nodes to cluster-heads by using ESPDA and the use of NOVSF Block-Hopping technique improves the security by randomly changing the mapping of data blocks to NOVSF time slots.

Manuscript published on 28 February 2013.
***** Correspondence Author (s)

Mr. J. John Major*, Electronics and Communication Engineering, Karunya University, Coimbatore, India.

Prof. Shajin Prince, , Electronics and Communication Engineering, Karunya University, Coimbatore, India.

Mr. Akuluri Rakesh, Electronics and Communication Engineering, Karunya University, Coimbatore, India.

Thus the advantages are redundancy rate increases, ESPDA bandwidth occupancy decreases and also NOVSF Block-Hopping technique that provide data communication security. The demerits of this technique is symmetric keys used in the security algorithms are not transmitted.

[5] to enhance the bandwidth usage and energy utilization by minimizing the transfer of redundant data. To reduce the number of bits transmitted, SRDA desires sensor nodes to send differential data instead of raw sensed data. Thus the advantages are deployment estimation and not performing any online key distribution and also SRDA yields significant savings in the energy consumption while preserving the data security.

[6] several studies attempt to provide confidentiality. That is, an aggregator can openly execute addition operations on encrypted binary data. CDA places more intensity on passive attacks. 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 summary. If sensors within the same cluster encrypt their sensing data with a common secret key, an adversary may decrypt the aggregated cipher text by compromising only one sensor.

[7] to perform a simple and provably secure additively homomorphic stream cipher which allows efficient aggregation of encrypted data. The advantages are influence of compromising a sensor is actually reduced. The disadvantages of the process is rekeying operations for each sensor cause this scheme to be impractical. A synchronization mechanism should be provided.

[8] proposed a data aggregation scheme based on addition homomorphic public-key encryption. It look likes more secure since every sensor stores only public key. The adversary cannot propel the same attack through compromising only one sensor. The adversary can still impersonate other legal sensors to send the forged cipertexts to the cluster head with the same public key. Authenticity of data is not founded.

[9] to show that aggregate Signatures give rise to verifiably encrypted signatures. Similar signatures enable the verifier to test that a given cipher text C is the encryption of a signature on a given message. To support the encrypted signatures are used in contract-signing protocols. The advantages are Key generation, aggregation, and verification requires no interaction. The drawbacks, chances are there to forger the message.

III. PRELIMINARIES

A. *Network Model*

A WSN is controlled by a base station (BS). A BS has large bandwidth, strong computing capability, stable power, and sufficient memory 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, storage, communication capability and computation. 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 better scalability of MAC (medium access control) or routing and efficient energy management etc. Each cluster has a cluster head (CH) responsible for collecting and aggregating sensing data from SNs within the same cluster. A CH sends the aggregation results to the BS. In a homogeneous Wireless Sensor Network, cluster heads act as normal SNs. On the other hand, cluster heads act as by powerful high-end 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. We consider the following three cases. They are outwardly compromising any SN or CH, Compromising SNs, Compromising CHs.

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 exists of four procedures: key generation (KeyGen), encryption (Enc), aggregation (Agg), and decryption (Dec).

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 schema consists of five procedures: key generation (KeyGen), signing (Sign), verifying (Verify), aggregation (Agg), and verifying aggregated signature (Agg-Verify).

The proposed systems are secure because sensing messages are encrypted. In RCDA-HOMO, all sensor encrypts their messages with PBS before transmits In RCDA-HETE, intracluster traffic is encrypted with pair wise keys. Beyond, our design generates the corresponding signature for each sensing data. Consequently, an adversary cannot transform messages and inject forged messages since he cannot sign forged messages without private keys.

If an adversary has the ability to compromise sensors, we allow for the following situations. An adversary can compact a sensor and perform it as a legal one. Detecting compacted sensors that still act normally is infeasible in all existing detection mechanisms in WSN. Also, if the value of a devise message is in a acceptable range, detecting it is still infeasible. An attacker can also try to manipulate the aggregated result. He may generate fake data; modify legal messages or act a part other sensors.

The proposed schemes are still secure against above attacks because of the signature required for each generated message. On the other hand, we examine the situation when an adversary compromises a cluster head in RCDA-HOMO. He cannot decrypt the aggregated ciphertext or each individual ciphertext because no decryption private key is stored in a cluster. The compromised cluster head may selectively drop some ciphertexts and signatures in the Aggregate procedure. This kind of foray which is called selective forwarding attack was described.

IV. PROPOSED WORK

A Recoverable Concealed Data Aggregation (RCDA) is introduced. In RCDA, the base station can recover each sensing data generated by all sensors even if all these data have been aggregated by the cluster heads (aggregators). The RCDA schemes for homogeneous and heterogeneous WSNs are discussed in this paper.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

A. (RCDA-HOMO) Homogeneous WSNs

There are four procedures in RCDA-HOMO: Setup, Encrypt-Sign, Aggregate, and Verify. The first Setup procedure is used to prepare and install all necessary secrets for the BS and each sensor nodes. When the sensor decides to send sensing data to its Cluster Head (CH), it performs Encrypt-Sign and then sends the result to the CH. When the CH receives all results from all its members, it activates and aggregate what it received, and sends the final results (aggregated cipher text and signature) to the Base Station (BS). The final procedure is Verify. Here, the BS first extracts individual sensing data by decrypting all the aggregated cipher text. Then, the BS verifies the integrity and authenticity of the decrypted data based on the aggregated signature.

B. (RCDA-HETE) Heterogeneous WSNs

There are five procedures in RCDA-HETE: Setup, Intracluster Encrypt, and Intercluster Encrypt, Aggregate, and Verify. In the first Setup procedure, all the necessary secrets are loaded to H-Sensor and L-Sensor. The Intracluster Encrypt procedure is involved, when L-Sensors wish to send their sensing data to the corresponding H-Sensor. In the third Intercluster Encrypt procedure, all H-Sensor aggregates the received data encrypt and then signs the aggregated result. Thereafter, if an H-Sensor receives the cipher texts and signatures from other H-Sensors on its routing path, then it activates the aggregate procedure. The final Verify procedure ensures the authenticity and integrity of all aggregated result.

C. Recovery property

The Recovery property provides two functionalities. First, the BS can able to verify the integrity and authenticity of sensing data. Second, the BS can perform aggregation aggregation operations on these data. However, in RCDA-HETE, the BS only recovers individual aggregated result generated by each cluster rather than all sensing data.

V. PERFORMANCE ANALYSIS AND SIMULATION RESULTS

In the proposed work, nodes are randomly deployed. Homogeneous and Heerogeneous network structure is created. Using the recovery property, the base station verifies all the data by the two processes. A network structure is created within the network. The performance of the network is evaluated in terms of parameters such as Throughput, Packet Delivery Ratio (PDR), and Delay parameters, Network structure created is defined as follows,

a) **Throughput** is the average rate of successfully transmitted data packets over the communication channel.
b) **Packet Delivery Ratio (PDR)** is defined as the ratio of the total number of successfully transmitted data packets to the total number of data packets sent from the source to the destination.
c) **Delay** is the time taken for a packet to travel across a network from source to destination.
d) **Processing delay** indicates the execution time for sensors to produce cipher texts and corresponding signatures before transmission.
e) **Aggregation delay** is also evaluated by measuring time spent on processing time on aggregating cipher texts and signatures in the proposed schemes.

In the simulation three different cases are considered and they are,

a) Nodes in static condition
b) Nodes in partially mobile state
c) Nodes in fully mobile state

Throughput is the average rate of successful message delivery over a communication channel. The throughput is measured in bits per second and in data packets per second. The system throughput or throughput is the sum of the data rates that are delivered to all stations in a network.

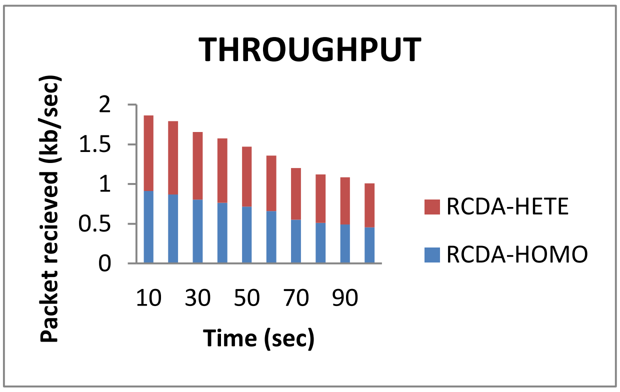

Fig.1 Packet Received versus Time.

Fig.1 shows that the throughput of the data is increased in the RCDA-HETE than in the RCDA-HOMO because the number of active nodes in the transmission is comparatively less in the case of the RCDA-HETE.

The **packet delivery ratio (PDR)** is defined as the ratio of the number of packets received by the destination and the number of packets transmitted by the source. Packet delivery ratio is plotted against number of mobile nodes in Once the sender has the receiver's address, it can direct the packet. The way the MAC header of that packet is addressed, depends on the network's topography. It depends on whether the source node and the destination node are separated by a router. This ratio directly affects the maximum throughput that the network can support. Packet delivery ratio increases with an increase in time for stable nodes.

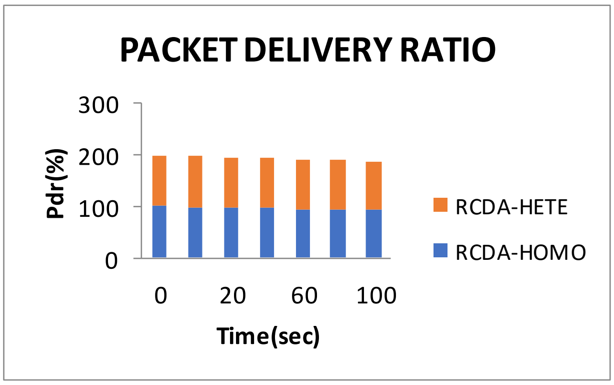

Fig.2 Packet delivery ratio with number of nodes

Fig.2 shows that the packet delivery ratio in the homogeneous network is greater than the heterogeneous network because the homogeneous network has low throughput which means that the packet delivery rate is less. The number of packets received at the destination, successful or unsuccessful packets are received.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

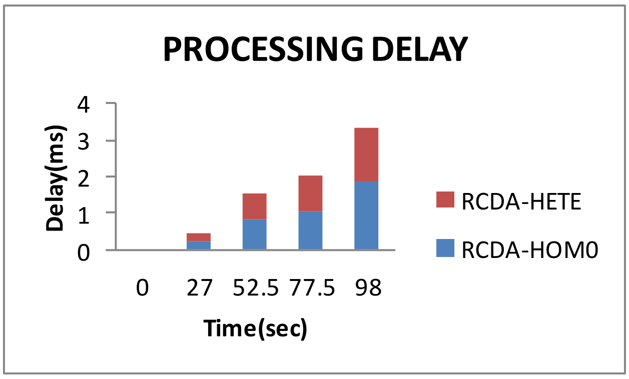

Fig.3 Delay variation versus Time

Processing delay indicates the execution time for sensors to produce cipher texts and corresponding signatures before transmission. From the above graph (Fig. 6.3), it is seen that delay in the homogeneous network is greater than the heterogeneous network because the execution time is more in the homogeneous networks. On comparing the both networks, heterogeneous network has better performance.

Aggregation delay is also evaluated by measuring time spent on processing time on aggregating cipher texts and signatures in the proposed schemes.

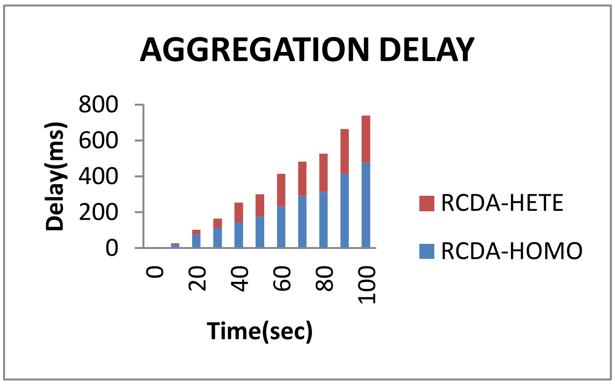

Fig.4 Delay variation versus Time

From the above graph (Fig. 4), it is seen that delay in the homogeneous network is smaller than the heterogeneous network because the execution time is more in the heterogeneous networks. The last delay, decryption delay, is not designed since the base station is considerably powerful as a workstation. Therefore, this delay is minute and can be ignored.

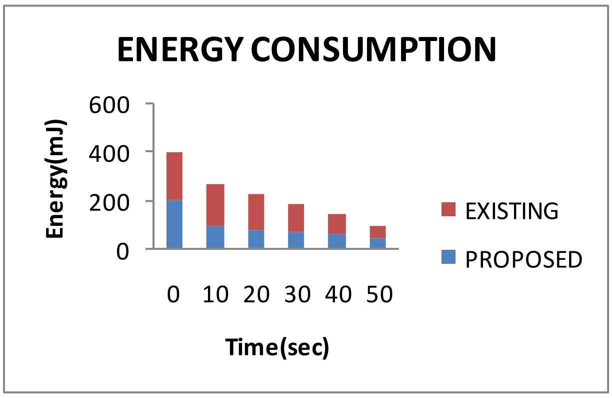

Fig.5 Energy versus Time

Fig.5 has been clearly stated the **energy consumption** in the existing scheme is greater than the energy consumption in the proposed scheme because the number of nodes involved in the transmission is less in proposed scheme. Wireless sensor network consists of large amount of sensor nodes, which are compact, light weighted and battery powered devices that can be used virtually in any environment. Sensor nodes must conserve their energy by all means and stay active in order to maintain the required sensing coverage of the environment. Hence by using this proposed network, the number of involved in the transmission of data is minimized and hence the energy will be conserved.

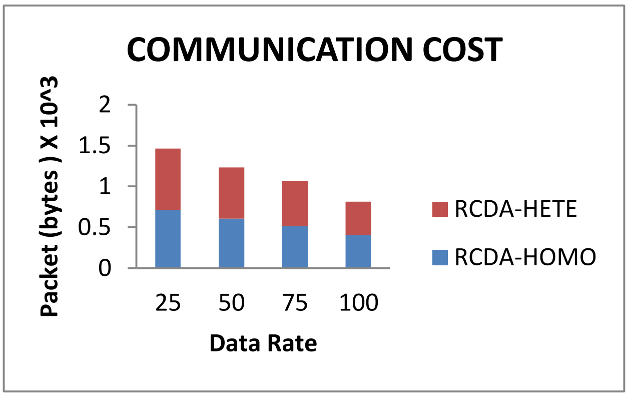

Fig.6 Packet versus Data Rate

Fig. 6 shows that **communication cost** of the RCDA-HOMO is greater than the RCDA-HETE. The cost of the sensor in the real time is compared with both networks. RCDA-HETE is greater compared to the RCDA-HOMO Even though signatures bring additional costs, the proposed schemes are still acceptable or WSNs after evaluation

VI. Conclusion

The proposed methods are recoverable concealed data aggregation schemes for homogeneous / heterogeneous WSNs. In the networks, cluster heads and sensor nodes are selected based on the energy and secure data communication is provided. A special feature is that the base station can securely recover all sensing data rather than aggregated results, but the communication overhead is still acceptable. The Recovery property attempts to provide two functionalities. First, BS can verify the integrity and authenticity of all sensing data. Second, BS can perform aggregation operations on these data. In the security analysis, demonstrate that the proposed schemes are secure under our attack model. Through experiments, show that the performance of the design is reasonable and affordable. The parameters like throughput, delay, energy consumption, communication cost, and packet delivery ratio are analysed.

REFERENCES

[1] R. Rajagopalan and P. Varshney, “Data-Aggregation Techniques Sensor Networks: A Survey,” IEEE Comm. Surveys Tutorials, vol. 8, no.4, pp. 48-63, Oct.-Nov. 2006.

[2] S. Madden, M.J. Franklin, J.M. Hellerstein, and W. Hong, “TAG: A Tiny Aggregation Service for Ad-Hoc Sensor Networks,” Proc. Fifth Symp. Operating Systems Design and Implementation, 2002G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in Plastics, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

[3] J.-Y. Chen, G. Pandurangan, and D. Xu, "Robust Computation of Aggregates in Wireless Sensor Networks: Distributed Randomized Algorithms and Analysis," IEEE Trans. Parallel Distributed Systems, vol. 17, no. 9, pp. 987-1000, Sept. 2006W.-K. Chen, Linear Networks and Systems (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[4] H. C¸ am, S. O ¨ zdemir, P. Nair, D. Muthuavinashiappan, and H. Ozgur Sanli, "Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks," J. Computer Comm., vol. 29, pp. 446-455, 2006.

[5] H. Sanli, S. Ozdemir, and H. Cam, "SRDA: Secure Reference- Based Data Aggregation Protocol for Wireless Sensor Networks," Proc. IEEE 60th Int'l Conf. Vehicular Technology (VTC '04-Fall), vol. 7, pp. 4650-4654, Sept. 2004.

[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 Computing, vol. 5, no. 10, pp. 1417-1431, Oct. 2006.

[7] C. Castelluccia, E. Mykletun, and G. Tsudik, "Efficient Aggregation of Encrypted Data in Wireless Sensor Networks," Proc. Second Ann. Int'l Conf. Mobile and Ubiquitous Systems, pp. 109-117, July 2005.

[8] E. Mykletun, J. Girao, and D. Westhoff, "Public Key Based Cryptoschemes for Data Concealment in Wireless Sensor Networks," Proc. IEEE Int'l Conf. Comm., vol. 5, pp. 2288-2295, June 2006.

[9] D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and Verifiably Encrypted Signatures from Bilinear Maps," Proc. 22nd Int'l Conf. Theory and Applications of Cryptographic Techniques (Eurocrypt), pp. 416-432, 2003.

[10] Chien-Ming Chen, Yue-Hsun Lin, Ya-Ching Lin, and Hung-Min Sun "RCDA: Recoverable Concealed Data Aggregation for Data Integrity in Wireless Sensor Networks" IEEE transactions on parallel and distributed computing, VOL. 23, NO. 4, APRIL 2012

Mr. J. John Major received his B.E degree from JACSI College of Engineering, Anna University in the year 2011. Currently pursuing his master's degree in communication systems. His area of interest includes Wireless Sensor Networks, Wireless Mess Networks and Antenna Design. Currently working on trust management in Wireless Sensor Networks.

Mr. Shajin Prince received his B.E degree from Karunya University in the year 2008. He received his M.tech from Karunya University in the year 2010. Currently working as Assistant Professor in Karunya University. He is also a part time research scholar in Anna University. His research area includes Audio signal processing and Multimedia compression.

Mr. Akuluri Rakesh received his B.E degree from Karunya University in the year 2010. Currently pursuing his master's degree in communication systems. His area of interest includes Wireless Sensor Networks, Wireless Mess Networks and Antenna Design. Currently working on security and providing high throughput in Wireless Mesh Networks.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)