Wireless and Mobile Technologies, 2013, Vol. 1, No. 1, 37-41
Available online at http://pubs.sciepub.com/wmt/1/1/7
© Science and Education Publishing
DOI:10.12691/wmt-1-1-7

Cb-SDA: Cluster-based Secure Data Aggregation for Private Data in WSN

Ajay Jangra[1,*], Priyanks[2], Richa[1]

[1]CSE department, University Institute of Engineering and Technology, kurukshetra University, Kurukshetra, India
[2]ECE department, University Institute of Engineering and Technology, kurukshetra University, Kurukshetra, India
*Corresponding author: er_jangra@yahoo.co.in

Received July 30, 2013; Revised October 18, 2013; Accepted December 17, 2013

Abstract As wireless senor network (WSN) has broad range application that needs privacy of sensed data while they transmit from source to base station. Providing robust and reliable data aggregation scheme with securing sampled data is a challenging problem in WSN. This paper discusses about secure data aggregation and proposed the new secure cluster based aggregation of private data scheme using the LEACH protocol. The proposed scheme (Cb-SDA) is based on the additive property of complex number to aggregate the sensor data in order to provide the privacy during their transmission over the base station and provide better security for wireless sensor network routing with efficient performance in terms of energy consumption, throughput, delay, bandwidth utilization, jitter etc. Simulation study evaluates the performance of proposed scheme and calculates aggregation privacy, communication overhead and accuracy in terms of throughput and compare to existing approach.

Keywords: *WSN, base station, clustering, data management in WSN, aggregator, SUM, opponent*

Cite This Article: Ajay Jangra, Priyanks, and Richa, “Cb-SDA: Cluster-based Secure Data Aggregation for Private Data in WSN.” *Wireless and Mobile Technologies* 1, no. 1 (2013): 37-41. doi: 10.12691/wmt-1-1-7.

1. Introduction

WSN consists of the upcoming technologies that have obtained remarkable consideration from the research community. The objectives of these networks are monitoring the physical environment, gather and transmit the information to other sink nodes. Typically the radio transmission range of the sensor network is several orders of magnitude which are smaller than the geographic extent of the intact network. Therefore, the data must be transferred hop-by-hop towards the sink in a multi-hop fashion. Energy consumption within the network will be reduced if the amount of data to be relayed is reduced [1]. Wireless sensor network contains an excellent range of minute electromechanical sensor devices that possess the sensing, communication and computing capabilities. These devices are typically used for accumulating sensory information such as measurement of temperature from an extended geographical space [2]. Many of the features of the wireless sensor networks give rise to challenging problems [3]. Sensors are typically power and resource constrained and suffering from restricted computation, power and communication resources. Sensors can offer engrained raw data as a result of these limitations data aggregation is crucial consideration for sensor networks. Data aggregation is taken as one of basic distributed data processing methods for conserving the energy and reducing the medium access layer contention in wireless sensor networks [4]. Data aggregation is presented as an important model for routing in the wireless sensor networks. The fundamental idea is to integrate the data from different sources, then reroute it with the removal of the redundancy and thereby minimizing the number of transmissions and conserving the energy [5]. The inbuilt redundancy in the raw data collected by different sensors can be avoided by the data aggregation. In addition, these operations use raw materials to obtain information specific to the application. To save the energy within the system for maximizing lifetime of the network, it is necessary for the network to sustain high level of the data aggregation [6]. As both data aggregation and security are vital for wireless sensor networks but providing secure data aggregation has been an noteworthy problem for researchers. Therefore, providing a reasonable guidance for building systems which execute private data aggregation is advisable. Basically two forms of security concern employed in it such that internal security and external security. Internal security is also called data privacy in which maintaining the privacy of sensor node from another participated senor node in the same network. On the other hand, external security is also called data security which protects them from recovering sensitive data. Generally, two methods can be employed for secure data aggregation in WSN; hop by hop encrypted data aggregation and end to end encrypted data aggregation or link level data aggregation [7].

A. Hop-by-Hop encrypted data aggregation.

In this method, an encryption of the data is accomplished by the sensing nodes and decryption by the aggregator. The aggregator aggregates the data and again encrypts the aggregation result. Finally, the BS obtains the final encrypted aggregation result and decrypts it.

B. End to End encrypted data aggregation or link-level data aggregation.

In this method, the aggregator nodes will only accomplish aggregation on the encrypted data and in between have not any decryption keys. It is renowned that end-to-end data encryption can capable to protect private communications between the data source and data sink as long as the same parties have agreement on encryption keys. Nevertheless, end-to-end encryption alone is not a beneficial candidate for private data aggregation.

2. Previous Work

A lot of work has been done on secure data aggregation in wireless sensor network that offer secure communication with trusted sensor. But still there may be likelihood that an opponent can compromise cryptography and modify the data. In [7] has classified the security problems, data integrity and confidentiality in data aggregation into two cases: end-to-end and hop-by-hop encrypted data aggregation. They have conjointly recommended two frame works for the above methods. The first method is end-to-end encrypted data aggregation has more computation cost on the sensor nodes however attains better security as compare to the framework for hop-by-hop encrypted data aggregation. In [8] has represented two data aggregation schemes to preserve the privacy for additive aggregation functions. Their very first scheme is CPDA (Cluster-based Private Data Aggregation) that provides the clustering mechanism and algebraic properties of polynomials. Their second scheme is SMART (Slice-Mix-Aggregate) that completely based on slicing techniques and also the associative property of addition. The aim of their work is to maintain the gap between collaborative data assortment by WSN and data privacy. They valuated the above schemes by privacy-preservation effectiveness, data aggregation accuracy and communication overhead. The Simulation results exhibit the effectiveness and efficiency of our schemes but the bandwidth consumption is more in the case of their proposed SMART technique. The authors [9] have presented SEA (a Secure Encrypted-data Aggregation) scheme in mobile wireless sensor networks i.e. MWSN environment. The design of data aggregation removes redundant sensor readings without using encryption and maintains data privacy and secrecy during transmission. As compared to traditional schemes, the proposed scheme leverages privacy and security and redundant instances of original readings will be summarized into a single packet. Hence, greater energy can be conserved. However there is no discussion of integrity in the proposed SEA scheme. This paper [10] has represented privacy-preserving data aggregation scheme for additive aggregation functions. The concentration of their work is to keep up the gap between collaborative data summarization by wireless sensor networks and data privacy. They have showed simulation results of their schemes and compared their performance to a traditional data aggregation scheme TAG in which no data privacy is provided. Simulation results show the efficiency and efficacy of their proposed schemes. However, because of additive properties employed in aggregating data, the communication overhead increases and become more complicated. In the paper[11] the author proposed energy efficient secure data aggregation protocol which provides the authentication and security to maintain the efficiency of aggregation. For this they divide the network into cluster. Simulation results showed that protocol has reduced energy consumption and attaining good packet delivery ratio. Therefore from the above previous work the major problem in PDA paper [8] in which communication overhead is high because of unnecessary messages generated during transmission. In order to solve this problem, this paper propose an efficient and a new cluster based Secure data aggregation of private data which uses algebraic property of complex number with arithmetic operation and make use of aggregation of sensor data with suitable privacy i.e. hide data from attacker and finally transmit data to the base station. Although private data can be overheard and decrypted by an attacker over sensor network however our propose scheme will protect sensitive information and propose scheme is built on the top of cluster leach protocol which support security and privacy using the same scheme.

3. Problem Statement

In this paper, Sensor Networks is represented as a connected graph G (v, e), where sensor nodes are modeled as the set of vertices v (v ЄV) and wireless links act as the set of edges e (e ЄE).If there exist edge between two nodes then they able to communicate in connecting graph. Data aggregation function is outlined as x(t)= f(d1(t), d2(t), dN(t)), where di(t) is the individual sensor reading at time t for node i. Typical functions of ‘f’may be sum, min, average, count and max . If d_i(i = 1, ,N) is given, the computation of x at a query server or data sink is feasible. However, because of the large data traffic, power consumption and bandwidth constraints, data aggregation techniques are required to preserve power and resources. In this paper, we concentrate on additive aggregation functions of complex number. Rather than using additive aggregation functions, the many other aggregation functions such as count, min, average, variance, max, standard deviation and any other measured data maybe used just like additive aggregation function sum. Here the scheme is using three varieties of nodes in the network such as Base station (BS), cluster head or aggregator node and source node. The BS is responsible for answering the queries and includes aggregation result. Here we are considering only one BS. Cluster based protocol consist of clusters made of multiple senor nodes and ultimate connected to BS. Cluster leaders act as aggregator node or intermediate aggregators which are responsible for forwarding queries and combining answers from their respective members and ultimately aggregated result to the BS.

4. System Design: Security Requirements

Protecting the data privacy in several WSN applications is a large concentration. The subsequent criteria summarize the fascinating characteristics of a private data aggregation scheme:

A. Privacy.

Every sensor in the network is made up of data that is only known to the sensors within the network. The private data aggregation scheme should be capable of handling only some extent of attacks and collusion among compromised nodes.

When malicious attack is applied to sensor network, it is possible that some nodes may collude to uncover the private data of other node. Moreover, wireless links can be intercepted by attackers to reveal private data. A good private data aggregation scheme must be robust to such attacks.

B. Accuracy.

The reliable aggregated data of sensor nodes is acceptable only when no other sensor node within the network have any knowledge about the value of any independent or single sensors. Accuracy is the unit for measuring the performance of secure data aggregation schemes. We can define accuracy by calculating the performance parameters such throughput and compare them with the existing approach to show the effectiveness of proposed scheme.

C. Efficiency.

The objective of data aggregation is to minimize the number of messages transmitted within the sensor network, thus minimize resource and power usage. Data aggregation attains bandwidth efficiency by using in-network processing. In this scheme, additional overhead is used to protect privacy. But a good private data aggregation scheme should maintain small overhead In brief, we have to calculate the overall network overhead for proposed approach.

5. Cluster Based Secure Data Aggregation (CB-SDA): the Prposed Algorithm

This paper present a new and efficient cluster based secure data aggregation for private data scheme which provide privacy and aggregation of sensor reading in WSN. To hide the private data from an opponent and other trusted sensor nodes at the time of transmission to the BS, our propose approach use algebraic equation by employing complex number. In other word, we use additive expression of complex number for hiding sensitive information and perform aggregation. As there are variety of aggregation function such as MIN, MAX, SUM, AVG, COUNT, STANDARD DEVIATION and many other moment of measured data but here we are using only SUM aggregation function for sake of simplicity like in [8]. To obtain the summarize reading of all sensor BS can use efficient cluster based leach protocol for disseminating the aggregated data.

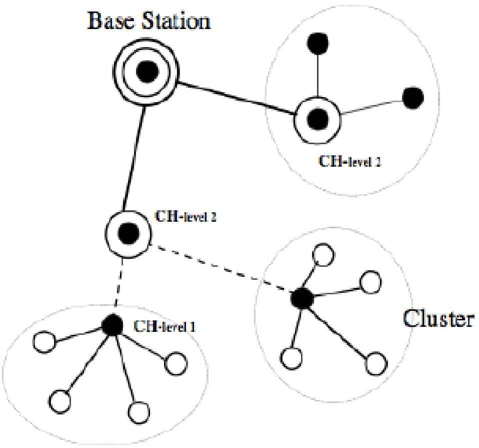

Figure 1. Cluster model

As depicted in figure1. Scheme takes general aggregated cluster model which is used for data collection such as target measuring and temperature measuring. This cluster model has following features: it has powerful BS which located anywhere in the network. A large number of sensor nodes with resource-constrained are distributed uniformly in network space and has different level based on hop count from BS. Each node can able to sense, aggregate and forward data to the BS. And finally it can switch into sleep mode when nodes do not require transmitting data in order to save energy.

Here an algorithm assumes that all sensor nodes use two types of share keys. The first type of key is shared between cluster head and BS and second symmetric key pair shares between all source nodes and their cluster head for data security. As sensor node has resource- constraint in term of memory, speed and power but we should be careful design an algorithm for wireless sensor networks so that execution of algorithm must consume less power and other resources. For this purpose, we are proposing efficient algorithm for SUM aggregation function. The Parameters for the algorithm are given in Table 1.

Table 1.

Parameters	Description
n	Node(1,2…n)
sd	Sampled data
Mask	Data sampling by sensor node is called masked when it is combined with a private real number
rs	Private real number
GenCmpxNum()	Function used for joining real no. with imaginary no. to obtain complex no
K_{ij}	K_{ij} is symmetric key for I and j sensor nodes
Ck	Cipher text
Enc	Encryption of data
Drc	Decryption of cipher data
SUM1 & SUM2	Sum value
Sisjoin	Function for separating real and imaginary number

Algorithm: CB-SDA (*Cluster based Secure data aggregation*) for private data.

Input: Use aggregation function SUM.

Output: Aggregation result using SUM.

```
Step 1. Create customized data from the data of the source nodes.
For n sensor nodes
        Sense   sd
        Mask    (sd, rs)
                A = sd+sr
        GenCmpxNum (A+Bi)
                $X_1$ = A+Bi
        Enc ($K_{i,j}$, $X_1$)
                $C_k = E_r (X_1)$
Transmit ($C_k$)
Step 2. Apply additive property of complex number on aggregator (cluster head) to obtain sum value or intermediate value.
    For every intermediate aggregators
        Drc ( (Kj, i, ($C_k$) )
        Add ()
$X' = X_1 + X_2$ // assuming aggregate data of two nodes
        Enc($K_{j,I}$, ($X'$) )
        $C_d = E_r(X')$
Transmit ($C_d$)
Step 3. Compute aggregation result at BS.
    Receive($C_{ipher}$) // ciphers of intermediate results
    For all ($C_{ipher}$)
    Drc( (K, ($C_{ipher}$) )
        Add()
$SUM_2 = X+X''$ // X is intermediate result and $X''$ reading of BS
Step 4.Extract actual SUM of the sensor from $SUM_2$ at the BS.
        Disjoin ($SUM_2$)
Take real unit of SUM and compute $SUM_1$ for all real seeds
        SUM= $SUM_2$-$SUM_1$
        Returning SUM value
```

A. Working of Algorithm.

The algorithm for Sum aggregation function is given above. In the **first** step, when BS receives the query with SUM from the user, it broadcast them to the whole wireless sensor network. The sampled data said to datum 'sd' of each sensor is first combine with private real number 'rs' to make it concealed within A. Next, function GenCpxNum () is used for joining private real data with imaginary expression Bi and make complex number X. After establishing the shared secret key with BS, BS gives the unique real and imaginary number to all sensor nodes. Hence both numbers are remaining private in sensor network. Then source node encrypts data using symmetric key and send cipher data to the aggregator node or cluster head. In this way at the end of first step our algorithm convert sample reading into complex expression and hence make the data private. In this scheme, any opponent, aggregator node and neighbouring node cannot recover the sensitive sampled data by using decryption. Therefore it can maintain security over existed scheme. In **second** step, when aggregator receives cipher data, it firstly converts into original form and add complex number to make it complex expression and it to the BS for further processing. Up to this step, a partial aggregation has been performed on private data. In the **third** step, BS unites all the intermediate information and decrypts them by using shared key between BS and cluster leader. Moreover, in order to generate the resultant i.e. SUM_2 it will do final aggregation by applying its own private data to the partial aggregate decrypted value. In **last** step, BS separate the real and imaginary part as received SUM_2 is combination of real and imaginary expression. Because original reading is concealed with real private unit so it is the responsibility of the BS to extract the original information by subtracting SUM_1 which is sum of real private data of sensor nodes from SUM_2 and return the resultant SUM to the user who request for the same query.

6. Simulation and Performance Analysis

We use cygwin setup for carrying out the simulation using ns-2 of our proposed work. In this research work we presented the new secure data aggregation scheme over the LEACH protocol. In our proposed cluster based scheme provides results into the better security for wireless sensor network routing with efficient performance in terms of throughput, jitter and communication overhead etc. For our simulation study we carried our working with LEACH protocol and compare the results between existing LEACH and proposed cluster based LEACH protocol. Existing LEACH do not having the mechanism of privacy preservation, hence we are considering the same things as baseline to design the proposed privacy preserving security scheme for data aggregation. Depending on the simulation studies and results, we evaluated following performance parameters:

A. Privacy.

In the proposed scheme, source node encrypts data using symmetric key and send cipher data to the aggregator node or cluster head. In this way at the end of first step our algorithm convert sample reading into complex expression and hence make the data private. In this scheme, any opponent, aggregator node and neighbouring node cannot recover the sensitive sampled data by using decryption. Therefore it can maintain security over existed scheme.

B. Communication Overhead.

Proposed scheme use information-hiding techniques and encrypted communication to guard data privacy. This leads to some communication overhead. In order to find out bandwidth efficiency of this scheme, we tend to simulated here proposed approach in NS2 using LEACH protocol. We did extensive simulations and collected results to compare this proposed scheme with existing LEACH. In our experiments we consider networks with 15 (10 common nodes and 2 cluster nodes) sensor nodes. These sensor nodes are arbitrarily scattered over a 30meters X 30meters area. The transmission range of a sensor node is 10 meters and data rate is 1 Mbps.

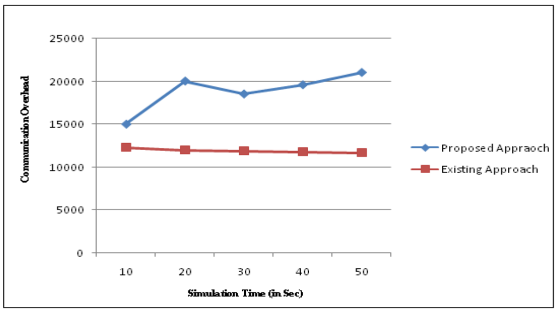

Figure 2. Communication Overhead for proposed and existing approach

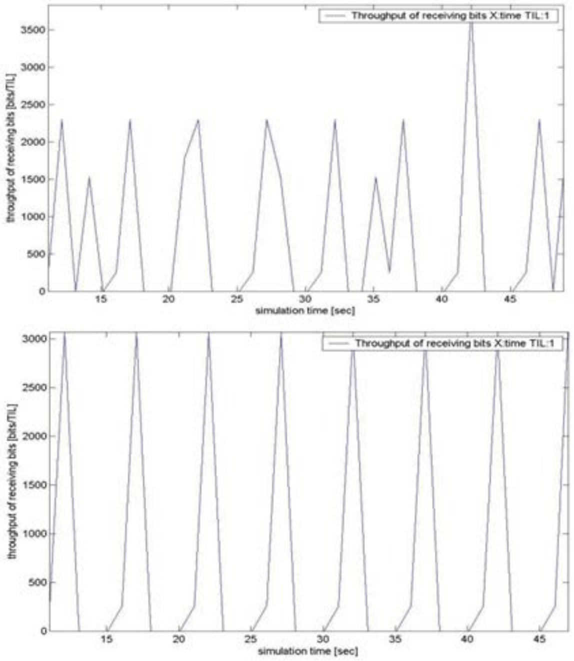

Figure 3. Throughput of Receiving Bits vs. Simulation Time of proposed and existed scheme

C. Aggregation Accuracy.

In ideal situations when there is no data loss in the network, proposed scheme should get 100 percent accurate aggregation results. But in WSN's because of processing delays and collisions, sometime messages may get dropped or lost. Hence this affects the data aggregation accuracy. We calculate accuracy on the basis of throughput. A higher throughput value means the collected sum using the specific aggregation scheme is more accurate.

As shown in simulation graph, throughput of the proposed scheme is increased as compared to the existed one.

7. Conclusion

This paper proposes a new and reliable scheme for securing the aggregate data where an environment produces sensitive data. The proposed scheme uses the additive property of complex numbers when the observed data collected are first converted to a complex number in the form before forwarding them to the query to the server. Therefore, it will defend private information stored in the sensor node from being known by its neighboring nodes including data aggregators in WSN. Additionally, it is still not feasible for opponent to recover the sensed data due to distinctive security protection technique. Since our scheme is constructed on the top of the existing reliable key management scheme and offer a way for information privacy by using complex numbers and believe that both of them work together to provide basic security properties like access control, confidentiality and message integrity along with preserving information privacy. Through the simulation analysis performance this paper has shown that our scheme is more efficient in terms of communication overhead than the existing work.

References

[1] Dorottya Vass, Attila Vidacs, "Distributed Data Aggregation with Geographical Routing in Wireless Sensor Networks", Pervasive Services, *IEEE International Conference* on July 2007.

[2] Jukka Kohonen, "Data Gathering in Sensor Networks", Helsinki Institute for Information Technology, Finland. Nov 2004.

[3] Gregory Hartl, Baochun Li, "Loss Inference in Wireless Sensor Networks Based on Data Aggregation", IPSN 2004.

[4] Zhenzhen Ye, Alhussein A. Abouzeid and Jing Ai, "Optimal Policies for Distributed Data Aggregation in Wireless Sensor Networks", Draft Infocom2007 Paper.

[5] Bhaskar Krishnamachari, Deborah Estrin and Stephen Wicker, "The Impact of DataAggregation in Wireless Sensor Networks", *Proceedings of the 22nd International Conferenceon Distributed Computing Systems,* 2002.

[6] Kai-Wei Fan, Sha Liu, and PrasunSinha, "Structure-free Data Aggregation in SensorNetworks", *IEEE Transactions on Mobile Computing,* 2007.

[7] Yingpeng Sang, Hong Shen, Yasushi Inoguchi, Yasuo Tan and Naixue Xiong, "Secure DataAggregation in Wireless Sensor Networks: A Survey", *Seventh International Conference onParallel and Distributed Computing, Applications and Technologies*, 2006.

[8] Wenbo He, Xue Liu, Hoang Nguyen, Klara Nahrstedt and Tarek Abdelzaher, "PDA: Privacy preserving Data Aggregation in Wireless Sensor Networks", 26th IEEE International Conference on Computer Communications. IEEE INFOCOM 2007.

[9] Shih-I Huang and Shiuhpyng Shieh, "SEA: Secure Encrypted-Data Aggregation in Mobile Wireless Sensor Networks", *International Conference on Computational Intelligence and Security* 2007.

[10] Prakash G L, S H Manjula, K R Venugopal and L M Patnaik, "Secure Data Aggregation Using Clusters in Sensor Networks", *International Journal of Wireless Networks and Communications*, Volume 1, Number 1 (2009), pp. 93-101.

[11] V. Bhoopathy, R.M.S parvathi "Energy Efficient Secure Data Aggregation protocol for Wireless Sensor Network" European journal of scientific research, Vol.50 No. 1(2011), pp 48-58.

[12] S. Madden, M. Franklin, J. Hellerstein, "TAG: a Tiny AGgregation Servicefor Adhoc Sensor Networks," in *in Proc. of the 33rd InternationalConference on OSDI,* December 2002.

[13] N. Alon, R.M. Karp, D. Peleg, and D. West, "A Graph Theoretic Game and Its Application to the K-Server Problem," in *SIAM J. Computing*, vol. 24,1995.

[14] Ajay Jangra, Richa, Swati, Rajesh Verma, "Vulnerability and security analysis of wireless sensor networks" Indian Journal of Applied Research and Engineering, 4 January, 2011 in IJARE.

[15] ArijitUkil, "Privacy Preserving Data Aggregation in Wireless Sensor Networks" IEEE Sixth International Conference on Wireless and Mobile Communications Sixth International Conference on Wireless and Mobile Communications Sixth International Conference on Wireless and Mobile Communications, 2010.

[16] Julia Albath, Sanjay Madria "Secure Hierarchical Data Aggregation in Wireless Sensor Networks" WCNC 2009 IEEE.

[17] JaydipSen, "A Robust and Secure Aggregation Protocol for Wireless Sensor Networks" 2011, sixth IEEE International Symposium on electronic design, test and application.