International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 8, pp: 344-350 www.ijaem.net ISSN: 2395-5252

A Study on Confidentiality and Authenticity in Data Aggregation for Wireless Sensor Network

Akshatha Y* and Dr. A.S Poornima†
Department of CSE, Siddaganga Institute of Technology Tumkur, Karnataka, India

Date of Submission: 10-10-2020 Date of Acceptance: 30-10-2020

ABSTRACT—In resource constrained wireless sensor networks, data transmission consumes more energy, to improve the overall network lifetime we need to decrease the data transmission. Towards this many mechanisms are proposed in the literature.In- network processing is one such popularly used mechanism which reduces number of data transmissions by computing the required aggregate within the network. For many applications secure data aggregation is very much needed, therefore secure data aggregation schemes are very much in demand in wireless sensor networks. In this paper we propose a lightweight secure data aggregation scheme. The scheme computes the average temperature of the monitoring area and the sensed temperature is aggregated within the network and securely delivered to central controlling node called Base Station. For security, we use mechanisms which are lightweight in terms of computation. The scheme ensures confidentiality of the aggregated value and also the authenticity of the participating nodes.

Keywords—Wireless Sensor Network, Data Aggregation, Au- thentication, Confidentiality

I. INTRODUCTION

Wireless sensor networks (WSN) are mainly used for sens- ing and collecting the data from deployed environment and forwarding it to the Base station (BS). A WSN consists of sensors which are small and resource constrained devices, here main focus is to optimize the energy utilization. In order to achieve energy optimization, one should focus on minimizing the amount of data transmitted and received within the network. In applications where the BS does not require individual sensor readings, in-network data aggregation can be used to elevate the network lifetime.

Data aggregation is defined as the process of aggregating the sensed data from multiple sensor nodes. Data aggrega- tion eliminates redundancy and transmits aggregated data to BS. Data aggregation fuses the data from several nodes at intermediate nodes and forwards aggregated data to BS. It is one of the widely used techniques in WSN to improve the network lifetime. Data aggregation is important, as WSN is resource constrained in nature. Secure data aggregation plays an important role in applications like military, agriculture and temperature monitoring. Area monitoring is one such important application of WSN. Let us consider an agriculture application where WSN is deployed for monitoring various parameters of the field. In this application sensed and aggre- gated data is forwarded to BS. Based on the received data BS has to take certain actions (like quantity of menure to be released, water to be regulated etc.,). In such application if security is not addressed a malicious node may insert its own values resulting in unexpected values received by BS forcing it to take some action when it is not required. Such false actions may create adverse effect in the overall functioning of the system. Therefore there is a strong requirement for secure data aggregation.

In the literature we have many articles addressing data aggre- gation/secure data aggregation schemes [1] - [2].

In the next subsection we going to discuss few of the schemes in detail. Here we are trying to design a lightweight secure data aggregation scheme for the following type of application. Here we consider an application of temperature monitoring in particular we are interested in knowing the average tempera- ture of the monitoring area. In order to improve the network lifetime in-network processing is adapted. Instead of every node forwarding sensed temperature to BS, it's aggregated at intermediate nodes. Based on the computed average tem- perature, BS has to take some action. In this application, if malicious node is successful in injecting its own value for calculating the average temperautre or is able to alter the values of other genuine nodes resulting

value at the BS may be an incorrect one. This incorrect value may force BS to take an action which is actually not required. Also outcome of such action may be costly. Therefore for such applications we need secure data aggregation algorithm.

In this paper, we propose a secure data aggregation scheme which address confidentiality and authentication. Here, confi- dentiality ensures that malicious node cannot alter the sensed values forwarded by genuine nodes. The authentication part of the algorithm resticts any malicious node from participating in the aggregation process. We also discuss about performance of the algorithm in terms of computation, communication and storage cost.

The rest of the paper is organized as follows:Existing schemes from the leterature are explained in Section II. In Section III Network model considered and various notations used are discussed. The proposed secure data aggregation algorithm is elaborated in Section IV. A brief discussion on security and performance analysis of the proposed algortihm is presented in Section V. Finally we conclude in Section VI.

II. LITERATURE SURVEY

We briefly discuss on the reviews of existing work on secure data aggregation in WSN.
In [1], Gaukhar Yestemirova and Sain Saginbekov addressed the data aggregation problem with multiple sinks is WSNs and also proposed an algorithm which reduces transmission of the number of data packets during data collection. This paper provided the solution for data aggregation problem by reducing the amount of data transmission. The total number of redundant packet transmission is reduced by allowing a few nodes to send more than once.

In [3], the authors have proposed energy-efficient data aggre- gation approach, the authors of this paper derived two network- specific algorithm, one algorithm efficiently clusters the net- work and other algorithm aggregates the data from sensing node to data node. Data node means the node which is present within the network, which can be connected to centralized storage like cloud. An intermediate data node must collect the data. Sensing nodes are clustered and every cluster is attached with the number of data nodes. Data node is selected on size and varsity. The clustering node is based on the LEACH (Low-Energy Adaptive Clustering Hierarchy) algorithm and transmission of data is based on LEDA (Low Energy Data Aggregation). This approach enhances the efficiency to a higher extent. Data transmission and aggregation done in terms of power and provides change in energy. A framework for aggregation for WSN based on tree formation with assump- tions for sensing priority and distributed nature is proposed in [4]. The algorithm proposed called Semi Distributed Heuristic Energy Efficient Aggregation Tree algorithm for WSN. This algorithm is a source independent aggregation protocol tree that is created by a semi-distributed manner which depends on the best first search technique. After tree creation, nodes are deployed. The table maintained at every node and values are updated when the node fails or when energy vanish at some aggregators.

This algorithm works in two phases:
- Formation of tree
- Physical data collection at aggregators

The proposed algorithm is for a heuristic distributed tree for WSN, which combines expandability of depth-first search and condensation of breadth. Formation of tree from sink node. Divide the sensor nodes around the sink in two sets: sensing set and communication set. Priority issues are combined to get the spanning tree in WSN. The technique is semi while it starts from the sink. Every node responsible for tree construction.

The authors of [2] proposed an algorithm called Secure and Concentric circle Itinerary based Data Algorithm (SCIDA), uses a secure channel to assure data privacy and reduces energy consumption which is because of encryption operation. The main advantage of SCIDA is, it doesn't require an encryption algorithm during data aggregation, which results in less consumption of energy and also extends the lifetime of the network. SCIDA extends energy efficiency that SCIDA does not maintain any network infrastructure. Data transmitted by a secure channel from one node to another node. There will be three categories in one node, like, first node, intermediate node, and last node. SCIDA algorithm consists of five phases,
- Itinerary design phase
- Initialization phase
- Aggregation request phase
- Data aggregation phase
- Final aggregation result returning phase

The authors of this paper proved that SCIDA is secure and scalable with low communication overhead.In [5], the authors proposed secure data aggregation with fault tolerance for WSN provides end to end confidentiality and also fault tolerance which occurs

during aggregation of data. To secure message while communicating in the network, authors used shared cryptography. In this approach, the data is divided into the number of shares and sent that shares through different disjoint paths between pairs of nodes. At the reception side, the original information is reconstructed by aggregating received data. This approach reduces the redundancy to withstand loss of shares, this loss is due to data loss which occurs during transmission data. This scheme ensures data aggregation even in one or more communication link failure or data loss during transmission. Low computational overhead by using graphical masking.

A novel technique called Secure Data Aggregation Model (SDAM) proposed in [6], which ensures secure communica- tion of data with very low cost which is in terms of resources. Encryption is very much required for secure communication. And encryption is done while aggregating data. SDAM con- sists of three modules, operation sub-layer, data aggregation sub-layer, and communication layer. The objective of this pro- posed scheme is to remove the confusion of having encrypted data before aggregation. Aggregation protocols work best on raw data (unencrypted data). The introduced scheme assures energy-efficient communication of data.

In [7], the authors proposed an energy-efficient clustering scheme with aggregation of data to extend the lifetime of the network. This algorithm work in two phases:

- Clustering is done
- Each node selects its cluster head

Every cluster member transmits data to cluster head for aggregating data and encrypts data then forwards data to Base station. RSA (Rivest-Shamir-Adleman) algorithm is used for encryption and decryption of data to obtain secure communi-cation. The proposed approach offers secure data aggregation to obtain the security issues like integrity and confidentiality. Hassan Harb et al. [8] proposed an efficient data aggregation technique for clustering based periodic WSN. To combine data at node level itself, this technique permits the cluster head to remove repeated set of data which are generated by closest nodes by implementing three data aggregation methods. These three methods are based on, set similarity function, one-way ANOVA (Analysis of Variance) model with statistical tests, and distance function. Two-layer algorithm introduced

- at node level
- at cluster head level

The objective of this paper is to compare three methods in terms of energy consumption, data accuracy, and data latency. After removing repeated data collected by each sensor, the authors have introduced three different data aggregation methods allowing cluster head to discard redundant set of data which is generated by the closest node.

A time series model proposed in [9], which is based on data aggregation scheme in WSN. And also introduced energy- efficient clustering and data aggregation for WSN. In this approach, there exist four phases, cluster head, cluster head selection, data aggregation, and maintenance. And the authors have addressed the problem of energy efficiency of data transmission in WSN and cluster-based aggregation method. Data aggregation is studied under three parts, find cluster tendency, find cluster, and secure aggregation. Hence the authors have proved that the clustering algorithm is energy efficient in WSN.

In [10], Secure End to End Data Aggregation in Wireless Sensor Networks (SEEDA) is proposed to assure data privacy with minimum number of data bit transmission. The scheme used hop-by-hop feature for data bit transmission and it ensures end to end data privacy. the proposed scheme used additive homomorphic encryption method to encrypt the data, that allows cipher texts when decrypted results in plain text.

III. NETWORK MODEL

We consider static clustered WSN consisting several sensing nodes S_i and Cluster Heads CH_i and BS. After deployment, each sensing node is associated to a cluster head (CH_i). Cluster information here is similar to the cluster formation discussed in [10]. Each sensing node shares a secret key k_i with its cluster head for secure communication between the sensing node and cluster head. Also every sensing node as well as CH will have their public and private key pairs denoted as PR and PU respectively.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 8, pp: 344-350 www.ijaem.net ISSN: 2395-5252

TABLE I SOME OF THE NOTATIONS USED IN THIS SCHEME

Notations	
S_i	Sensing Node
C	Cluster
H_i	Head
T_i	Temperatu
k_i	re Secret
$T_i \oplus k_i$	key
PRS_i	Temperature encrypted using secret key
PUS_i	Private key of sensing nodes S_i
PR_{CH_i}	Public key of sensing nodes S_i
PU_{CH_i}	Private of cluster head CH_i
$\{T_i \oplus k_i\}_{PRS_i}$	Public key of cluster head CH_i
$\{T_i \oplus k_i\}_{PRCH_i}$	Temperature $T_i \oplus k_i$ is encrypted using private key of sensing nodes PRS_i
n	Temperature $T_i \oplus k_i$ is encrypted using private key of cluster head PR_{CH_i}
	Number of sensing nodes in a cluster

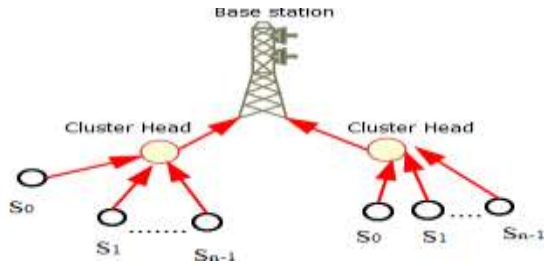

Fig. 1. Static clustered wireless sensor network.

Fig. 1 shows a simple clustered network as considered in this paper, where entire monitoring area is divided into clusters with one cluster head attached to n number of sensing nodes. Here communication pattern is sensing nodes to cluster head and cluster head to base station(BS). The notations used in the paper are depicted in Table 1.

IV. SECURE AGGREGATION ALGORITHM

In this section, we are going to discuss in detail about the scheme proposed for achieving confidentiality and au- thentication in the aggregation process. Here, confidentiality is to protect the secrecy from unauthorized users/attackers and authentication is to ensure that malicious nodes are not participating in the aggregation process. Sensing nodes S_i senses temperature T_i of particular region. Sensed temperature is encrypted using secret key k_i: T_i k_i and is forwarded to CH_i. XOR operation in secret key k_i is used to reduce encryp- tion overhead. As the secret key k_i is known only to sensing node S_i and its respective cluster head CH_i confidentiality is maintained. Further, XORed temperature value T_i is encrypted using private key PRS_i of sensing node: T_i k_i PRS and is forwarded to cluster head CH_i. Encryption using private key PRS_i ensures authentication. The detailed steps are given in Algorithm 1.

V. SECURITY AND PERFORMANCE ANALYSIS

In this section we discuss about the security goals achieved in the aggregation process.

A. Security Analysis

Identification of malicious sensing nodes by respective CH:

- CH receives $\{T_i \oplus k_i\}_{PRS_i}$ as PRSi sensing node as PRSi is known only to Si this ensures indeed the message has come from sensing node Si and it confirms sensing node Si is not malicious which ensures authenticity.
- Ti is also encrypted using ki which is shared between sensing node Si and the cluster head CHi. This en- cryption is possible only by sensing node Si and only respective CH is capable of decrypting the message which ensures confidentiality.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 8, pp: 344-350 www.ijaem.net ISSN: 2395-5252

TABLE II PERFORMANCE ANALYSIS OF THE PROPOSED SCHEME

	Storage	Communication	Computation
Sensing Node	k_i, PU_{S_i} ,PR_{S_i}	1 message sent	1 XOR operation 1 Public key encryption using ECC
Cluster Head	n secret keys k_i n public keys of sensing node PU_{S_i} PU_{CH_i} PR_{CH_i}	n messages received 1 message sent	1 XOR operation 1 Public key encryption using ECC n decryption operation
Base Station	n secret keys k_i m public keys of cluster head PU_{CH_i}	m messages received	m decryption operation

Algorithm 1 Secure Data Aggregation Scheme

1: Every sensing node Si senses the temperature Ti and con- struct the message by encrypting it with secret key ki and further encrypting the message with private key of sensing node PRSi : Ti ki P RS , ki ensures confidentiality and PRSi ensures authenticity.
2: Upon receiving message from sensing nodes, CHi will
decrypt the message using corresponding public key of sensing node PUSi to obtain Ti ki .
3: Now the CHi aggregates the message as TCHi = T1
k1 T2 k2 T3 k3 ..... Ti ki .
4: Aggregated message TCHi by i cluster head is encrypted using private key of cluster head PRCHi and forwarded to BS TCHi PRCH m where m is the number of nodes contributed.
5: Upon receiving aggergated message from cluster head, BS decrypts the message by using respective public key of cluster head PUCHi to get the aggregated message.
6: BS calculate the average temperature as

$$AT = \frac{\sum_{i=1}^{m} T_{CH_i}}{N}$$

private keys PUSi , PRSifor secure communication with CH. CH stores secret key ki where N is the number of nodes contributed across all clusters.
simple XOR operation first to encrypt the sensed temperature value T_i and this is again encrypted using its private key PR_{S_i} before forwarding. The cluster head CH_i aggregates of all the sensing nodes in the cluster, the public and private keys PUCHi , PRCHi of its own and also stores the public keys of sensing nodes PUSi which are indeed required to decrypt the message, as message has been encrypted using private key of sensing nodes. BS stores secret key ki of all the cluster heads and public keys of all the cluster head PUCHi .

Communication: Communication cost is calculated in terms of number of messages exchanged between sensing nodes and CH and also with CH and BS. The number of messages exchanged between sensing nodes and CH is one message sent from each sensing node to CH and CH receives n messages from n sensing nodes, since all sensing nodes transmits the message. The number of messages exchanged between CHs and BS is one message sent from each CHs and BS receives m messages where there can be m CHs within the network that transmits aggregated messages to BS.

Computation: The computation cost is computed based on the additional cryptographic operations performed on the data to achieve security. Each node has to perform one , where N is the number of nodes contributed across all clusters.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 8, pp: 344-350 www.ijaem.net ISSN: 2395-5252

Identification of malicious CH by BS:

- CH aggregates the message which come from sensing nodes in its cluster; $T_1 \oplus k_1, T_2 \oplus k_2, T_3 \oplus k_3, \ldots\ldots, T_i \oplus k_i$ and the aggregated message $\{T_{CHi}\}$ is encrypted using PR_{CHi} the private key of cluster head CH_i.and forwards the same to BS.
- As BS receives $\{T_{CHi}\}, PR_{CH}$ from CH where PR_{CHi} is known only to CH it ensures the message has come from genuine CH. The message is decrypted using PU_{CHi} public key of CH_i.

B. Performance Analysis

In this section, we analyse the proposed scheme with respect to storage, communication and computation cost.

Storage: Here, we discuss the storage required to store secret key which is used for encryption of the message. Each sensing node stores secret key k_i used for encrypting the message. In addition to this, sensing nodes needs to store public and all the received message by performing XOR operation and this aggregated message is again encrypted using its own private key PR_{CHi} before forwarding to BS. BS receives the aggregated message from cluster head CH_i and decrypts using public keys PU_{CHi} of respective cluster heads to obtain original message.

VI. CONCLUSION

Data aggregation is important technique in improving life- time of the resource constrained network like WSN. In this paper we have introduced lightweight scheme for securely aggregating data. The proposed scheme uses XOR operation for encrypting the sensed data which reduces computation at each node. Also authentication is achieved using ECC to reduce the computation. In future we are planning to study the actual computation and communication cost of the scheme by simulating the scheme using the NS3 simulator.

REFERENCES

[1]. G. Yestemirova and S. Saginbekov, "Efficient data aggregation in wireless sensor networks with multiple sinks," in 2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA). IEEE, May 2018. [Online]. Available: https://doi.org/10.1109/aina.2018.00029

[2]. T. Wang, J. Zhang, Y. Luo, K. Zuo, and X. Ding, "An efficient and secure itinerary-based data aggregation algorithm for WSNs," in 2017 IEEE Trustcom/ Big Data SE/ICESS. IEEE, Aug. 2017. [Online].

[3]. Available: https://doi.org/10.1109/trustcom/bigdatase/iceess.2017.268

[4]. S. Manishankar, P. R. Ranjitha, and T. M. Kumar, "Energy efficient data aggregation in sensor network using multiple sink data node," in 2017 International Conference on Communication and Signal Processing (ICCSP). IEEE, Apr. 2017. [Online]. Available: https:// doi.org/ 10.1109/ iccsp.2017.8286397

[5]. S. M. Al-Tabbakh, "Novel technique for data aggregation in wireless sensor networks," in 2017 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC). IEEE, Oct. 2017. [Online]. Available: https:// doi.org/ 10.1109/iintec.2017.8325904

[6]. H. Annapurna and M. Siddappa, "Secure data aggregation with fault tolerance for wireless sensor networks," in 2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT). IEEE, Dec. 2015. [Online]. Available: https:// doi.org/ 10.1109/erect.2015.7498982

[7]. M. B. H. Frej and K. Elleithy, "Secure data aggregation model (SDAM) in wireless sensor networks," in 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). IEEE,

[8]. Dec. 2015. [Online]. Available: https:// doi.org/ 10.1109/icmla.2015.116

[9]. M. B. A, K. R. Dayananda, and S. D. H, "Energy efficient clustering scheme with secure data aggregation for mobile wireless sensor networks (EECSSDA)," in 2016 Online International Conference on Green Engineering and Technologies (IC-GET). IEEE, Nov. 2016. [Online]. Available: https://doi.org/10.1109/get.2016.7916636

[10]. H. Harb, A. Makhoul, S. Tawbi, and R. Couturier, "Comparison of different data aggregation techniques in distributed sensor networks," IEEE Access, vol. 5, pp. 4250–4263, 2017. [Online]. Available: https://doi.org/10.1109/access.2017.2681207

[11]. P. K.Hirani, "Energy-efficiency based clustering anddata ag- gregation for wireless sensor networks," vol. 119 – No.21, 2015. [Online]. Available: https:// pdfs.semanticscholar.org/0b0b/ df1637e6d47ca9094fc0716a720542339743. pdf

[12]. A. Poornima and B. Amberker, "SEEDA:

Secure end-to-end data aggregation in wireless sensor networks," in 2010 Seventh International Conference on Wireless and Optical Communications Networks - (WOCN). IEEE, Sep. 2010. [Online]. Available: https://doi.org/10. 1109/ wocn. 2010.5587353