An Efficient Aggregation Method for Wireless Sensor Networks in the Presence of Collusion Attacks

Sudheer K. T[#1] and Neethu Francis[*2]

[#] *PG Scholar, Department of CSE, KMP College of Engineering, Odakkali, Kerala, India.*

[*] *Assistant Professor, Department of CSE, KMP College of Engineering, Odakkali, Kerala, India.*

***Abstract*— In WSN the computational power of very low power processors dramatically increases, mostly driven by demands of mobile computing. When the cost of such technology drops, WSNs will be able to afford hardware which can implement more sophisticated data aggregation and trust assessment algorithms. The aggregation of data from multiple sensor nodes is done at the aggregating node, by simple method such as averaging. However such aggregation is known to be highly vulnerable to node compromising attacks. Generally, WSNs are highly susceptible to such attacks due to absences of tamper resistant hardware. Iterative Filtering technique simultaneously aggregate data from multiple sources, usually in a form of corresponding weight factors. In our proposed method, we demonstrate that several existing iterative filtering algorithms, while significantly more robust against collusion attacks than the simple averaging methods, are nevertheless susceptive to a novel sophisticated collusion attack is resolved by using hop by hop authentication mechanism in data aggregation of with help of robust iterative algorithm for ensures the data confidentiality. The experimental results proves that obtained results is better than existing system.**

Index Terms*— Wireless Sensor Network, Aggregation,, Security, Collusion attack and hop by hop authentication scheme.

I. Introduction

The wireless sensor network is defined as the highly distributed networks of small, lightweight wireless node, deployed in large numbers to trust the environment or system by the measurement of physical parameters such as temperature, pressure or relative humidity. In the WSN, the data from the sensor nodes are collected by means of data aggregation. Sensory information is collected by the nodes. WSN consists of a base station and the number of nodes. The aggregator node is used to aggregate the data from multiple sensor nodes and then the data is forwarded to the base station.

There is several security challenges can be faced during the aggregation of data. Due to this wireless aggregation, eavesdropping and packet injection are occurred. Providing security in the sensor network is more difficult than the mobile adhoc network. To achieve the security in WSN, they perform various cryptographic operations like encryption, decryption and authentication and so on. For any cryptographic operation they must use any of the key like symmetric key or asymmetric key. If symmetric key is used then it is very difficult to design for security purpose. If asymmetric key is used then it is too expensive. For applying any of the encryption scheme then it has extra bits, memory required, delay occurred and so on. In the existing system, various algorithms are used to achieve the security during data aggregation. Many algorithms focus only on the specific attacks or problems. The iterative filtering algorithm is only concentrate on collusion attack.

The technique must be robust in the presence of non-stochastic errors, such as faults and malicious attacks and besides aggregating data; also provide an assessment of the reliability and trustworthiness of the data received from the sensor nodes. Identification of a new sophisticated collusion attacks against IF based reputation system which reveals a severe vulnerability of techniques. The novel method for estimation of sensor errors which is effective in a wide range of sensor faults and not susceptible to the described attack. Design of an efficient and robust aggregation method inspired by the MLE, which utilizes an estimate of the noise parameters obtained. Enhanced IF schemes able to protect against sophisticated collusion attacks by providing an initial estimate of trustworthiness of sensors using inputs [1].

The performance of IF is validated by simulation on synthetically generated data sets. The simulation results illustrate that the robust aggregation technique is effective in terms of robustness against the novel sophisticated attack scenario as well as efficient in terms of the computational cost. The sensor errors are estimated based on biased and unbiased readings in specified location. IF provides both higher accuracy and better collusion resistance than the other methods.

A. SECURITY IN WIRELESS SENSOR NETWORK

During the transmission of data the wireless sensor network must need the security. This security is also needed for every data as well as the nodes for which transferring the data. The security is needed while transmitting the data for wireless communication. The following information discuss that why security is needed.

Providing security in sensor networks is more difficult because of limited number of resources.

Security is needed at the design time to ensure that operation safety, secrecy of sensitive data and privacy for people in the sensor environment.

II. LITERATURE SURVEY

In [2] Chan H., Perrig A., and Song D. discussed secure hierarchical in-network aggregation in sensor networks. The first algorithm for provably secure hierarchical in network data aggregation. The algorithm is guaranteed to detect any manipulation of the aggregate by the adversary beyond what is achievable through direct injection of data values at compromised nodes. In other words, the adversary can never gain any advantage from misrepresenting intermediate aggregation computations. The algorithm incurs only O(Dlog2 n) node congestion, supports arbitrary tree-based aggregator topologies and retains its resistance against aggregation manipulation in the presence of arbitrary numbers of malicious nodes. The main algorithm is based on performing the SUM aggregation securely by first forcing the adversary to commit to its choice of intermediate aggregation results, and then having the sensor nodes independently verify that their contributions to the aggregate are correctly incorporated. They show how to reduce secure MEDIAN, COUNT, and AVERAGE to this primitive.

In [3] Ho J.-W., Wright M., and Das S. introduce fast zone-based node compromise detection and revocation in wireless sensor networks using sequential hypothesis testing. However, they require each sensor node to be attested periodically, thus incurring substantial overhead. To mitigate the limitations of the existing schemes, they propose a zone-based node compromise detection and revocation scheme in wireless sensor networks. The main idea behind this scheme is to use sequential hypothesis testing to detect suspect regions in which compromised nodes are likely placed. In these suspect regions, the network operator performs software attestation against sensor nodes, leading to the detection and revocation of the compromised nodes. Additionally, the detection problem using a game theoretic analysis, derive the optimal strategies for the attacker and the defender, and show that the attacker's gain from node compromise is greatly limited by the defender when both the attacker and the defender follow their optimal strategies.

In [4] Roy S., Conti M., Setia S., and Jajodia S. discussed a secure data aggregation with a large sensor network, in-network data aggregation significantly reduces the amount of communication and energy consumption However, this aggregation framework does not address the problem of false sub aggregate values contributed by compromised nodes resulting in large errors in the aggregate computed at the base station, which is the root node in the aggregation hierarchy. This is an important problem since sensor networks are highly vulnerable to node compromises due to the unattended nature of sensor nodes and the lack of tamper-resistant hardware. Thorough theoretical analysis and extensive simulation study show that the algorithm outperforms other existing approaches. Irrespective of the network size, the per-node communication overhead in the algorithm is O(1).

In [5] Tang L.-A., Yu X., Kim S., Han J., Hung C.-C., and Peng W.-C. Introduce trustworthiness analysis of sensor networks Cyber-Physical System (CPS) which integrates physical devices with cyber components to form a situation-integrated analytical system that responds intelligently to dynamic changes of the real world scenarios. In the paper, they propose a method called True-Alarm which finds out trustworthy alarms and increases the feasibility of CPS. True-Alarm estimates the locations of objects causing alarms, constructs an object-alarm graph and carries out trustworthiness inferences based on linked information in the graph. Extensive experiments show that True-Alarm filters out noises and false information efficiently and guarantees not missing any meaningful alarms.

Y. Sun et al. [6], accomplish data trustworthiness by extending Josang's trust model. Based on the multilayer aggregation architecture of network, they design a trust-based framework for data aggregation with fault tolerance with a goal to reduce the impact of erroneous data and provide measurable trustworthiness for aggregated results.

H.-S. Lim et al. [7], addressed the important and challenging problem of assuring trustworthiness of sensor data in the presence of malicious adversaries. They developed a game theoretic defense strategy to protect sensor nodes from attacks and to guarantee a high level of trustworthiness for sensed data. The objective of the defense strategy is to ensure that sufficient sensor nodes are protected in each attack/defense round.

III. EXISTING SYSTEM

Due to limited computational power and energy resource, aggregation of data from multiple sensor nodes is done at the aggregating node is usually accomplished by simple methods such as averaging. Aggregation reduces the amount of network traffic which helps to reduce energy consumption on sensor nodes. It however complicates the already existing security challenges for wireless sensor networks and requires new security techniques tailored specifically for various purposes. Providing security to aggregate data in WSN is highly vulnerable to node compromising attacks. Since WSN are usually unattended and without tamper resistant hardware, they are highly susceptible to such attacks. Thus, ascertaining trustworthiness of data and reputation of sensor nodes is crucial for WSN.As the performance of very low power processors dramatically improves; future aggregator nodes will be capable of performing more sophisticated data aggregation algorithms, thus making WSN less vulnerable

IV. LIMITATIONS

The major limitation in the existing methods, to address this security issue and proposes an improvement for iterative filtering techniques by providing an initial approximation for such algorithms which makes them not only collusion robust, but also more accurate and faster converging. This algorithm does not handle packet drop attack and not efficient for centralized approach.

V. PROPOSED SYSTEM

A. *ADVERSARY MODEL*

The past researchers [8] develops the attack models by considering the fact that they cannot rely on cryptographic methods for preventing the attacks, since the adversary may extract cryptographic keys from the compromised nodes. The authors in, considers Byzantine attack model, where the adversary can compromise a set of sensor nodes and insert any false data through the compromised nodes [9].

Following are some assumptions made in this model

a) Sensors are deployed in a hostile unattended environment with some physically compromised nodes.

b) When a sensor node is compromised, all the information which is inside the node becomes accessible by the adversary. System cannot depend on cryptographic methods for preventing the attacks because the adversary may extract cryptographic keys from the compromised nodes [10].

c) Through the compromised sensor nodes the adversary can send false data to the aggregator with a purpose of changing the aggregate values.

d) All compromised nodes can be under control of a single adversary or a colluding group of adversaries, enabling them to launch a sophisticated attack.

e) The adversary has enough knowledge about the aggregation algorithm and its parameters.

f) The base station and aggregator nodes cannot be compromised by adversary node.

B. COLLUSION ATTACK SCENARIO

In this scenario ten sensors are assumed that report the values of temperature which are aggregated using suitable aggregation algorithm. Most of the algorithms employ simple assumptions about the initial values of weights for sensors [5]. In suitable adversary model, an attacker is able to mislead the aggregation system through careful selection of reported data values. The collusion attack scenarios are as follows:

1) In scenario 1, all sensors are trustworthy and the result of the aggregation algorithm is close to the actual value.

2) In scenario 2, first an adversary compromises two sensor nodes, and alters the readings of these values such that the simple average of all sensor readings is twisted towards a lower value. As these two sensor nodes report a lower value, aggregation algorithm penalizes them and assigns to them lower weights, because their values are far from the values of other sensors.

3) In scenario 3, an adversary compromise three sensor nodes in order to launch a collusion attack. It listens to the reports of sensors in the network and instructs the two compromised sensor nodes to report values far from the true value of the measured quantity. It then computes the twisted value of the simple average of all sensor readings and commands the third compromised sensor to report such skewed average as its readings. In other words, two compromised nodes twist the simple average of readings, while the third compromised node reports a value very close to such twisted average.

C. PROPOSED APPROACH

The main goal of data aggregation algorithm is to gather and aggregate data in an energy efficient manner so that network life time is enhanced. Wireless Sensor Network offers an increasingly, attractive method of data gathering in distributed system architectures and dynamic access via wireless connectivity. Iterative Filtering technique provides a solution for a major problem regarding with data aggregation in WSN.IF, simultaneously aggregate data from multiple sources and provide trust assessment of these sources, usually in a form of corresponding weight factors assigned to data provided by each source. By demonstration it is proved that iterative filtering techniques are more robust against collusion attacks than the simple averaging methods, to a novel sophisticated collusion attack. To address this security issue, an improvement for iterative filtering techniques is done by providing an initial approximation for such technique which makes them not only collusion robust, but also more accurate and faster converging. We process the hop by hop authentication, it is scalable authentication scheme based on elliptic curve cryptography (ECC). While enabling intermediate nodes authentication, our proposed scheme allows any node to transmit an unlimited number of messages without suffering the threshold problem. In addition, our scheme can also provide message source privacy

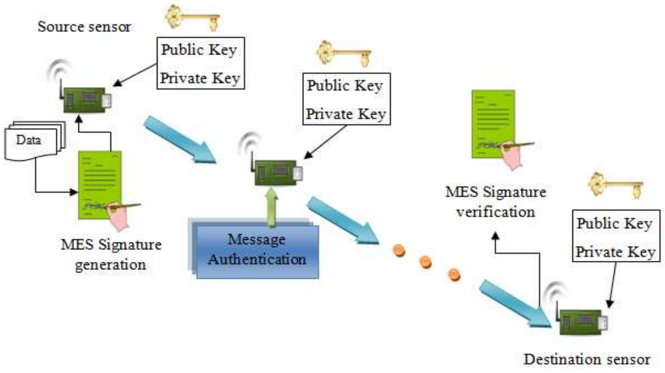


Fig.1The architecture diagram

After registration in the network if the user is valid they can enter into the existing network topology. The user must register their login credentials and to select the assigning weight factors depending on the number of data have to be used. By using IF, the sensor error is estimated in a wide range of sensor faults and not susceptible to the described attack. It utilizes an estimate of the noise parameters obtained from sensor nodes. The enhanced IF schemes able to protect against sophisticated collusion attacks by providing an initial estimate of trustworthiness of sensor using input. Our contribution to address vulnerability of IF algorithm is to employ the results of the proposed hop by hop authentication method, we also propose hop-by-hop message authentication scheme for protect the data. We are using ElGmal signature for message authenticate. Along this signature we can provide the secure for data packet and also using the signature we can detect the adversaries. The message receiver should be able to verify whether the message sent by the authorized node and also verify the message has been modified by the adversaries. Every forwarder can verify the message is authenticated or not. If the forwarder detect the intruder or find the message has been modified, forwarder will drop the packet or change the routing path. Along this proposed scheme, we can get accurate data without modifying and also can easily detect the adversaries.

The four modules for secure data aggregation using IF are:

A. Node Configuration.
B. Data aggregation in multiple sources.
C. Hop by Hop message authentication scheme
D. Signature generation algorithm
E. Signature verification algorithm

1) *NODE CONFIGURATION*

In this network, the nodes are static and fixed. The sensor nodes are divided into disjoint clusters, and each cluster has a cluster head which acts as an aggregator. Data are periodically collected and aggregated by the aggregator. In this project we assume that the aggregator itself is not compromised and concentrate on algorithms which make aggregation secure when the individual sensor nodes might be compromised and might be sending false data to the aggregator. We assume that each data aggregator has enough computational power to run an IF algorithm for data aggregation with message authentication scheme.

2) *DATA AGGREGATION IN MULTIPLE SOURCES*

This module specifies the data aggregation from multiple sources. Data aggregation is any process in which information is gathered and expressed in a summary form, for purposes such as statistical analysis. A common aggregation purpose is to get more information about particular groups. The network is formed and the aggregate node collects many data from multiple nodes. It is also reduce the data traffic.

3) *HOP BY HOP MESSAGE AUTHENTICATION SCHEME*

We also propose hop-by-hop message authentication scheme for protect the data. We are using ElGmal signature for message authenticate. Along this signature we can provide the secure for data packet and also using the signature we can detect the adversaries. The message receiver should be able to verify whether the message sent by the authorized node and also verify the message has been modified by the adversaries

Let p > 3 be an odd prime. An elliptic curve E is defined by an equation of the form:

$E : y^2 = x^3 + ac + b \ mod \ p,$

Where a, b $\in F_p$, and $4a^3 + 27b^2 \not\equiv 0 \ mod \ p$. The set E($F_p$) consists of all points (x,y) $\in F_p$ on the curve, together with a special point $\mathbf{O}$, called the point at infinity.

Let G = (x_G, y_G) be a base point on E(F_p) whose order is a very large value N. user A selects a random integer $d_A \in$[1, N-1] as his private key. Then, he can compute his public key Q_A from $Q_A = d_A \times G$.

4) *SIGNATURE GENERATION ALGORITHM*

We utilize the signature generation algorithm, in order to make secure key generation using the algorithm given below:

For Alice to sign a message m, she follows these steps:

1. Select a random integer $k_A.1 \leq k_A \leq N - 1.$
2. Calculate r = $x_A mod$ N, Where (x_A, y_A) = $k_A \ G$. If r = 0, go back to step 1.
3. Calculate $h_A \overset{l}{\leftarrow} h(m, r)$, where h is a cryptographic hash function, such as SHA-1, and $\overset{l}{\leftarrow}$ denotes the l leftmost bits of the hash.
4. Calculate s = r$d_A h_A + k_A$ mod N. If s =0, go back to step 2.
5. The signature is the pair (r,s).

5) *SIGNATURE VERIFICATION ALGORITHM*

We utilize the signature verification algorithm, in order to make secure key verification using the algorithm given below:

For Bob to authenticate Alice's signature, he must have a copy of her public key Q_A then he:

1. Checks that $Q_A \neq \mathbf{O}$, otherwise invalid
2. Checks that Q_A lies on the curve
3. Checks that $nQ_A = \mathbf{O}$

After that, Bob follows these steps to verify the signature:

1. Verify that r and s are integers in [1, N – 1]. If not, the signature is invalid.
2. Calculate $h_A \overset{l}{\leftarrow} h(m, r)$, where h is the same fucntion used in the signature generation.
3. Calculate (x_1, x_2) = sG - r$h_A Q_A$mod N.
4. The signature is valid if r = x_1 mod N, invalid otherwise.

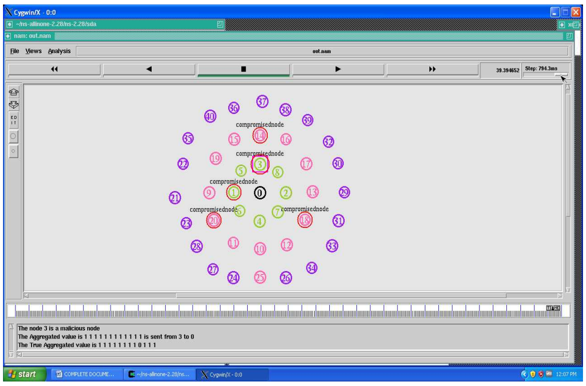
VI. PERFORMANCE EVALUATION

A. *Simulation Parameters*

The NS2 tool [16] is used to study the performance of our proposed method. We employ the IEEE 802.11 [17] MAC with a channel data rate of 11 Mb/s . We choose the two evaluation metrics: number of deviation of nodes, number of nodes performed in the MAC layer and the average number of bits send in the node for a packet to be transmitted from the source to destination,.

Table I. Stimulation parameters

Parameter	Value
Application Traffic	10 CBR
Transmission rate	4 packets/s
Packet Size	512 bytes
Channel data rate	11 Mbps
Area	700m*700m
Simulation time	800



```
The node 3 is a malicious node
The Aggregated value is 1 1 1 1 1 1 1 1 1 1 1 1 is sent from 3 to 0
The True Aggregated value is 1 1 1 1 1 1 1 1 0 1 1 1
```

Fig.2 Network Topology

B. *Simulation Results*

We used the performance metrics to validate the proposed algorithm with results obtained in this papers are shown in Figure 3,4 and 5.

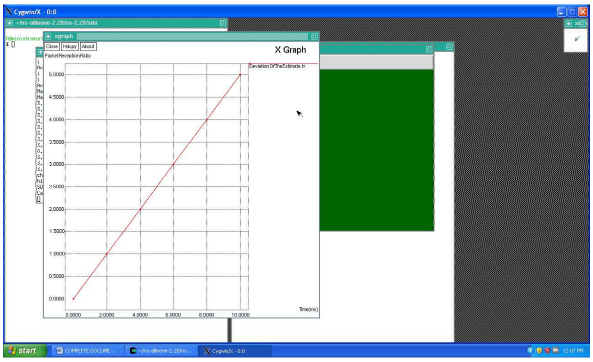


Fig.3 Deviation in the nodes

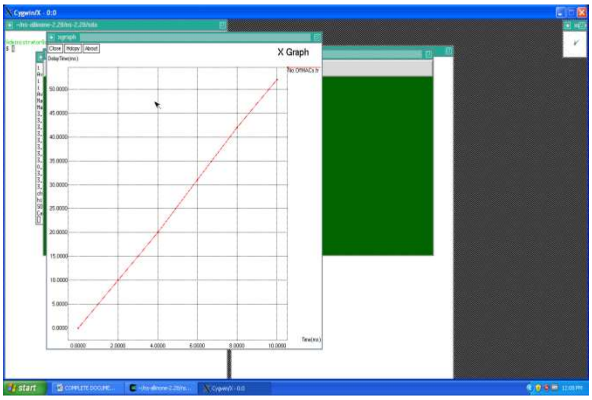


Fig. 4 Nodes in the MAC layer

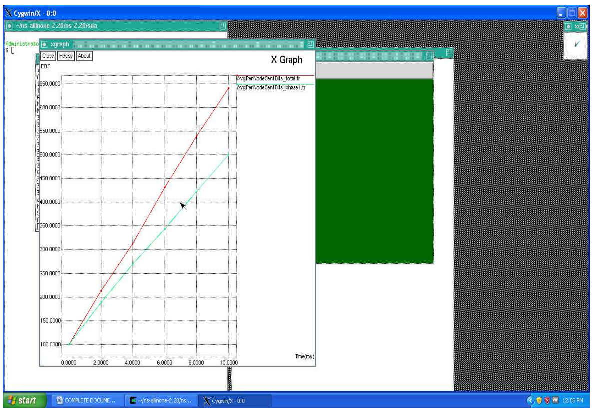


Fig. 4 Number of nodes sends bites

Thus the proposed scheme is very significant and effective when comparing with existing methods.

VII. CONCLUSION

In wireless sensor network computational cost and energy need high level for transmitting the data. So that the data aggregation technique is used in WSN. This technique is done by using various simple methods such as averaging but this data aggregation is highly vulnerable. We propose the hop by hop authentication, it is scalable authentication scheme based on elliptic curve cryptography (ECC). While enabling intermediate nodes authentication, our proposed scheme allows any node to transmit an unlimited number of messages without suffering the threshold problem. In addition, our scheme can also provide message source privacy in the data aggregation.

REFERENCES

[1] S.Ozedemir and Y.Xiao,(Aug.2009)"Secure data aggregation in wireless sensor networks: A comprehensive overview, " Computer Network, vol.53,no.12,pp.2022-2037.

[2] L.Wasserman, All of statistics: a concise course in statistical inference. New York: Springer.

[3] K.Hoffman,D.Zage and C.Nita-Rotaru (Dec.2009), "A survey of attack and defence techniques for reputation systems," ACM Comput .Surv, vol.42 ,no.1, pp.1:1-1:31.

[4] H-S.Lim, Y-S. Moon and E. Bertino (2010) "Provenance-based trustworthiness assessment in sensor networks," in proceedings of seventh International Workshop on Data Management for sensor Networks, ser. DMSN, pp.2-7

[5] Y.Zhou,T.Lei and T.Zhou (2010) "A robust ranking algorithm to spamming", CoRR, vol. abs/1012.3793.

[6] P.Laureti, L.Moret, Y-C.Zhang and Y-K.Yu(2006), "Information filtering via Iterative Refinement, "EPL (Europhysics Letters),vol75,pp.1006- 1012.

[7] R-H.Li, J.X.Yu,X.Huang and H.Cheng(2012),"Robust reputation based ranking on bipartite ranking networks", in SDM'12,pp.612-623.

[8] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," IEEE/ACM Trans. Netw., vol. 14, no. 2, pp. 316–329, Apr. 2006.

[9] S.Ganeriwal, L.K. Balzano ,and M.B.Srivastava, "Reputation based framework for high integrity sensor networks," ACM Trans. Sen. Netw., vol. 4, no. 3, pp. 15:1–15:37, Jun. 2008.

[10] M. Li, D. Ganesan, and P. Shenoy, "PRESTO: feedback-driven data management in sensor = networks," in Proceedings of the 3rd conference on Networked Systems Design & Implementation - Volume 3, ser. NSDI'06, 2006, pp. 23–23.