False Data Detection Using MAC pairs in Wireless Sensor Networks

V.M.Sivagami[1], K.S.Easwara Kumar[2]

[1] Research Scholar, [2]Professor

Department of Computer Science & Engineering

Anna University, Chennai

Abstract: Wireless sensor networks are vulnerable to many types of security attacks, including false data injection, data forgery, and eavesdropping. Sensor nodes can be compromised by intruders, and the compromised nodes can distort data integrity by injecting false data. False data can be injected by compromised sensor nodes in various ways, including data aggregation and relaying. Data confidentiality prefers data to be encrypted at the source node and decrypted at the destination. However, data aggregation techniques usually require any encrypted sensor data to be decrypted at data aggregators for aggregation. The basic idea behind the false data detection algorithm is to form pairs of sensor nodes such that one pair computes a message authentication code (MAC) of forwarded data and the other pair mate later verifies the data using the MAC. Data aggregation is implemented in wireless sensor networks to eliminate data redundancy, reduce data transmission, and improve data accuracy.

***Keywords*- Data aggregation and authentication protocol (DAA), Data integrity, network-level security, Message Authentication Code(MAC) and Sensor networks.**

1. INTRODUCTION

Wireless sensor networks are vulnerable to many types of security attacks, including false data injection, data forgery, and eavesdropping [1]. The transmission of false data depletes the constrained battery power and degrades the band- width utilization. False data can be injected by compromised sensor nodes in various ways, including data aggregation and relaying. The existing false data detection techniques consider false data injections during data forwarding only and do not allow any change on the data by data aggregation. To support data aggregation along with false data detection, the monitoring nodes of every data aggregator also conduct data aggregation and compute the corresponding small-size message authentication codes for data verification at their pair mates. To support confidential data transmission, the sensor nodes between two consecutive data aggregators verify the data integrity on the encrypted data rather than the plain data.

Sensor nodes are often deployed in a hostile environment and it may be captured or compromised by the adversaries and secret information such as symmetric key may be revealed to the adversaries [2]. Thus adversaries can easily inject false data reports of non-existing events or faked readings. Such an attack is called false data injection attack. It may not only cause false alarms due to bogus sensing reports but also drain out the limited energy of the nodes forwarding these reports, thus reduce the lifetime of the sensor networks. Three schemes have been proposed to detect and contain such attacks.SEF is a pre deployment scheme in which each node randomly picks some secret keys from one partition of a global key pool before deployment. All nodes have pre determined probability to detect and filter false report. It has limited filtering capacity. Hence a post deployment scheme that requires each node periodically establish pair wise keys with others that are multi hop away from it. This scheme can drop false report within the fixed number of hops. This requirement is impractical for the sensor network with high dynamic topological changes, which are due to nodes failures or nodes switching their state between active and sleeping mode to save energy. Commutative Cipher based En-route Filtering (CCEF) scheme in which each node preloads a distinct authentication key before deployment. When the reports are needed the base station distributes a session key to the cluster head and a witness key to the forwarding nodes respectively. It suffers from dynamic topology problem by requiring the same fixed path from messages in both directions between the base station and the cluster head.

Data aggregation [3] usually involves the fusion of the data from multiple sensors at intermediate nodes and transmission of aggregated data to the base station. Data aggregation attempts to collect the most critical data from the sensors and make it available to the sink in an energy efficient manner with maximum data latency. The following issues have been addressed for secure data aggregation. Some sensor nodes may be compromised and transmit wrong data values to the aggregator that corrupts the aggregation result. The aggregator may be compromised and report maliciously aggregate values to the home serve or the sink. Estimation errors introduced by the sampling techniques used by the aggregator to compute the result.

Data fusion is used to process the collected information before they are sent to the base station or the observer of the sensor network. The security of data fusion process is studied .Data fusion is employed in order to reduce the traffic load from the entire sensor to the base station. To reduce energy consumption in the scheme, minimum length needed for the message authentication code to achieve a ore defined level of security. The results show that the number of both used for MAC's[4] does not increase linearly with the number of witnesses.

Recent advances in wireless communication and electronics have enabled the development of low-cost, low power, multi functional sensor nodes that are small in size and communicate undeterred in short distances. The design of sensor networks [5] is influenced by many factors including fault tolerance, scalability, production cost, operating environment, sensor network topology, hardware constraints, transmission media and power consumptions. The main task of a sensor node in a sensor field is to detect events, perform quick local data processing and then transmit data. Power consumption can hence be divides in to three domains-sensing, communication and data processing.

Data aggregation is essential to reduce data redundancy and to improve data accuracy. False data detection is essential for the protection of data integrity and efficient utilization of battery power and bandwidth. It enhances the network lifetime because communication constitutes 70% of the total energy consumption of the network. Data confidentiality is required in sensor network applications.

2. EXISTING SYSTEM

In the existing system, false data detection techniques consider the false data injections only during data forwarding. An SEF scheme enables relaying nodes and base station to detect false data with a certain probability. In 10 hops, SEF is able to drop 80%–90% of the injected false reports. In the interleaved hop-by-hop authentication scheme, any packet containing false data injected by compromised sensor nodes is detected by those T+1 sensor nodes that collaborate to verify data integrity. In the interleaved hop-by-hop authentication scheme, sensor nodes are not allowed to perform data aggregation during data forwarding. The Commutative Cipher based En-route Filtering (CCEF) scheme drops false data en-route without symmetric key sharing. In CCEF, the source node establishes a secret association with base station on a per-session basis, while the intermediate forwarding nodes are equipped with a witness key. With the use of a commutative cipher, a forwarding node can use the witness key to verify the authenticity of the reports without knowing the original session key.

The first and foremost limitation of existing system is the value of 'T' depends strictly on several factors such as geographical area conditions, modes of deployment, transmission range of sensor nodes, power management, and node density of the network. Second, the pair wise key establishment among non neighboring nodes takes more time than that among direct neighboring nodes. Therefore such key establishment process is more vulnerable to node compromise attacks. Finally, group communication schemes are vulnerable to those attacks where an adversary who compromises a legitimate group member seizes some or all past group keys as well as the current group key and discloses the secrecy of the data.

3. PROPOSED SYSTEM

DAA protocol aims to integrate false data detection with the data aggregation and data forwarding. For every data aggregator the corresponding small-size MAC codes are computed and Data verification at the pair mates are made. For confidentiality, the data aggregators verify the data integrity on the encrypted data.DAA provides secure data aggregation, data confidentiality, and false data detection by performing data aggregation at data aggregators and their neighboring nodes and verifying the aggregated data during data forwarding between two consecutive data aggregators.

3.1 FALSE DATA DETECTION

This paper aims at providing network security and efficiency. The mainstay of our work is to integrate the False Data Detection with Data Aggregation and Confidentiality Using data

aggregation and authentication protocol (DAA).Compromised sensor nodes can distort the integrity of data by injecting false data.

Previously known techniques on false data detection do not support data confidentiality and aggregation, even though they are usually essential to wireless sensor networks. However, our work has presented the novel security protocol DAA to integrate data aggregation, confidentiality, and false data detection.

DAA appends two FMACs to each data packet. To reduce the communication overhead of algorithm SDFC, the size of each FMAC is kept fixed. Each FMAC consists of T+1 subMACs to safeguard the data against up to T compromised sensor nodes. We start with the architecture of the system

3.2 ARCHITECTURE DIAGRAM

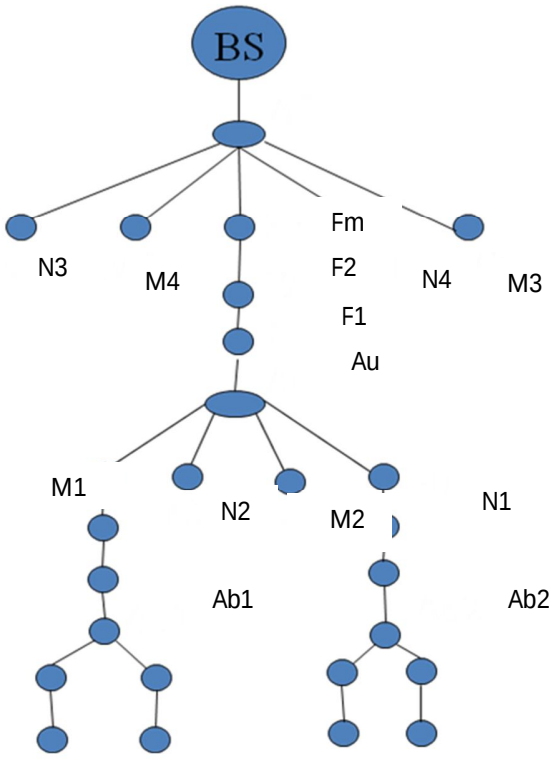

Fig.1 Architecture of Sensor Nodes

Au -Current data aggregator

Af -Forward data aggregator of Au

F1,F2..Fm-Forwarding nodes of Au

M3&M4 -Monitoring nodes of Af

M1&M2 -Monitoring nodes of Au

Ab1&Ab2-Backward aggregators of Au

The Backward aggregators aggregate the data in the appropriate nodes .Monitoring nodes are selected from the neighboring nodes that are present in the network using the MNS algorithm. Thus backward aggregators send the intended data to the Au via these selected monitoring nodes. Data is then forwarded through F1,F2..Fm. These are then sent to the monitoring nodes of Af and finally reaches the destination.

3.3 NETWORK TOPOLOGY

The data aggregators are chosen in such a way that:

1) There are at least T nodes, called forwarding nodes, on the path between any two consecutive data aggregators; and

2) Each data aggregator has at least T neighboring nodes, so they can form T pairs with the forwarding nodes on the path between two consecutive data aggregators.

In order to ensure that there are at least nodes between any two consecutive data aggregators, we assume that the secure data aggregator selection protocol (SANE) is employed as follows. Sensor nodes are scattered over a large area to form small sets of nodes in close proximity from each other. These sets are called sectors, and the sector size (i.e., the number of nodes in a sector) depends on the value of . The protocol SANE is first run to select candidate data aggregators. Since sector size is determined based on the value of T, the number of intermediate nodes between any two consecutive candidate aggregators is expected to be around If it happens that there are less than intermediate nodes between two consecutive candidate aggregators, one of these candidate aggregators drops its candidacy, and then the protocol SANE is run again. This process is repeated until there are at least intermediate nodes between any two consecutive data aggregators.

We assume a fully decentralized network of equal functionality and capability microsensor nodes. All nodes in the network are stationary. Each node can have the role of a sensing node, an aggregator node or a forwarding node. The sensor nodes are scattered over a large area so that they form small sets of nodes in close proximity from each other. We call these sets sectors. Nodes in the same sector are pre-configured with the same sector ID and are said to belong to the same set S. For example, sensor

nodes released from the same parachute would belong to the same sector. We further assume a MAC scheme for broadcast and unicast communication. Election protocol messages are exchanged only among nodes in the same sector. These messages are either sent via unicast among nodes in the same sector or disseminated to all nodes in a sector using a simple sector-aware controlled flooding scheme. With this flooding scheme, nodes re-broadcast first seen messages only if the source of the message belongs to the same sector as they do. We denote the i-th sensor node in a sector S by si, and the aggregator node during the t-th epoch as At. For each sensor si, we denote by Ni C S, the set of nodes from which it has received valid election contributions during an election round. A contribution refers to a node's input to the election process. An election contribution of a node can be sent in one or two messages, or it can be aggregated with other node contributions as they are propagated in the sector.

With respect to security, a SANE protocol should attain:

* **Non-manipulability** - A party's contribution in the election process should not be able to influence the decision of honest nodes towards the election of preselected nodes. We further call a protocol strongly non-manipulable, if in addition to the above property, no party has the ability to prevent the election of a preselected node. Unlike non-manipulability, strong non-manipulability is desirable but not required by a SANE protocol.

*** Authentication** - Each party should consider only the contributions of a restricted set of nodes. Contributing nodes should be capable of proving that they belong to this set. We note here that our protocol descriptions do not explicitly address authentication, however this is a property that can be efficiently achieved by incorporating existing WSN authentication solutions.

* **Unpredictability** - An election protocol is predictable if the adversary can forehand know the order in which nodes are elected as aggregators. This information could facilitate adversarial actions. For example, an adversary with the ability to compromise only a few nodes at a time, could use this knowledge to prevent a cluster of sensor nodes from transmitting information to its consumer during selected periods.

3.4 MONITORING NODE SELECTION

Each data aggregator is monitored by its T neighboring nodes out of total n neighboring nodes. T Neighbors of a data aggregator A are selected as monitoring nodes to perform the data aggregation and to compute subMACs of the aggregated data. The monitoring nodes are selected by the Monitoring Node

Selection (MNS) algorithm. The selection of T monitoring nodes for each data aggregator in Algorithm MNS is to assign indices to the neighboring nodes in some order and then compute T indices by applying modulus operation to the sum of some random numbers generated by the neighboring nodes. Any neighboring node index is equal to one of these T indices becomes a monitoring node. The data aggregator and all neighboring nodes are involved with the selection of monitoring nodes to minimize the adverse impact of a compromised node.

MNS protects a compromised data aggregator from affecting the monitoring node selection. The monitoring nodes are selected by all neighboring nodes. To affect the selected monitoring nodes, a compromised data aggregator must change the random numbers before broadcasting them.

3.5 PAIRS FORMING

The Following 2T+1 pairs of nodes are formed enabling the nodes of every pair to share a distinct Symmetric Key.

One pair is formed by the current data aggregator and the forward data aggregator (AA-type).

T pairs are formed by the monitoring nodes of the Current Data Aggregator and the neighboring nodes of the Forward Data Aggregator (MN-type).

T pairs are formed by the monitoring and forwarding nodes of the Current Data Aggregator (MF-type).

To establish pairs among monitoring nodes and forwarding nodes, Af sends out a “pair mate discovery message” M to Au along with its neighboring node list. Af also adds the MAC of neighboring node list using the key it shares with Au. Message M is forwarded by the nodes on the path between Af and Au, and each node that forwards M appends its ID to M. When Au receives M, it has the IDs of its forwarding nodes and neighboring nodes of Af. Then, computes the MAC of the concatenated IDs using and broadcasts the MAC forwarding, then DAA would form only T+1 pairs (i.e., one pair of AA-type and T pairs of MN-type) to detect false data at data aggregators and their neighboring nodes. However, the second step of DAA forms 2T+1 pairs to benefit from the fact that false data detection during data forwarding allows false data to be dropped as early as possible.

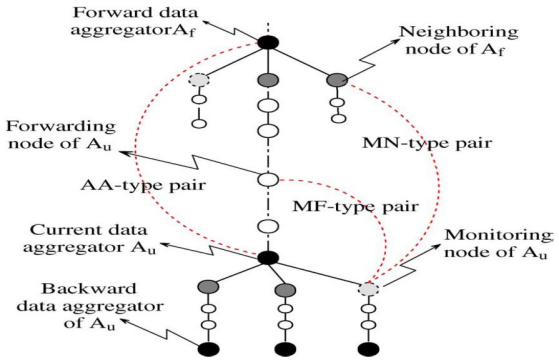

Fig.2. Node pairs between two consecutive data aggregators Au and Af. Three types of node pairs are formed: 1) an AA-type pair by data aggregators Au and Af ; 2) an MF type pair by the monitoring node of Au and a forwarding node of Au .

3.6 DATA AGGREGATION AND FALSE DATA DETECTION

This section introduces Algorithm SDFC to provide false data detection, secure data aggregation and data confidentiality for the third step of DAA. To provide data confidentiality, transmitted data are always encrypted and forwarding nodes perform the data verification over the encrypted data. Prior to this third step of DAA, monitoring nodes of every data aggregator are selected, and 2T+1 pairs are formed. To verify data integrity and detect false data injections, one pairmate computes a subMAC, and the other pairmate verifies the subMAC. subMACs are computed for both plain and encrypted data. subMACs of plain data are used to detect false data injections during data aggregation, whereas subMACs of encrypted data are used to detect false data injections during data forwarding. To detect any false data that the current data aggregator Au can inject during data aggregation, the monitoring nodes of Au also aggregate the incoming data of Au and compute subMACs for the plain aggregated data, so that the forward data aggregator Af and its neighboring nodes verify the subMACs. Similarly, to detect those false data that can be injected during data forwarding, the monitoring nodes of Au compute subMACs for the encrypted aggregated data and then their pairmates of forwarding nodes verify the subMACs. The main steps of SDFC are:

1) Whenever some data are received by a data aggregator, the authenticity of data is verified by the data aggregator and its neighboring nodes;

2) The data aggregator and its monitoring nodes aggregate the data independently of each other;

3) Each monitoring node computes one subMAC for the encrypted data and the other subMAC for the plain data;

4) The data aggregator collects these subMACs from its monitoring nodes to form the FMACs of the encrypted and plain data, appends the FMACs to the encrypted data, and transmits them;

5) The forwarding nodes verify the data integrity of the encrypted data; and finally

6) The neighboring nodes of the next aggregator verify the integrity of the plain data. Each data aggregator forms two FMACs: one FMAC for the encrypted data, and the other FMAC for the plain data. Each FMAC consists of T+1 subMACs computed by the data aggregator Au and its T monitoring nodes. In the formation of FMACs, data aggregator Au determines the order of subMACs in anyway and inform each forwarding node about its subMAC location individually.

Consequently, an intruder cannot know in advance the exact location of subMAC bits for a given forward node. Therefore, to inject a false message, an attacker has to try all possibilities for a 32-bit FMAC. Thus, if an intruder wants to inject messages at a forwarding node to consume its energy, only (1/2)^32 of randomly generated messages at a forwarding node can be accepted and forwarded.

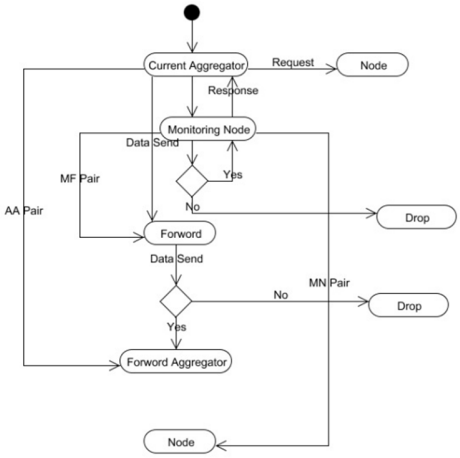

Fig.3 Flowchart for Algorithm SDFC.

4.METHODOLOGY

Nodes are created to form the network. Each node has node information, aggregator information and data transfer to describe its details.

4.1. NODE INFORMATION

Node information displays the name of the node, the input that the user entered for each node. And it also displays the randomly generated memory, Battery life and mobility which determine the strength of every node that is created in the network.

4.2. AGGREGATOR INFORMATION

Based on the memory, battery and mobility that was generated earlier current aggregator is selected and displayed. The node that has the highest sum of these three gets elected. This is the node that connects the backward and forward aggregators for sending the intended information in the designed network.

4.3. DATA TRANSFER DETAILS

Once nodes are created in the network, based on the range that is given to the nodes neighbors that are in close proximity to each other are detected and displayed. This also has the destination column to indicate the path via which information has to flow. The information is browsed and sent. To display this to the user, send and receive columns are used. The inject procedure is used for sending the malicious data in the network.

4.4. SELECTION OF MONITORING NODE

Among all nodes, the aggregator node is selected and monitoring node selection process starts . The MNSelect procedure asks for the generation of random numbers for all the neighboring nodes of the Au. Each neighbor sends two random numbers and they are displayed in the node details table. The neighbors are numbered in ascending order and index calculation is made to detect the monitoring node .

4.5 PAIRWISE KEY ESTABLISHMENT

Any node other than the Au is selected and the other Mobile node's Au is selected as the destination node. Pairwise procedure is applied on the mobile node to form three types of pairs namely AA, MN and MF. The keys are sent back and forth and appropriate messages are displayed. The generated keys are stored in the database for reference.

4.6. SELECTION OF SOURCE TO SEND THE INFORMATION

Now the source node is selected and the forwarding node's name will be given in the destination node. The respective file is browsed and the information is sent. The files that have an extension of txt, java, html and jsp are transmitted successfully.

4.7. ENSURING SECURED TRANSMISSION

A random number is generated and a group key is established to overcome any node compromise attacks. The encrypted data is sent to the destination. The key numbers are referred and entered from the key holder file for confidential transmission.

4.8. SUCCESSFUL TRANSMISSION

If there is no false data detected, then the information is transferred successfully and a message packet is sent to the source node. SubMACs are computed for the plain and encrypted data .The forwarding nodes verify data integrity of the encrypted data and neighboring nodes verify the integrity of plain data. When the integrity verification fails, false data are dropped.

5. EXPERIMENTAL ANALYSIS

In the existing system, it has been mentioned that DAA is simulated using QualNet network simulator for an area of 100*100 m and 100 sensor nodes with a transmission range of 15 m. Some nodes are designated as data aggregators and distributed into the network area uniformly. Data are assumed to be generated mainly by the nodes located at the edges of the network, although any node is allowed to sense events and generate data.

But we implemented it using the Bluetooth concept which directly detects the neighboring nodes that have been created in individual systems so that it can be formed as a full network. By doing so the data transmission can be made from any node making it as a source and then to forward it to the desired destination, the concept can be extended easily.

6. CONCLUSION

Thus every sensor node in the network is capable of detecting false data during data aggregation and data forwarding. Our scheme has improved the network security and efficiency during the data transmission in the wireless sensor networks. We have overcome the limitations of the existing systems such as taking a very minimal time for performing the pair wise key establishment and also making the group key very securable by protecting it from node compromise attacks. As we are dealing with a wireless network we have a constraint that firewalls and antivirus has

to be disabled before forming the network. This is essential as without this, the sensor nodes cannot communicate via different systems.

7. FUTURE ENHANCEMENT

Existing work has provided bounds on lifetime for networks with specific network topologies and source behaviors. It would be interesting to extend this work to more general topologies such as cluster based sensor networks.

Another interesting domain of research is the application of source coding theory for data gathering networks. The sensor data are usually highly correlated and energy efficiency can be achieved by joint source coding and data compression. Although some research has been pursued in this direction, there is significant scope for future work.

8.REFERENCES

[1] Suat Ozdemir, Member, IEEE, and Hasan Çam, Senior Member, IEEE, " Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks",Vol.18,no.3,June 2010.

[2] Z. Yu and Y. Guan, "A dynamic en-route scheme for filtering false data in wireless sensor networks," in Proc. IEEE INFOCOM, Barcelona, Spain, Apr. 23–27, 2006, pp. 1–12.

[3] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: A survey," IEEE Commun. Surveys Tutorials, vol. 8, no. 4, 4th Quarter 2006.

[4] W. Du, J. Deng, Y. S. Han, and P. K. Varshney, "A witness-based approach for data fusion assurance in wireless sensor networks," in Proc. IEEE GLOBECOM, 2003, pp. 1435–1439.

[5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Commun. Mag., vol. 40, no. 8, pp. 102–114, Aug. 2002.

[6] M. Sivrianosh, D.Westhoff, F. Armknecht, and J. Girao, "Non-manipulable aggregator node election protocols for wireless sensor networks," in Proc. IEEE WiOpt, Cyprus, Apr. 2007, pp. 1–10.

[7] F. Ye, H. Luo, S. Lu, and L. Zhang, "Statistical en-route detection and filtering of injected false data in sensor networks," in Proc. IEEE INFOCOM, 2004, vol. 4, pp. 2446–2457.

[8] S. Zhu, S. Setia, S. Jajodia, and P. Ning, "Interleaved hop-by-hop authentication against false data injection attacks in sensor networks," ACM Trans. Sensor Netw., vol. 3, no. 3, Aug. 2007.

[9] H. Yang and S. Lu, "Commutative cipher based en-route filtering in wireless sensor networks," in Proc. IEEE VTC, 2004, vol. 2, pp.1223–1227.

[10] D. Seetharam and S. Rhee, "An efficient pseudo random number generator for low-power sensor networks," in Proc. 29th Annu. IEEE Int.Conf. Local Comput. Netw., 2004, pp. 560–562.