Indian Journal of Science and Technology, *Vol 8(14), 62529, July 2015*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Hierarchical Approach to Secure Data Aggregation in Wireless Sensor Networks

S. Saranya* **and N. Bharathi**

School of Computing, SASTRA University, Thanjavur-613401, Tamilnadu, India; saranyaselvan91@gmail.com

Abstract

The main objective of this paper is to aggregate the data in a secure way such that the tampering caused by the adversaries can be avoided and these secure authentication mechanisms ensure that the base station does not accept forged aggregation value. Previous data aggregation schemes make sure that base stations do not receive and agree with the false aggregation values. The authentication mechanisms are used to detect malicious nodes. But these mechanisms may not be efficient and they can be easily bypassed by hacking the keys used for encryption. A secure mechanism for aggregating the data has been put forward in which the two stages of verification are performed. In the first stage, Level based scheme is applied as the verification technique at each level. In this scheme, the keys are updated dynamically which enhances the security. Next stage of verification is the encryption of the packets by using its private key. The proposed secure data aggregation mechanism improves the performance and efficiency of the aggregation process. Analysis of experimental results indicates that the performance of the proposed scheme is twice that of the previous schemes.

Keywords: Hierarchical Approach, Level Based Scheme (LBS), Secure Data Aggregation, Verification, Wireless Sensor Network (WSN)

1. Introduction

Wireless sensor networks are employed in various security related applications. There are several sensor nodes in a sensor network such that the nodes can interact with each other and with the base station. These sensor nodes are distributed over various remote areas and critical regions in which the nodes can detect the data and process the detected data before transmitting the data to the base station. Nodes in a region may identify common events which will lead to increased level of data redundancy. Sensor nodes utilize battery power for sensing and processing the received data. This leads to the consumption of more energy which leads to the reduction of its lifetime.

A protocol has been put forward which can securely calculate the average as well as median of the given data. In order to facilitate the heterogeneous network to deal with the packet synchronization, spanning tree based on attribute[1] has been proposed. Clue hop[2] is another technique which has been put forward for the efficient clustering of sensor network. Various techniques have been put forward for improving the lifetime of the sensor node. They are data aggregation based on the awareness of link[3] and data collection based on the awareness of energy[4].

Aggregation[5] of related data plays an important role in specific applications.

The drawbacks are overcome by aggregating the data in a secure way such that this scheme can check whether any malicious node attack has occurred or not, along the path from leaf node to root. Aggregation of data using in-network approach has been implemented[6]. But this scheme cannot identify the malicious node that has caused the tampering. To overcome this limitation, Hongjuan Li et al.[7] have proposed a technique for securely aggregating the data along with the malicious

**Author for correspondence*

node detection such that malicious aggregators are identified effectively. An example aggregation tree is described in Figure 1. Another approach to overcome the drawback has been proposed which uses homomorphic encryption[8]. Protocol for secure aggregation has been proposed for several aggregation functions which are called secure hop-by-hop data aggregation protocol[9].

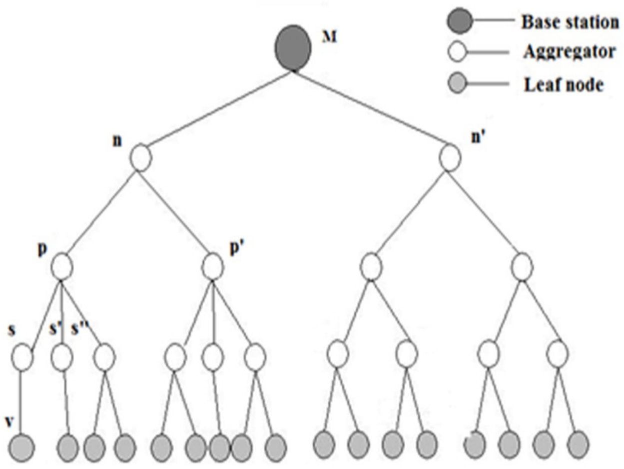

Figure 1. An Aggregation Tree.

Data aggregation can improve the lifetime of the network and reduce the consumption of energy. But these aggregation techniques are designed without considering the security. So in order to overcome this limitation, secure data aggregation[10–12] has been proposed. Synopsis diffusion is a framework for secure aggregation to calculate the aggregates accurately. Przydatek et al.[13] has proposed a framework which can aggregate the information in a secure way. Various techniques have been put forward for improving the lifetime of the sensor node.

Witness-based technique[14] has been put forward to ensure the validity of the data before sending it to the base station from the nodes that perform fusion of the received data. Data fusion values are checked by selecting specific nodes that perform the fusion of the received data, as witnesses. To ensure security, authentication mechanisms[15–18] have been designed. Persistent authentication scheme15 has been put forward to identify and avoid the tampering of the data by the nodes in the network. Digital watermarking approach[19] is another authentication mechanism that supports in-network operations. Another technique that has been used is a secure encrypted-data aggregation scheme for enhancing security and privacy.

2. Proposed Architectural Framework

Hierarchical approach to secure data aggregation has been proposed which can sense the malicious sensor nodes. Figure 2 shows architectural framework of proposed scheme. Initially, the leaf nodes collect the data value and send it to its corresponding parent node. Aggregation of data does not take place in this level. Before transmitting the packet to its parent node, the level key of the corresponding leaf node and parent node is checked. If the verification succeeds, then the packet is transmitted. If the decryption of the packet is performed successfully, then the packet is accepted by the parent node else the packet is discarded.

The intermediate nodes perform the level key verification in the same way before transmitting the packets. If the verification succeeds, the packet is accepted and the packet is decrypted. When the transmission has to be performed to the base station, the level key verification has to be performed and after the successful verification, decryption of the packet is performed. If it succeeds, packet is accepted by the base station. This procedure is repeated for each level of the hierarchical structure starting from level n-1 to level 0.

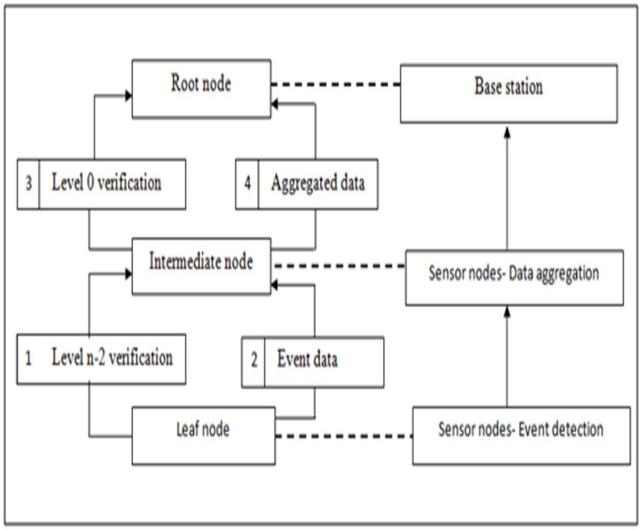

Figure 2. Architectural Framework.

The packet format that is being sent during the aggregation process is as follows:

< ID, data>

Where ID is the identifier of the node, data indicates the aggregated value which is calculated by considering the subtree's leaves. The level to level security is managed by level key. The packet is encrypted by the respective private key such that if any malicious node tries to forge the aggregation value, it must be able to decrypt the packet.

The packet of node s_i is represented by:

$<s_i d_i>$

$d_i={}^r s_i$, if s_i is a leaf node such that ${}^r s_i$ is the value that is obtained by the node $s_i.d_i = \sum_{p=1}^{t} d_p$, if s_i is an intermediate node such that it has d_p child nodes (where p = 1, 2,….., t). The process of aggregation takes place in the following way for each of the three types of the node.

2.1 Leaf Node Event Detection

2.1.1 Verification Phase

The main idea is that the data aggregation process takes place only after the verification in two ways. The first stage of verification is level based scheme. In this scheme, when two nodes need to get involved in communication, the level keys of the respective nodes are verified. If they are inconsistent, then the communication is denied. The main feature of this scheme is that the level key for each level gets updated dynamically after a certain time interval. This dynamic updating of the level keys ensures level wise security. It prevents the attacks by malicious nodes.

When a leaf node has to initiate communication with its parent node, initially level key of the leaf node and parent node is checked. If the verification succeeds, then the packet is received. The packet is decrypted to obtain the aggregated data. If the decryption succeeds, the packet is accepted. Verification takes place in the following way.

$node_{l_{n-1}} \rightarrow node_{l_{n-2}} : < E_{l_{n-1}}$ (id)>

Where $E_{l_{n-1}}$ is level key of the leaf node's level.

2.1.2 Data Collection Phase

Each leaf node collects the data value and sends it to its corresponding parent node. The process of aggregating the values begins from this node but aggregation of data does not take place in this level. In this level, only the data collection process is carried out. The packet that is sent by the leaf node v to its parent node s is shown below:

$node_{l_{n-1}} \rightarrow node_{l_{n-2}} : < E_{pr}$ (id,data)>

Where, data is the information collected by the leaf node v. $node_{l_{n-1}}$ indicates the packet transmitted by the node v at level n-1 and $node_{l_{n-2}}$ indicates the receiver node at level (n-2).

2.2 Intermediate Node Data Aggregation

2.2.1 Verification Phase

The value is forwarded to the intermediate nodes after the data collection process at the leaf node level. Before transmitting the packet to the node that acts as the parent node, the respective level key of corresponding sender and a receiving parent node is checked. If the verification succeeds, then the packet is transmitted. The intermediate nodes initially check the received value before performing further aggregation process. After the transmission, the decryption of the packet is performed. If the process succeeds, then the packet is accepted by the parent node else the packet is discarded.

The verification is carried out as follows:

$node_{l_{n-2}} \rightarrow node_{l_{n-3}} : < E_{l_{n-2}}$ (id)>

Where $E_{l_{n-1}}$ is level key of the sender node's level.

2.2.2 Data Aggregation Phase

Further aggregation process is carried on after successful verification. In this way of verification, the validity of the packet is ensured. Consider that node s is the child node of node p. The data that is sent by the intermediate node s to its parent node p is shown as follows:

$node_{l_{n-2}} \rightarrow node_{l_{n-3}} : < E_{pr}$ (id,data)>

Where data is the aggregation value calculated over the child nodes of node s. $node_{l_{n-2}}$ indicates the packet transmitted by the node s at level n-2 and $node_{l_{n-3}}$ indicates the receiver node at level n-3.

2.3 Base Station Data Aggregation

2.3.1 Verification Phase

The level keys of the base station and its child nodes are verified. This verification ensures the level wise security. If this verification succeeds, then the base station receives the aggregation value from its child nodes. It first decrypts the packet and if it is done correctly, then the packet is accepted by the base station else the packet is discarded. The verification takes place in the following way:

$node_{l_1} \rightarrow node_{l_0} : < E_{l_1}$ (id)>

Where E_{l_1} is level key of the sender node's level l_1 .

2.3.2 Data Collection Phase

The final process of aggregating the received values is performed at this level. The aggregation value at the base station is as shown below:

BS : < E_{pr} (id,data) >

Where, data is the aggregation value calculated over the child nodes of node n.

3. Extensions

3.1 Malicious Nodes

The transmitted packet is accepted only after verification at each stage of transmission. But in some cases, malicious nodes may transmit a forged aggregation value and retains the original aggregation value. This may lead to the acceptance of forged aggregation value by the receiver node which may lead to the acceptance of incorrect aggregation value by the base station. This can be overcome by re-verification by the grandparent nodes after receiving the packets from its child node.

The injection of false aggregation values by the malicious node is avoided by the grandparent node which can verify the received values by obtaining the values from the grandchild nodes. Initially, the level based scheme is applied and the decryption of the packet is performed. On successful verification, the grandparent node obtains the aggregated value. But before accepting the value for further aggregation, the grandparent nodes obtain the aggregated values from all the grandchild nodes (child nodes of grandparent node's child).

The grandparent node compares the received value from its child nodes and the value received from its grandchild nodes. If they match, then the aggregated values are accepted by the grandparent node and further aggregation is performed. Otherwise, if there is any inconsistency in the values, the received packet is discarded by the grandparent node. As indicated in Figure 3, parent node Q3 acts as a malicious node by sending false aggregation values to its parent node a. Node Q3 tampers the aggregation result ten received from its child nodes [R1, R2] and sends twenty two to its parent node P. To avoid this tampering by the parent node Q3, grandparent node P receives the aggregation result ten from its grandchild nodes [R1, R2]. This value is compared with the value received from its child node Q3. This indicates the inconsistency involved in it. Thus, leads to the discarding of the received packets.

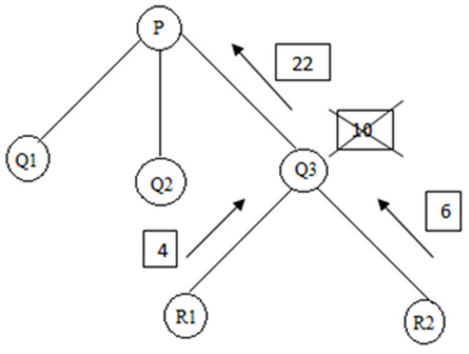

Figure 3. Malicious Aggregator.

4. Experimental Evaluation

NS-2 simulator has been used to perform a simulation setup and its study and thus, analyze the performance level.

4.1 Simulation Setup

The nodes are scattered randomly over a particular region such that the entire network is organized into a hierarchical structure. Size of the network n varies in the range of 40-300 nodes. If the network size is n<150, the area which is distributed is 200×200m^2. If the network size is in the range of 150<n<250, the area which is distributed is 300×300m^2. If the network size is in the range of 250<n<300, the area which is distributed is 400×400m^2. All the sensor nodes must get involved in the network. For that communication range of the sensor nodes can be adjusted. In this setup, 0.5W is set for transmitting and receiving the packets per unit time. Total battery power per node is set as 100 J. In this, the performance level of MAI and LBS [level based scheme] are compared.

4.2 Simulation Results

The number of packets being transmitted at each level of the hierarchy is shown in Figure 4. The number of packets being transmitted must be constant at each level. If any variation occurs at any level, it indicates that there is packet loss in that particular level and it indicates that the previous level contains a malicious node. In the Figure 4 level 2 collects 37 packets and at level 1 it receives the same number of packets. But at level 0, it receives only 35 packets which indicate the presence of malicious node in the level 1.

Performance comparison of MAI (Malicious Aggregator Identification) and LBS (Level Based Scheme) in terms of the packet loss is shown in Figure 5. In level

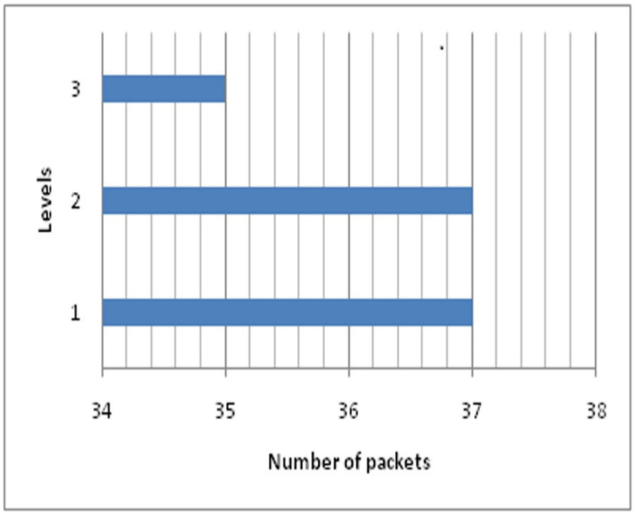

Figure 4. Malicious Node Detection.

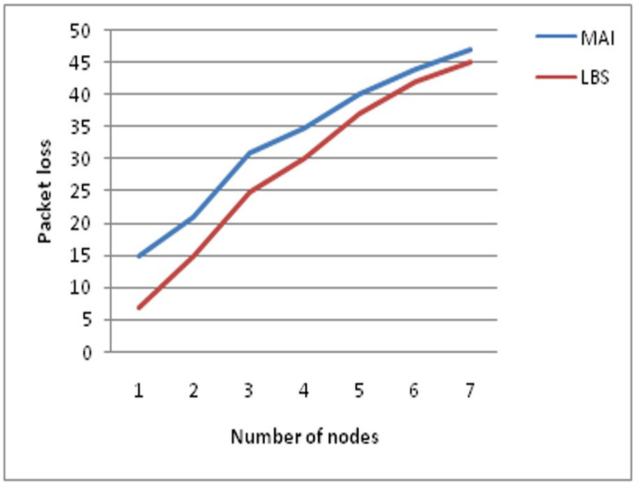

Figure 5. Performance under Different Network Size.

based scheme, the additional verification is performed when compared to MAI. In LBS, the level keys are verified at each stage of acceptance of received packets. The level keys are updated dynamically after a particular time interval. This feature reduces the packet loss as malicious nodes can be detected effectively when compared to the previous scheme. This enhances the security and improves the performance of the proposed technique.

5. Conclusion

In this work, security is ensured by implementing two stages of verification which makes sure that base stations do not agree with the false aggregation values and detects malicious nodes. Level based scheme has been applied in which the main feature is that the key for each level is updated dynamically after an interval of time. The experimental results indicate that the proposed system exhibits twice the performance of the previous schemes.

6. References

1. Jayalakshmi R, Baranidharan B, Santhi B. Attribute based spanning tree construction for data aggregation in heterogeneous wireless sensor networks. Indian Journal of Science and Technology. 2014; 7(S5):76–9.
2. Shanmugasundaram T, Nachiappan A. Multi-layer support based clustering for energy-hole prevention and routing in wireless sensor networks. Indian Journal of Science and Technology. 2015; 8(S7):236–46.
3. Sruthi K, Umamakeswari A. Link aware data aggregation mechanism based on passive clustering in wireless sensor network. Indian Journal of Science and Technology. 2014; 7(8):1236–42.
4. Baranidharan B, Akilandeswari N, Santhi B. EECDC: Energy Efficient Coverage Aware Data Collection in Wireless Sensor Networks. Indian Journal of Science and Technology. 2013; 6(7):4903–7.
5. Krishnamachari B, Estrin D, Wicker S, editors. The impact of data aggregation in wireless sensor networks. 2002 Proceedings 22nd International Conference on Distributed Computing Systems Workshops; IEEE; 2002.
6. Chan H, Perrig A, Song D, editors. Secure hierarchical in-network aggregation in sensor networks. Proceedings of the 13th ACM conference on Computer and communications security; ACM; 2006.
7. Li H, Li K, Qu W, Stojmenovic I. Secure and energy-efficient data aggregation with malicious aggregator identification in wireless sensor networks. Future Generation Computer Systems. 2014; 37:108–16.
8. Ben OS, Trad A, Youssef H, Alzaid H, editors. Secure data aggregation in wireless sensor networks. 2013, 12th Annual Mediterranean Ad Hoc Networking Workshop (MED-HOC-NET); IEEE; 2013.
9. Yang Y, Wang X, Zhu S, Cao G. SDAP: A secure hop-by-hop data aggregation protocol for sensor networks. ACM Transactions on Information and System Security (TISSEC). 2008; 11(4):18.
10. Roy S, Conti M, Setia S, Jajodia S. Secure data aggregation in wireless sensor networks. IEEE Transactions on Information Forensics and Security. 2012; 7(3):1040–52.
11. Cam H, Ozdemir S, Nair P, Muthuavinashiappan D, Sanli HO. Energy-efficient secure pattern based data aggregation for wireless sensor networks. Computer Communications. 2006; 29(4):446–55.

12. Ozdemir S, Xiao Y. Secure data aggregation in wireless sensor networks: A comprehensive overview. Computer Networks. 2009; 53(12):2022–37.
13. Przydatek B, Song D, Perrig A, editors. SIA: Secure information aggregation in sensor networks. Proceedings of the 1st International Conference on Embedded networked sensor systems; ACM; 2003.
14. Du W, Deng J, Han YS, Varshney PK, editors. A witness-based approach for data fusion assurance in wireless sensor networks. 2003 GLOBECOM'03 IEEE Global Telecommunications Conference; IEEE; 2003.
15. Ghosal A, Singh JP, editors. Secure data aggregation using some degree of persistent authentication in sensor networks. Conference on Mobile and Pervasive Computing (CoMPC); 2008.
16. Ma D, Tsudik G, editors. Forward-secure sequential aggregate authentication. 2007 SP'07 IEEE Symposium on Security and Privacy; IEEE; 2007.
17. Huang S-I, Shieh S, Tygar J. Secure encrypted-data aggregation for wireless sensor networks. Wireless Networks. 2010; 16(4):915–27.
18. Jariwala V, Jinwala D. A novel approach for secure data aggregation in wireless sensor networks. 2012. arXiv preprint arXiv:12034698.
19. Zhang W, Liu Y, Das SK, De P. Secure data aggregation in wireless sensor networks: A watermark based authentication supportive approach. Pervasive and Mobile Computing. 2008; 4(5):658–80.