Securing Node Capture Attacks for Hierarchical Data Aggregation in Wireless Sensor Networks

Bhoopathy, V. and R.M.S. Parvathi

1Department of CSE, Anmai Mathammal Sheela Engineering College, Tamil Nadu, India
2Department of CSE, Sengunthar College of Engineering, Tamil Nadu, India

Abstract - Serious security threat is originated by node capture attacks in hierarchical data aggregation where a hacker achieves full control over a sensor node through direct physical access in wireless sensor networks. It makes a high risk of data confidentiality. In this study, we propose a securing node capture attacks for hierarchical data aggregation in wireless sensor networks. Initially network is separated into number of clusters, each cluster is headed by an aggregator and the aggregators are directly connected to sink. The aggregator upon identifying the detecting nodes selects a set of nodes randomly and broadcast a unique value which contains their authentication keys, to the selected set of nodes in first round of data aggregation. When any node within the group needs to transfer the data, it transfers slices of data to other nodes in that group, encrypted by individual authentication keys. Each receiving node decrypts, sums up the slices and transfers the encrypted data to the aggregator. The aggregator aggregates and encrypts the data with the shared secret key of the sink and forwards it to the sink. The set of nodes is reselected with new set of authentication keys in the second round of aggregation. By simulation results, we demonstrate that the proposed technique resolves the security threat of node capture attacks.


I. INTRODUCTION

1.1 Wireless Sensor Networks

Wireless sensor networks consist of the latest technology that has attained notable consideration from the research community. Sensor networks consist of numerous low cost, little devices and are in nature self organizing ad hoc systems. The job of the sensor network is to monitor the physical environment, gather and transmit the information to other sink nodes. Generally, radio transmission ranges for the sensor networks are in the orders of the magnitude that is lesser that of the geographical scope of the unbroken network. Hence, the transmission of data is done from hop-by-hop to the sink in a multi-hop manner. Reducing the amount of data to be relayed thereby reduces the consumption of energy in the network. [1].

Wireless sensor network consists of a huge number of tiny electromechanical sensor devices that are capable of sensing, computing and communicating. These electromechanical sensor devices can be made use for gathering sensory information, like measurement of temperature from an extensive geographical area [2].

Many features of the wireless sensor networks have given rise to challenging problems [3]. The most important three characteristics are:

- Sensor nodes are exposed to maximum failures.
- Sensor nodes which make use of the broadcast communication pattern and have severe bandwidth restraint.
- Sensor nodes have inadequate amount of resources.

1.2 Data Aggregation

Data aggregation is considered as one of the basic dispersed data processing measures to save the energy and minimize the medium access layer contention in wireless sensor networks [4]. It is used as an important pattern for directing in the wireless sensor networks. The fundamental idea is to combine the data from different sources, redirect it with the removal of the redundancy and thereby reducing the number of transmissions and also saves energy [5]. The inbuilt redundancy in the raw data gathered from various sensors can be banned by the in-network data aggregation. In addition, these operations utilize raw materials to obtain application specific information. To conserve the energy in the system thereby maintaining longer lifetime in the network, it is important for the network to preserve high incidence of the in-network data aggregation [6].

1.3 Hierarchical Secure Data Aggregation

The following are the issues that are related to the security in the data aggregation of WSN [7]:

- Data Confidentiality: In particular, the fundamental security issue is the data privacy that protects the transmitted data which is sensitive from passive attacks like eavesdropping. The significance of the data confidentiality is in the hostile environment, where the wireless channel is more prone to eavesdropping. Though
cryptography provides plenty of methods, such as the process related to complicated encryption and decryption, like modular multiplication of large numbers in public key based on cryptosystems, utilizes the sensor’s power speedily.

- Data Integrity: It avoids the modification of the last aggregation value by the negotiating source nodes or aggregator nodes. Sensor nodes can be without difficulty compromised because of the lack of the expensive tampering-resistant hardware. The otherwise hardware that has been used may not be reliable at times. A compromised message is able to modify, forge and discard the messages.

Generally, in wireless sensor networks for secure data aggregation, two methods can be used. They are hop by hop encrypted data aggregation and end to end encrypted data aggregation [7].

- Hop-by-Hop encrypted data aggregation: In this technique, the encryption of the data is done by the sensing nodes and decryption by the aggregator nodes. The aggregator nodes aggregate the data and again encrypt the aggregation result. At the end, the sink node that obtains the last encrypted aggregation result decrypts it.

- End to End encrypted data aggregation: In this technique, the aggregator nodes in between does not contain any decryption keys and can only perform aggregation on the encrypted data.

1.4 Node Capture Attacks
The process of getting hold of the sensor node through a physical attack is termed as node capture attack. For example: uncovering the sensor and adding wires in any place. This attack essentially differs from getting hold of a sensor via certain software bug. Since sensors are typically supposed to operate the same software, specifically, the operating software which discovers the suitable bug permits the adversary to manage the entire sensor network. Distinctly, the node capture attacks can be set over a small segment of adequately large network, [8]

The blend of passive, active and physical attacks by an intellectual adversary results in node capture attack. The adversary initializes an attack by gathering the data’s about WSN by overhearing something on message exchanges. This is performed either locally to single adversarial device or via entire network with the help of several adversarial devices organized in the entire network. Along with passive learning, the adversary dynamically takes part in network protocols, inquiring the network regarding the information and injecting malicious information in the network.

The adversary performs the physical attacks, following active and passive learning. To enhance the function of the attack related to certain attack objective, the gathered information can be utilized to aid the adversary in choosing the sensor node. [9]

There are two types of node captures possible:

- Random node capture
- Selective node capture

The above node captures varies in the key distribution information to the attacker. The attacker should minimum capture hundreds of sensor nodes during selective node capture attacks. [12]

1.5 Problem Identification
In sensor node compromise technique, there is a initiation of node capture attack where the adversary physically captures the sensor nodes, removes them, compromises and redeploys them in the network. Following the redeployment of the compromised nodes, it builds up a variety of attacks through compromised nodes. The forceful attacker weakens the sensor network protocols along with the formation of clusters, routing and data aggregation and hence resulting in recurrent disruption of network operations. Therefore, the node capture attacks are unsafe and need to be identified as soon as possible for reducing the damages caused by them. [10]

During the node capture attacks, the adversary attempts to tamper the node physically for extracting the secrets of the cryptography. Based on the security architecture of the network, this type of attack is highly destructive and furthermore results in influential insider attacks. [11]

A security issue of WSN corresponds to node capture attack which leads to compromise in the communication of a whole sensor network. [13]

An Energy Efficient Secure Data Aggregation Protocol for wireless sensor networks, we incorporate the authentication and security to maintain the efficiency of the data aggregation. Whenever a sensor node wants to send data to another node; first the sensor node encrypts the data using a key and sends it to the aggregator. For integrity of the data packet, a MAC based authentication code is used [14]. The security problem of WSN such as node capture attacks is not taken into consideration. This node capture attack is harmful for network communication in network data aggregation, routing and so on.

Secure Authentication Technique for Data Aggregation in Wireless Sensor Networks, during first round of data aggregation, the aggregator upon identifying the detecting nodes selects a set of nodes randomly and broadcast a
unique value which contains their authentication keys, to the selected set of nodes. When any node within the set wants to send the data, it sends slices of data to other nodes in that set, encrypted with their respective authentication keys. Each receiving node decrypts, sums up the slices and sends the encrypted data to the aggregator [15]. The security problem of WSN such that hierarchical data aggregation is not considered.

We propose a Securing Node Capture Attacks for Hierarchical Data Aggregation in Wireless Sensor Networks

II. RELATED WORKS

Kashif Kifayat et al [13] proposed a novel and distinct Structure and Density Independent Group Based Key Management Protocol (DGKE). The protocol offers a better secure communication, secure data aggregation, confidentiality, and resilience against node capture and replication attacks using reduced resources. The drawback of this approach is that security issues are not considered which impacts significantly on key management.

Yupeng Hu et al [16] proposed a robust authentication scheme (RAS) for filtering false data in wireless sensor networks. In RAS, each big event is divided into several small event chunks, every one of which is endorsed by witness nodes both with dynamic authentication tokens from one-way hash chain and their secret keys pre-loaded from the key pool. This way, compromised nodes, even in possession of all endorsement keys for the data reports will not able to fabricate or modify the reports.

Mohamed Hamdy Eldefrawy et al [17] proposed a key distribution protocol based on the public key cryptography. The protocol establishes pairwise keys between nodes according to a specific routing algorithm after deployment, instead of loading full pair-wise keys into each node. The proposed scheme comes to circumvent the shortage of providing the re-keying property of nodes.

Eitaro Kohno et al [18] proposed a new method resilient to node capture attacks. This method utilizes secret sharing scheme to disperse confidential information without the need of a secret key. This method is implemented on the motes nodes and it is more effective as the number of hops-to-sink node increases. On the other hand the increased overhead is observed on short hop node. They have also shown a countermeasure capable of reducing excess dispersals without degrading the resilience against node capture attacks.

Mauro Conti et al [19] proposed two efficient and distributed solutions. In the first proposal, Simple Distributed Detection (SDD), the attack is detected using only information local to the nodes. The second solution, the Cooperative Distributed Detection (CDD), exploits node collaboration to improve the detection performance. CDD outperforms both SSD in a meaningful scenario. Moreover, the proposed solutions do not rely on any specific routing protocol—we only use direct range communications and message flooding.

Ka-Shun Hung et al [20] investigated the effects of different node capture attack patterns on state-of-the-art key management schemes. They proposed two recovery strategies, namely link replacement strategy and node replenishment strategy to replace the compromised region, respectively. This proposed approach achieves significant improvement in terms of network resilience.

Haowen Chan et al [21] Secure hierarchical in-network data aggregation is guaranteed to identify 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 system incurs only O(Δlog2 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.
III. PROPOSED WORK

3.1 System Architecture

3.2 Algorithm

Algorithm Node_Capture_Attack (node, aggregator, key, cluster, AGG_adv)

// u_i is a member node in cluster C_j where j = 1 to n.
// A_j is the aggregator of the cluster C_j.
// AGG_adv represents Aggregator Advertisement Message

// R_1 is the first round of aggregation.
// TS_1 is R_1's respective time stamp.
// A_j possesses a secret key (k_j^{sec}) which is shared with the sink.

A_j \xrightarrow{AGG_adv} u_i

// In R_1, the aggregator broadcasts the AGG_adv to all the nodes.

u_i \xrightarrow{ACK} A_j

// u_i sends acknowledgment (ACK) message to A_j.
// ACK = \{w_i, g\} Where w_i = node's ID, g = node's category.
// Based on ACK messages, the A_j selects c nodes (c \leq n) randomly.

Set Q = \{u_1, u_2, ..., u_c\}. // selected c nodes are represented by the set Q

A_j \xrightarrow{V} Q

V = [(w_1, K_{w_1}), (w_2, K_{w_2}), ..., (w_c, K_{w_c})]

// the A_j broadcasts a set of unique values V to all nodes in Q.
// V consists of the node ids of Q and their authentication key.
// K_{w_i} denotes the authentication keys of the corresponding node w_i.

u_2 \xrightarrow{encr(1to(c-1))} u_3

X = 1 + 2 + ... + c. // X represents data which sliced into c pieces.

// assume u_2 wants to send the data to any node. First u_2 send encrypted data to nearest node u_3.
// In c slices, one of them is kept inside that node itself.

X (1 to (c-1)) \xrightarrow{decr(1to(c-1))} u_3

// u_3 waits for a time t, which assures that all slices of this round of aggregation are received.

1 + 2 + ... + (c-1) = S_c // sums up the received slices

u_3 \xrightarrow{encr(S_c)} A_j

// S_c is again encrypted with the authentication key of the respective node and sent to the A_j.

A_j \xrightarrow{MAC(ED, TS_1)} Sink

// A_j aggregates and encrypts the data with the shared key k_j^{sec} and forwards it to towards sink.
// The message in the form MAC (ED, TS_1) where TS_1 = time stamp, ED = encrypted data.

If (TS_1 \rightarrow expires)

{ R_1 \rightarrow ends
R_2 \rightarrow starts
TS_2 \rightarrow begins
}
Consider the node 2 in Figure 2. When it wants to send data to its neighboring nodes, it slices the data (X) into 8 pieces (since network size u=8). It holds the one of the slices with it. The remaining slices are encrypted with their respective authentication keys and sent to rest of the nodes.

When the node 1 receives the encrypted data slice from node 2, it decrypts the slice using its authentication key K_1. Then Node 1 waits for reception of the rest of the slices until time t. When t expires, the node 1 stops receiving the data slice. After complete decryption of the received slices, the node 1 sums them up along with the slice within it and this sum is represented as S_1.

\[ S_1 = C_{11} + C_{21} + C_{41} + C_{81} \]

Similarly the summed data of other nodes are as follows.

\[ S_2 = C_{12} + C_{22} + C_{32} + C_{42} \]
\[ S_3 = C_{23} + C_{33} + C_{43} + C_{53} \]
\[ S_4 = C_{14} + C_{24} + C_{34} + C_{44} + C_{54} \]
\[ S_5 = C_{35} + C_{45} + C_{55} \]
\[ S_6 = C_{66} + C_{76} + C_{86} \]
\[ S_7 = C_{47} + C_{57} + C_{67} + C_{77} + C_{87} \]
\[ S_8 = C_{18} + C_{68} + C_{78} + C_{88} \]

The node 1 encrypts S_1 with k_1 and sent to the aggregator A_1. The aggregator encrypts the data with the secret shared key (k'_sec) and forwards it to the sink.

### IV. SIMULATION RESULTS

#### 4.1 Simulation Setup

The performance of SNCAHDA approach is evaluated through Network Simulator Version-2 Ns-2 [20] simulation. A random network deployed in an area of 351 X 351 m is considered. Initially 30 sensor nodes are placed in square grid area by placing each sensor in a 50x50 grid cell. 4 phenomenon nodes which move across the grid (speed 5m/s) are deployed to trigger the events. 4 aggregators are deployed in the grid region according to our protocol. The sink is assumed to be situated 100 meters away from the above specified area. In the simulation, the channel capacity of mobile hosts is set to the same value: 2 Mbps. The distributed coordination function (DCF) of IEEE 802.11 is used for wireless LANs as the MAC layer protocol. The simulated traffic is CBR with UDP source and sink. The number of sources is fixed as 4 around a phenomenon.
### Table 2: Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Nodes</td>
<td>30</td>
</tr>
<tr>
<td>Area Size</td>
<td>351 X 351</td>
</tr>
<tr>
<td>Mac</td>
<td>802.11</td>
</tr>
<tr>
<td>Routing protocol</td>
<td>DSDV</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>50 sec</td>
</tr>
<tr>
<td>Traffic Source</td>
<td>CBR</td>
</tr>
<tr>
<td>Packet Size</td>
<td>50 bytes</td>
</tr>
<tr>
<td>Rate</td>
<td>50 bytes</td>
</tr>
<tr>
<td>Transmission Range</td>
<td>150m</td>
</tr>
<tr>
<td>No. of events</td>
<td>4</td>
</tr>
<tr>
<td>No. of Sources</td>
<td>1, 2, 3, 4, 5</td>
</tr>
<tr>
<td>No. of attackers</td>
<td>1, 2, 3, 4, 5</td>
</tr>
<tr>
<td>Speed of events</td>
<td>5 m/s</td>
</tr>
</tbody>
</table>

### Table 4: Comparison of SNCAHDA and SATDA based on sources

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SATDA</th>
<th>SNCAHDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay</td>
<td>High</td>
<td>Comparatively low</td>
</tr>
<tr>
<td>Packet Delivery Ratio</td>
<td>Slightly low</td>
<td>High</td>
</tr>
<tr>
<td>Energy</td>
<td>High</td>
<td>Comparatively Low</td>
</tr>
<tr>
<td>Packet Drop Ratio</td>
<td>Comparatively High</td>
<td>Low</td>
</tr>
<tr>
<td>Throughput</td>
<td>Slightly Low</td>
<td>High</td>
</tr>
</tbody>
</table>

### 4.2 Performance Metrics

The performance of Securing Node Capture Attacks for Hierarchical Data Aggregation in Wireless Sensor Networks (SNCAHDA) protocol is compared with our previous work Secure Authentication Technique for Data Aggregation (SATDA) protocol [15]. The performance is evaluated mainly, according to the following metrics.

- **Average end-to-end delay**: The end-to-end delay is averaged over all surviving data packets from the sources to the destinations.
- **Average Packet Delivery Ratio**: It is the ratio of the number of packets received successfully and the total number of packets transmitted.
- **Average Energy**: It is the average energy consumption of all nodes in sending, receiving and forward operations.
- **Average Packet Loss**: It is the average number of packet dropped at each receiver.
- **Throughput**: It is the number of packets successfully received by the receiver.

### Table 3: Comparison of SNCAHDA and SATDA based on attackers

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SATDA</th>
<th>SNCAHDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay</td>
<td>Low</td>
<td>Comparatively high</td>
</tr>
<tr>
<td>Packet Delivery Ratio</td>
<td>Slightly low</td>
<td>High</td>
</tr>
<tr>
<td>Energy Consumption</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Packet Drop Ratio</td>
<td>Comparatively High</td>
<td>Low</td>
</tr>
<tr>
<td>Throughput</td>
<td>High</td>
<td>More High</td>
</tr>
</tbody>
</table>

### A. Based on Attackers

In our initial experiment, we vary the number of attackers as 1, 2, 3, 4 and 5.

#### Fig 3: Attackers Vs Delay gives the average end-to-end delay for both protocols when the number of nodes is increased. We can see that the average end-to-end delay of our proposed SNCAHDA protocol is less than the existing SATDA protocol.

#### Fig 4: Attackers Vs Delivery ratio gives the packet delivery ratio for both protocols when the number of nodes is increased. We can see that the packet delivery ratio of our...
proposed SNCAHDA protocol is higher than the existing SATDA protocol.

**Fig 5:** Attackers Vs Energy gives the energy consumption for both protocols. We can see that the energy consumption of our proposed SNCAHDA protocol is less than the existing SATDA protocol.

**Fig 6:** Attackers Vs Packet drop ratio gives the Packet drop ratio for both protocols. We can see that the Packet drop ratio of our proposed SNCAHDA protocol is less than the existing SATDA protocol.

**Fig 7:** Attackers Vs Throughput gives the throughput for both protocols. We can see that the Throughput of our proposed SNCAHDA protocol is higher than the existing SATDA protocol.

**Fig 8:** Sources Vs Delay gives the average end-to-end delay for both protocols when the number of sources increased. We can see that the average end-to-end delay of our proposed SNCAHDA protocol is less than the existing SATDA protocol.

**Fig 9:** Sources Vs Delivery ratio gives the packet delivery ratio for both protocols. We can see that the packet delivery ratio of our proposed SNCAHDA protocol is higher than the existing SATDA protocol.

**Fig 10:** Sources Vs Energy gives the energy consumption for both protocols. We can see that the energy consumption of our proposed SNCAHDA protocol is less than the existing SATDA protocol.

**B. Based on Sources**

In the second experiment, we vary the number of sources as 1, 2, 3 and 4.
Fig 11: Sources Vs Drop gives the Packet drop ratio for both protocols. We can see that the Packet drop ratio of our proposed SNCAHDA protocol is less than the existing SATDA protocol.

Fig 12: Sources Vs Throughput gives the throughput for both protocols. We can see that the Throughput of our proposed SNCAHDA protocol is higher than the existing SATDA protocol.

V. CONCLUSION
In this paper, we have proposed Securing Node Capture Attacks for Hierarchical Data Aggregation in Wireless Sensor Networks. During first round of data aggregation, the aggregator upon identifying the detecting nodes selects a set of nodes randomly and broadcast a unique value which contains their authentication keys, to the selected set of nodes. When any node within the set wants to send the data, it sends slices of data to other nodes in that set, encrypted with their respective authentication keys. Each receiving node decrypts, sums up the slices and sends the encrypted data to the aggregator. The aggregator aggregates and encrypts the data with the shared secret key of the sink and forwards it to the sink. In the second round of aggregation, the set of nodes is reselected with new set of authentication keys. By simulation results, we have shown that the proposed approach rectifies the security threat of node capture attacks in hierarchical data aggregation.

REFERENCES
of information assurance and security, Vol 2, Issue 2, June 2007


[22] Network Simulator: www.isi.edu/nsnam/ns