A Secure Data Transmission Mechanism for Sensor Network Communication

Inshil Doh† · Kijoon Chae‡

ABSTRACT

For reliable sensor network communication, secure data transmission mechanisms are necessary. In our work, for secure communication, we cluster the network field in a hexagonal shape and deploy nodes according to Gaussian distribution. After node deployment, clusterheads and gateway nodes in each cluster play the role of aggregating and delivering the sensed data with security information all the way to the base station. Our mechanism decreases the overhead and provides good performance. It also has resilience against various routing attacks.

Key Words: Sensor Networks, Security, Cluster, Secure Routing

1. Introduction

Sensor network is considered to be the core technology for our ubiquitous computing environment because it can be applied efficiently in various applications such as battlefield surveillance, medical monitoring, emergency response, and so on. However, to apply the technology in real environment, security is the critical issue. Sensor network is more vulnerable to various attacks because of its basic constraints such as energy, short transmission range, low computation power, and so on[1]. However, many security mechanisms developed for the Internet or ad-hoc network cannot be applied directly to WSNs (Wireless Sensor Networks) because of the characteristics of WSN. Among many research areas in sensor networking, a lot of studies have been done in routing area. They are usually focusing on how to find the routes to base station and how to manage the routes efficiently. However, without security mechanisms, data cannot be delivered safely to the destination, and the network can be collapsed. Several secure routing mechanisms have been proposed recently, however, they still have shortcomings such as overhead or performance problems. In our work, we propose a clustered network, and by adding security information to data and by sending the data through multiple routes, we provide secure data transmission with good performance and make the network resilient to various routing attacks.

The paper is organized as follows. After introducing
the motivation of our work in chapter 1, we summarise the related work in chapter 2, and mention the network model and the basic assumptions in chapter 3. We propose a secure data transmission mechanism in chapter 4. Chapter 5 shows the simulation result and analysis of our proposal considering possible attacks in sensor routing. Finally, conclusions are mentioned in chapter 6.

2. Related Work

For sensor network routing, the major research topic is how to find the routes to destinations and how to maintain the routes which have been found. However, when security is considered, the more important aspect is how to defend against various attacks which are possible in sensor network and how to keep the throughput above certain degree. When compared to routing protocols, researches for secure routing in sensor network are insufficient. Several researches which have been proposed are as follows.

ARAN[2] and SAODV[3] use public-key cryptography, which is not memory and energy efficient for sensor networks. SEAD[4], Ariadne[5], and SRP[6] use symmetric cryptography or hashing, but require maintenance of routing tables by distance vector algorithms. However, for large-scale networks, maintaining routing tables requires a lot of consumption of memory and energy, and it is also vulnerable to security attacks. SPINS[7] and TinySec[8] provide secure channels for use by otherwise unsecured protocols. They are still inadequate defenses when nodes are compromised. INSSENS[9] is designed to tolerate node compromise and uses a variety of efficient mechanisms to establish routing. However, it is based on centralized topology collection and route computation. Some geography-based routing algorithms have also been proposed. In addition to plain geographic forwarding such as GF[10] and IGF[11], GPSR and descendents[12, 13] extend GP to route around voids by traversing faces of a planar subgraph until greedy forwarding can resume. ZRF[14] divides the network into variable size zones and allows different algorithms for intra-zone and inter-zone routing. It is lightweight and efficient, but does not consider security. SIGF[15] designed by Wood et al., achieves secure routing properties using local keys and nondeterministic selection of forwarding nodes. But it assumes that all nodes know their own geographic locations.

3. Network Model and Assumptions

For our proposal, nodes are deployed according to Two-dimensional Gaussian distribution model. Before node deployment, sensor network field is clustered in hexagonal shape. Nodes are deployed from the helicopter in the air at each deployment point (x, y) where,

\[ f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}} \]

The height of the helicopter is dependent on the value \( d \), and according to above equation, more nodes are deployed near the deployment points, as in Fig. 1.

After nodes are deployed, CH (clusterhead) and GW (gateway) nodes are selected through neighbor detection process. This process is described in detail in the next chapter. All nodes are static, and collaborative attacks by multiple adversaries are not considered. We also assume that there is no adversary in node deployment and key setup phases.

![Node Deployment Model according to Two-Dimensional Gaussian Distribution for a Cluster with Deployment Point(50,50)](image)

4. A secure Data Transmission Mechanism

4.1 Prerequisites for the Proposal

4.1.1 Network Clustering and Nodes Election

When nodes are deployed with proper key information, they detect their own neighbors and elect CH and GW nodes. Every node broadcasts the number of neighbor nodes from the same cluster and the number of neighbors from different clusters. Nodes with more number of neighbors from the same cluster usually become CHs. And when some nodes have neighbors from different clusters, they usually become GW nodes. We need multiple number of CHs because of the energy depletion problem. This is discussed in 4.5. When some nodes are decided to be CHs or GWs, they setup pairwise keys
which they need to use for secure communication. Key setup process is out of the scope of this work.

The number of clusters varies according to network scale. Let \( d = a \cdot \sigma \) denotes the distance between two neighboring deployment points in each hexagonal cluster as in Fig. 2. Given a fixed \( a \), the value of \( a \) determines the size of clusters and further affects the degree of connectivity \( P_c \). According to normal distribution, 99.87% of nodes are located within 3\( \sigma \) from their deployment point[19]. If \( a \) is too large, it causes the deployed network partitioned and lowers the degree of connectivity. If \( a \) is too small, the size of grids becomes small and it makes more nodes fall into those non-neighboring grids and more pairs of neighbors share no secret key, thus, also lowers the degree of connectivity. Therefore, a high degree of connectivity can only be achieved when the value of \( a \) is in an appropriate range. When the shortest length between two non-neighboring clusters is larger than 6\( \sigma \), we can guarantee that transmission range of a sensor node does not reach non-neighboring clusters, which means the value \( a \) is about 2\( \sqrt{3} \)[19]. Considering this condition and network scale, we determine the number of clusters and the number of sensor nodes in each cluster.

4.1.2 Required keys

We have proposed a pairwise key establishment mechanism using deployment information. Several pairwise keys are used for secure sensor communications.

- Keys between every pair of sensor nodes in a cluster. Every pair of sensor nodes in the same cluster can establish pairwise keys using preassigned key-related information. Every node can compute pairwise keys with any node it wants to communicate with. Especially, all nodes need to setup pairwise keys with the CH in the cluster in which the sensor nodes are located.

- Keys between every pair of neighboring sensor nodes from different clusters from each other. Neighboring sensor nodes which are located in different clusters from each other have pairwise keys. These nodes play the role of gateway nodes.

- Keys between every pair of CHs. Every CH can compute pairwise keys with other CHs using preassigned information for secure communication.

- Keys between every node and the BS (Base Station). Every node is predistributed respective pairwise key with the BS before deployment.

(Fig. 2) Clustered Network Architecture

4.2 Secure Data Transmission Mechanism

As in Fig. 2, every cluster is hexagonally clustered before deployment, and it has three dimensional coordinates \((x, y, z)\). Using the coordinates, CHs decide next routing paths dynamically in a greedy manner, and deliver the related packets with proper security information. In the initiating cluster, the basic transmission method is broadcasting, which means every node in the same cluster knows that an event has been sensed by some nodes in the same cluster. This is important for detecting selective forwarding or sinkhole or blackhole attacks. Every CH gathers the sensed data from its own member sensor nodes and then delivers the aggregated data to the CH through the other CHs. Every CH computes the difference between its own coordinates and those of the BS, and then chooses the next CH. The steps of transmitting packets are shown in Fig. 3.

(Fig. 3) Data transmission mechanism

4.2.1 Event sensing and event packet transmission to CH by each sensor node

Sensor nodes sensing events generate data packets and deliver them to their own CHs with two MAC values computed using the MAC keys derived from pairwise keys with the BS(MACSN-BS) and CH(MACSN-CH), respectively. Some of these values are selected by the CH and finally used by the BS to check the authenticity of the initiating CH. Intermediate sensor nodes just deliver
the data to their CH without verifying the MAC values.

4.2.2 Report packet generation

When CHs receive the event packets, they verify the MAC_{SN-CH} values and then generate a new report packet which has next intermediate destinations. Gateway nodes are randomly selected among the multiple gateway nodes by the CH and the next clusters are selected according to the path selection rules which are described in the next sub-section. In this step, the CH adds MAC_{CH-BS}, MAC_{SN-BS}', MAC_{CH-CID} (Table 1). Among many MAC_{SN-BS}'s from each sensor node, some are selected according to threshold and then contracted to MAC_{SN-BS}'. This MAC_{SN-BS}' and the IDs of each sensor node whose MAC_{SN-BS} has been selected by the CH and used to compute MAC_{SN-BS}' are included in the report packet. The initiating CH adds its own MAC with the BS to the report packet. These MAC_{CH-BS}'s are used by the BS to check the authenticity of the data in the report packet. Each CH also adds MAC_{CH-CID} value which will be verified by the next intermediate CH. All pairs of neighboring sensor nodes have pairwise keys of their own. They may just deliver the data to their CH without MAC verification, or they can authenticate the data through hop-by-hop manner. This is decided by the security policy. Fig. 4 shows the format of the event packet and the report packet by sensor nodes and CHs, respectively. MAC_{SN-BS}' is the contracted form to reduce the length of packets as follows.

$$MAC_{SN-BS}' = MAC_{SN-BS} \oplus MAC_{SN-BS} \oplus MAC_{SN-BS} \oplus ...$$

A CH chooses more than two different paths and GW nodes to next CH to prevent selective forwarding attack by insider attackers. In this way, even an attacker drops the packets in one of the path, another packet can be delivered safely through the alternate path. If we want a higher security level, we can add MAC_{CH-GW} to report packet for the GW node to verify the authenticity of the report packets.

4.2.3 Path selection and transmission of report packets

Every cluster is assigned to coordinates (i,j,k). When a CH tries to send a report packet, it computes the difference between the coordinates of itself and those of the BS. It needs to increment or decrement i, j, or k values. As in Fig. 5, when the cluster is located at the same axis with the BS, it needs to increment or decrement i, j, or k to get to the BS and sends the packet toward next three CHs. Each cluster has no common axis with the BS, it chooses next two clusters towards BS by changing i, j, or k values. CHs pick the operation randomly every time they need to make the movement to decrease the energy consumption of certain nodes on the route to the BS. In this multipath routing method, we can provide resilience against several attacks. Each CH can choose several movements according to directions in Fig. 5. Through some of the options, we expect efficient movement which is similar to diagonal movement for square clusters by moving just one step. This is possible because we choose hexagonal shape and modified coordinates. Each CH can deliver the report packet to the next cluster on the route, and the next cluster can choose another cluster to send the packets. And finally, CHs(CHp, CHq, or CHr) which are next to
cluster in which BS is located transfer the report packet to the cluster which is including the BS.

4.2.4 Verification by the intermediate CH
When the CH at a cluster gets the report packet, it checks the MAC_{CH-CH} value, decides the next CH, generates new MAC_{CH-CH} for the next CH, and delivers the packet toward the CH2.

4.2.5 Verification by the BS
When the report packet is delivered to BS through previous steps, BS finally checks MAC_{CH-BS}, MAC_{CH-BS}, and MAC_{CH-BS}'. Because it gets the list of the initiating sensor nodes, it can generate MAC_{CH-BS} values, contract them, and then compare the contracted value with MAC_{CH-BS}'. If the two values are not the same, BS decides that the packet has been compromised by one of the CHs or sensor nodes and discards it. When all the MAC values are verified, BS accepts the report packet as authentic one.

4.3 Routing loop problem
Because each CH needs two or three directions according to its relative position to the destination node in a greedy fashion, the packets cannot move far from the destination, but move toward the destination in each movement. Even some compromised CHs try to move farther from the original destination, the next CH can correct the direction of the routes if it is not colluding with the former one. So, routing loops are not formed. If there are more correctly functioning CHs than compromised nodes, the packets can get to the destination.

4.4 Node addition and deletion problem
When some nodes are added in the network field later, they need to detect neighbor nodes. And in this phase, to prevent Hello-flood and Sybil attack, they need to follow several steps in Fig. 6. When a node broadcasts Hello messages in its transmission range, the neighboring nodes which have heard the Hello message reply with nonces. After the original node gets the reply with nonce, it computes the hashed value of the nonce and broadcasts the value again. When the original node verifies the hashed nonce, it confirms the node as its neighbor. After checking if some nodes are authentic neighbors, new nodes establish pairwise keys with neighbors using related information.

4.5 Candidate CHs
CHs usually consume more energy than the other nodes because they need to generate report packets and relay the packets to BS. Because the CHs are selected among normal sensor nodes, their lifetimes are shorter than the other nodes. To make the lifetime of the network longer, several number of candidate CHs are required. In the CH selection process, a couple of nodes which have more neighbor nodes than the others are selected as candidate CHs. And when the old CH consumes all energy it has, another candidate CH which has been sleeping notifies that it is the new CH and starts playing the role of the new aggregator.

5. Performance analyses

5.1 Simulation model
To evaluate the performance, we simulate our proposal in the presence of several compromised nodes, using a simulator ns-2[17]. Simulation model is shown in Fig. 7, and the metric used is described in <Table 2>. As in Fig. 7, Events occur in cluster 3 periodically, and CH3 aggregates the event sensing data and delivers the report packet to BS through other CHs.

![Simulation Model](Fig. 7)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>210</td>
</tr>
<tr>
<td>Terrain</td>
<td>150 x 140 meters</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>7</td>
</tr>
<tr>
<td>Node deployment</td>
<td>Two dimensional Gaussian Distribution</td>
</tr>
<tr>
<td>Simulation time</td>
<td>100 sec</td>
</tr>
<tr>
<td>Transmission range</td>
<td>10m</td>
</tr>
<tr>
<td>Application</td>
<td>CBR</td>
</tr>
<tr>
<td>Radio transmission model</td>
<td>Two-Ray</td>
</tr>
<tr>
<td>Packet transmission rate</td>
<td>5 pkt/sec, 10 pkt/sec</td>
</tr>
<tr>
<td>PHY, MAC</td>
<td>822.15-4[30]</td>
</tr>
<tr>
<td>$l$</td>
<td>$2\sqrt{3}\sigma$</td>
</tr>
</tbody>
</table>
5.2 Simulation Result

We consider two kinds of attacks, one of which is selective forwarding which can cover several other attacks such as sinkhole and blackhole attacks. The other one is sybil attack, and we simulate ID stealth and ID fabrication, respectively.

As in Fig. 8, more packets drop as more traffic is generated and more number of attackers exist. Especially, when the number of selective forwarding attackers are more than 20, which means about 10% of the nodes are compromised, delivery ratio drops as low as 60%. However, in any case, our proposal shows good performance under these attacks.

Fig. 9 shows performance with sybil attackers, one of which is fabricating IDs and the other one is stealing the other nodes' IDs. In both cases, performance degrades as more number of sybil nodes attack the network. Compared to selective forwarding attacks, sybil attacks have the worse influence on the delivery ratio. With the number of attacker nodes fifteen and packet transmission rate 5 pkt/sec, delivery ratio drops as low as 60%. Under this situation, our proposal still shows almost perfect packet delivery ratio. The result shows that ID fabrication has a bit more serious harm than ID stealing attack as the number of attackers increases.

5.3 Defense against routing attacks

Researchers have identified several severe routing protocol attacks[18]. We summarize them and describe how our proposal defends against the attacks. Attacks can be carried out by three different kinds of nodes. First kind is the sensor nodes which relay the report packets to another CH. Second one is the initiating CH, and the third one is intermediate CHs which relay the report packets all the way to the BS. Event sensing sensor nodes do not make big effect because we need more than certain number of event sensing nodes to make the initiating CH generate a report packet. Even if some of the nodes are compromised and make wrong data, CH can filter the modified data considering the other data gathered.

- Routing loop An attacker injects malicious routing information that causes other nodes to form routing loops. Packets injected into the loop (both by legitimate and malicious nodes) are then sent in a circle, wasting precious communication and battery resources. As mentioned previously, in our proposal, CHs deliver the packets in a greedy method toward the BS. And at the clusters next to the BS, only one cluster in which BS is located is chosen. Thus routing loops are not formed in our proposal.

- Bogus routing By spoofing, altering, or replaying routing information, attackers are able to attract or redirect network traffic, increasing end-to-end delay, etc.

  - The initiating CH cannot generate bogus report packets because it needs to add not only the MAC_{CH-BS} by itself but also MAC_{SN-BS}, which is computed using MAC_{SN-BS} from normal sensor nodes and also the ID list of each sensor node.

  - The intermediate CH cannot modify the original report packets because initiating CH has added MAC which can be verified by the BS at the final stage.

  - When the intermediate sensor node tries to modify the report packet, next intermediate CH detects this when verifying MAC_{CH-CH}.

- Selective forwarding Attackers selectively forward packets instead of faithfully forwarding all received packets or completely dropping all packets.

In our proposal, we use multi-path to relay the report packets to BS, and even if an intermediate CH or an intermediate sensor node selectively drops the report packets, some packets can be delivered to BS through the alternate path. If an initiating CH tries not to deliver a report packet, this is detected by the member sensor nodes which have sensed that there has been an event because the event sensing information is broadcasted to all sensor nodes in the cluster. The fact that the CH did not generate a report packet is delivered to BS by member sensor nodes through flooding.
• Blackhole attack and Sinkhole attack In blackhole attack, a malicious node advertises a short distance to all destinations, attracting traffic meant for those destinations. In sinkhole attack, the attacks typically work by advertising attractive routing information from a compromised node to its neighbors. For both cases, our proposal lets the CH deliver the report packets to multiple neighbor CHs all the way to BS. So, even one of the routes are corrupted, the packets can be transmitted through another way. In addition, routing information is not stored, so the attackers cannot attract traffic with fake routing information.

• Sybil attack A malicious node fabricates or steals multiple fake identities to perform attacks. In geographic routing protocols, fake identities can claim to be at multiple locations. In our proposal, through the handshake Hello exchange phase, every node gets the authentic neighbor IDs and detects messages from the fake ID.

• Wormhole attack In wormhole attack, an attacker tunnels packets from one location to another one in the network. Wormhole attack can be defended through the same way as sinkhole or blackhole attack in our proposal.

• HELLO flood attack In HELLO flood attack, an adversary with a powerful transmitter reaches every node in the network, and pretends to be a neighbor. If an attacker broadcasts Hello messages with high transmission power to other nodes, the messages reach non-neighboring nodes. In our proposal, nodes receiving these broadcast messages do not just consider the nodes as neighbors but reply with nonce values and these reply messages with normal transmission power cannot reach the attacker. As a result, the attacker cannot make non-neighboring nodes consider the attacker as their neighboring node.

• General DoS (Denial-of-Service) attack Normal sensor nodes cannot carry out DoS attack because their CH needs certain number of event sensing packets from different sensor nodes to make report packets. The initiating or intermediate CH cannot fabricate the original event sensing packets from normal sensor nodes because BS verifies the MAC\textsubscript{SN-BS}' to check the authenticity of the report packet. They cannot replay the old messages because the Timestamp is included in the message and the MAC value is computed including the Timestamp.

5.4 Overhead analysis

For storage overhead, nodes do not need large memory because they do not keep the routes to BS. For delivering the sensed data to CH, normal sensor nodes just flood the event packets in their own clusters, and for delivering the report packets to next cluster they just keep the information of their own neighbor nodes. CHs need to keep the list of the GW nodes in its own cluster, which does not cost much.

For computing and verifying MACs, extra energy is required. However, verifying one MAC costs the energy as much as for transmitting one byte data[21]. BS needs a bit more energy for verifying the report packets which have been delivered to it. Usually, BS has more energy and computation ability than normal nodes, and that overhead is ignorable.

When needed, nodes check their respective neighbor nodes. For defending against attacks such as sybil and hello flood attack, we use handshake method, and it needs a bit more communication overhead. However, considering the damage from the attacks, it is worth consuming the energy.

6. Conclusion

Sensor network is very vulnerable to various attacks because of its basic constraints. So for reliable communication, secure data transmission mechanisms are essential. In this work, we cluster the sensor network field as hexagons before node deployment, and in every cluster, the CH aggregates sensed data and generates a report packet with several MAC values. At each CH, the MAC values are verified and regenerated for the next step. The report packets with these MAC values are delivered through multiple routes all the way to the base station. To prevent the sybil or Hello flooding attacks, a handshake method is used for identifying neighbors. Simulation result shows that our proposal not only provides good performance with low overhead but also defends against various routing attacks which are possible in sensor network.

References


도인실

e-mail: iodoh@ewha.ac.kr

1993년 2월 이화여자대학교 전자계산학과

1995년 8월 이화여자대학교 전자계산학과

2000년 3월 ~ 2002년 9월 삼성 SDS

배치정보: 네트워크 보안, 인터넷 네트워크, 접속 네트워크 보안, 유비쿼터스 컴퓨팅

체기준

e-mail: kicha@ewha.ac.kr

1982년 2월 연세대학교 수학과 학사

1984년 2월 미국 Syracuse University 컴퓨터공학과 석사

1990년 2월 미국 North Carolina State University 컴퓨터공학과 박사

1990년 9월 ~ 1992년 9월 한국 대학연구원 컴퓨터공학과 조교수

1992년 9월 ~ 현재 이화여자대학교 컴퓨터공학과 교수

배치정보: 네트워크 보안, 인터넷/무선망/고속통신망

포토볼 설계 및 성능분석, 센서네트워크, 홈 네트워크, 유비쿼터스 컴퓨팅