The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

Performance Evaluation of Immunity-based Statistical En-route Filtering in Wireless Sensor Networks

Yuji Watanabe

Graduate School of Natural Sciences, Nagoya City University,
1 Yamanohata, Mizuho-cho, Mizuho-ku, Nagoya 467-8501 Japan
(Tel : 81-52-872-5037; Fax : 81-52-872-5037)
(yuji@nsc.nagoya-cu.ac.jp)

Abstract: Statistical en-route filtering (SEF) schemes can detect and eliminate false data injection attack in wireless sensor networks. However, SEF do not address the identification of compromised nodes injecting false reports. In this paper, we propose an immunity-based SEF to identify compromised nodes and achieve the earlier detection of false reports. In the proposed scheme, each node has a list of neighborhood and assigns credibility to each neighbor node. Each node can update the credibility of neighbor node based on success or failure of filtering and communication, and then use the updated credibility as the probability of next communication. Some simulation results show that the immunity-based SEF outperforms the original SEF.

Keywords: wireless sensor networks, statistical en-route filtering, immunity-based approach**.**

I. INTRODUCTION

In the last decade, wireless sensor networks have paid much attention because of the popularization of sensor nodes that are smaller, cheaper, and more intelligent [1]. In large-sized wireless sensor networks including a lot of sensor nodes, a detected event report can be sent to base station (user) using multi-hop communication where intermediate nodes forward the report. Wireless sensor networks may also be deployed in potentially hostile environment, so that the issue of security must be addressed. Attackers can compromise sensor nodes to inject false data reports of non-existing or bogus events using the compromised nodes. Such an attack is called *false data injection attack* [2]. The attack may cause not only false alarms but also the depletion of the limited energy of the nodes forwarding these reports to the base station.

Several research efforts [2-7] have proposed schemes to overcome such attack. The *statistical en-route filtering* (SEF) scheme [3] can probabilistically filter out false reports en-route in the dense deployment of large sensor networks. In SEF, assuming that the same event can be detected by multiple nodes, forwarding nodes along the way to base station can statistically detect false reports. SEF has achieved the early detection of false data reports with low computation and communication overhead. There are several revised en-route filtering schemes, for example, the dynamic en-route filtering [4], the multipath en-route filtering [5], the ticket-based en-route filtering [6], and LEDS [7]. However, these schemes do not address the identification of compromised nodes injecting false reports. If the compromised nodes are successfully identified, then neighbor nodes of the compromised nodes can drop false reports at an earlier stage.

For the detection of fault nodes on networks, an *immunity-based diagnostic model* [8] has been proposed inspired by the *Jerne's idiotypic network hypothesis* [9]. In the diagnosis, each node has the capability of testing the neighbor nodes, and being tested by the adjacent others as well. Based on the test outcomes, each node calculates its credibility. However, compromised nodes can not only output bogus test outcomes but also calculate the credibility at random.

In this paper, we propose an immunity-based SEF to identify compromised nodes. In the proposed scheme, each node has a list of neighborhood and assigns credibility to each neighbor node. Each node can not only update the credibility of neighbor node based on success or failure of filtering and communication but also use the updated credibility as the probability of next communication. We carry out some simulations to evaluate the performance of the proposed scheme. Some results show that the immunity-based SEF outperforms the original SEF.

II. SENSOR NETWORK MODEL

Following the previous studies on SEF, we also consider a large sensor network composed of a lot of sensor nodes and a base station which is a data collection center. We further assume that the sensor

nodes are deployed in high density, so that an event (sensing target) can be detected by multiple surrounding nodes. Because it is useless for each of the detecting nodes to send the event report (e.g., the location, the time, and the type of event) to the base station, one of them is elected as the cluster head. The cluster head collects and summarizes all the received event reports, and forward a synthesized report toward the base station. The report potentially traverses a large number of hops.

We assume that the attacker can compromise a node to obtain the security information installed in the node. Once compromised, the node can be used to inject false data reports of bogus events. However, we consider the attacker cannot defeat the base station because the base station has powerful security. Furthermore, this paper does not focus on various other attacks, for instance, false negative attacks and Dos attacks, by the compromised node.

III. STATISTICAL EN-ROUTE FILTERING (SEF) [3]

SEF can probabilistically filter out false reports en-route. SEF exploits collective decision-making by multiple detecting nodes and collective false detection by multiple forwarding nodes. SEF consists of three major components: (1) key assignment and report generation, (2) en-route filtering, and (3) base station verification.

1. Key assignment and report generation

The process of key assignment and report generation is as follows:

1) The base station (BS) maintains a global key pool of N secret keys $\{K_i, 0 \le i \le N-1\}$, divided into n non-overlapping partitions. Each partition has m keys. In other words, $N = m\,n$.
2) Before each sensor node is deployed, it stores randomly chosen k ($k < m$) keys from a randomly selected partition in the key pool.
3) When an event appears, multiple surrounding nodes can detect the event. A cluster head (CH) is elected from the detecting nodes to generate the event report.
4) Each of the nodes that detected the event generates a keyed message authentication code (MAC) M_i using the event report E and randomly selected K_i, one of its k stored keys. Each detecting node then sends $\{i, M_i\}$, the key index and the MAC, to the CH. K_i is secret while M_i is public.
5) The CH collects all the $\{i, M_i\}$s from the detecting nodes and randomly chooses T MACs from distinct partitions. This set of multiple MACs acts as the proof that the report is legitimate. Then the CH sends the final report attached T key indices and T MACs like $\{E, i_1, M_{i1}, i_2, M_{i2}, \ldots, i_T, M_{iT}\}$ toward the BS.

Fig.1 illustrates the example of the key assignment and report generation in SEF. In this figure, the BS has a global key pool of N = 12 keys divided into n = 4 partitions, each of which has m = 3 keys. Each sensor node randomly picks k = 2 secret keys from one partition of the key pool. After each detecting node endorses the event report by producing a keyed MAC using one of its stored 2 keys, the CH collects all the MACs from the detecting nodes and attaches randomly selected T = 3 MACs, that is, M_2, M_9 and M_{10} to the event report E.

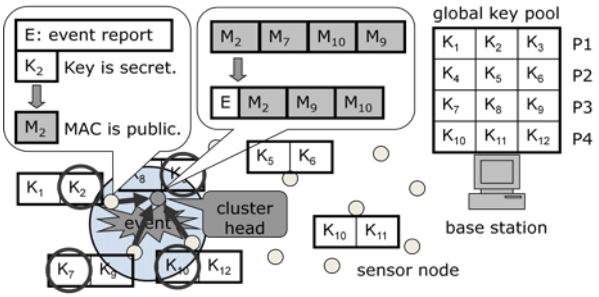

Fig.1. Example of the key assignment and report generation in SEF with 12 keys, 4 partitions, 3 keys in each partition, 2 keys in each node, and 3 MACs attached to event report.

2. En-route filtering

In en-route filtering process, intermediate forwarding nodes verify the correctness of the MACs probabilistically and drop a report with forged MACs en-route. The en-route filtering process is as follows:

1) Since a legitimate report carries exactly T MACs produced by T keys of distinct partitions, a report with less than T MACs or more than one MACs in the same partition is dropped.
2) Because of the randomized key assignment, each forwarding node has certain probability to possess one of the keys that are used to produce the T MACs. If forwarding node finds out that it has one of the T keys in the report, it reproduces the MAC using its stored key and compares the result with the corresponding MAC attached in the report. If

the attached MAC is different from the reproduced one, the report is dropped.

3) When intermediate node does not have any of the T keys, the node forwards the report to the next hop.

The key assignment ensures that each node can produce only *partial* proof for a report. A single compromised node has to forge MACs to assemble a seemingly complete proof in order to forward false reports. In Fig.2, since a malicious node has 2 keys from only partition 1, it needs to forge the other 2 MACs, M_9 and M_{10}. The report with the forged MACs is dropped because the correctness of the MACs can be verified by the intermediate node with K_{10}.

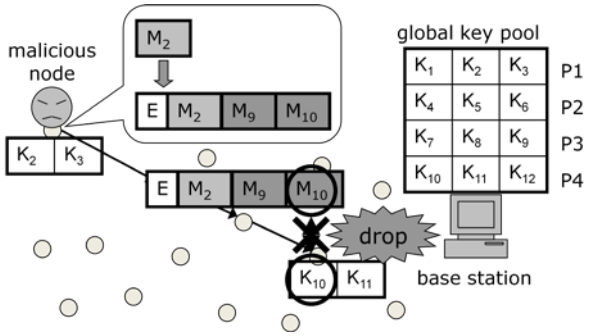

Fig.2. Case that a false report with forged MACs from a malicious node is dropped by the intermediate forwarding node.

3. Base station verification

Due to the statistical nature of the detection mechanism, a few bogus reports with invalid MACs may escape en-route filtering and reach the BS. In base station verification process, the BS further verifies the correctness of each MAC and eliminates false reports that elude en-route filtering.

IV. PROPOSED IMMUNITY-BASED SEF

The original and revised SEFs do not deal with the identification of compromised nodes injecting false reports. Simple trace back is futile if the compromised nodes tell a lie that the false reports are received from the other nodes. To detect fault nodes in networks, the *immunity-based diagnostic model* [8] is a promising approach. In the diagnosis, each node has the capability of testing the neighbor nodes, and being tested by the adjacent others as well. Based on the test outcomes, each node calculates its own credibility. However, malicious nodes can not only output bogus test outcomes but also calculate the credibility at random.

Therefore, we propose an immunity-based SEF scheme to identify compromised nodes. In the proposed scheme, each node has a list of neighborhood and assigns a state variable $R \in [0,1]$ indicating *credibility of neighbor* to each neighbor node. Note that each node does not have its own credibility. Node j updates the credibility R_{ji} of the previous neighbor node i sending the event report based on its filtering result and the reply from next neighbor node k as follows:

$$R_{ji}(t+1) = \begin{cases} R_{ji}(t) + \Delta_s & \text{if node } j \text{ receives the reply from next node } k \\ R_{ji}(t) - \Delta_f & \text{if node } j \text{ does not receives the reply from next node } k \\ R_{ji}(t) - \Delta_d & \text{if node } j \text{ drops the report using SEF} \end{cases} \quad (1)$$

The initial value of credibility $R_{ji}(0)$ is 1. If credibility $R_{ji}(t)$ is over 1 (under 0), it is set to 1 (0). The values of the parameters Δ_s, Δ_f and Δ_d should be chosen through mathematical analysis and simulation.

For example, in Fig.3, node i increases the credibility R_{ih} of the previous node h because the reply from next node j can be received. However, node j decreases the credibility R_{ji} of the previous node i because next node k drops the event report using SEF and does not reply to node j. Since node k filters out the report by itself, the credibility R_{kj} of the previous node j also decreases.

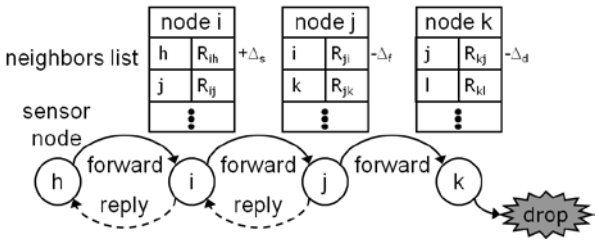

Fig.3. An immunity-based SEF scheme for identifying compromised nodes.

Only the credibility update process cannot achieve the identification of compromised nodes. For instance, in Fig.3, if node h is compromised, false reports are still forwarded toward node k. Therefore each node uses the updated credibility as the probability of next communication. In the same example, node i has

adversely higher probability of receiving the report from compromised node h because of the increase of the credibility R_{ih}. However, since node j has lower probability of receiving the report from node i, node i may fail to communicate with node j at next stage, and then the credibility R_{ih} of the previous node h in the neighbors list of node i decreases. Although the credibility R_{kj} of the previous node j in the neighbors list of node k decreases at first, if node j sends legitimate reports received from the other previous nodes to node k, the credibility R_{kj} can be recovered. By iterating the credibility update and the communication based on the updated credibility, our scheme will be expected to inhibit neighbor nodes of compromised nodes from forwarding false reports.

V. SIMULATION RESULTS

We carry out some simulations to evaluate the performance of the proposed scheme. Simulation conditions are the same as [3]: 340 nodes are uniformly distributed in a field which size is 200 x 20 m^2. One base station and one event source sit in opposite ends of the field, with about 100 hops in between. The BS has a global key pool of 1000 keys divided into 10 partitions, each of which has 100 keys. Each node has 50 keys, and 5 MACs are attached to event report. The results are averaged over 10 network topologies.

Fig.4 shows the percentage of dropped false reports as a function of the number of forwarding nodes for immunity-based approach ($\Delta_s = \Delta_f = \Delta_d = 0.02$) and SEF, respectively in case that one node is compromised and 1000 bogus reports are sent by the compromised node. Results show that as false reports are forwarded, more and more reports are dropped: about 100% bogus reports are detected within 20 forwarding nodes for both methods. Furthermore, about 65% false reports are dropped by the original SEF within 5 intermediate nodes, while about 73% reports are filtered out by the immunity-based SEF. We confirm that the immunity-based SEF can achieve the earlier detection of false reports than the original SEF.

VI. CONCLUSION

In this paper, we proposed an immunity-based SEF scheme for identifying compromised nodes in wireless sensor networks. Some results show that the immunity-based SEF outperforms the original SEF. In future, the proposed scheme will be additionally combined with other security mechanisms for higher security level.

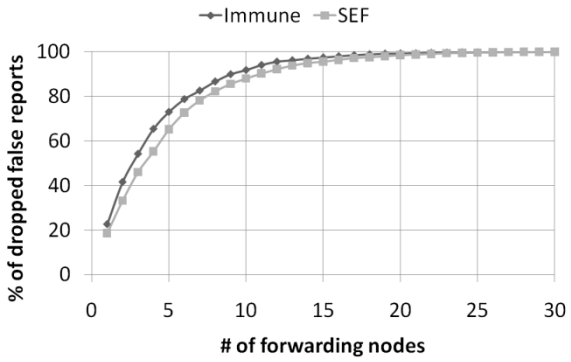

Fig.4. Percentage of dropped false reports as a function of the number of forwarding nodes for immunity-based approach and SEF, respectively.

ACKNOWLEDGEMENTS

This work was partly supported by a Grant-in-Aid for Scientific Research (C) (22500063) from Japan Society for the Promotion of Science.

REFERENCES

[1] Yick J, Mukherjee B, Ghosal D (2008), Wireless Sensor Network Survey. Computer Networks 52: 2292-2330

[2] Zhu S, Setia S, Jajodia S, Ning P (2004), An Interleaved Hop-by-Hop Authentication Scheme for Filtering of Injected False Data in Sensor Networks. IEEE Symposium on Security and Privacy, 259-271

[3] Ye F, Luo H, Lu S, Zhang L (2005), Statistical En-Route Filtering of Injected False Data in Sensor Networks. IEEE Journal on Selected Areas in Communications 23(4):839-850

[4] Yu Z, Guan Y (2006), A Dynamic En-route Scheme for Filtering False Data Injection in Wireless Sensor Networks. Proceedings of the 25th IEEE Conference on Computer Communications, 1-12

[5] Kim MS, Cho TH (2007), A Multipath En-Route Filtering Method for Dropping Reports in Sensor Networks. IEICE Transactions on Information and Systems E90-D(12):2108-2109

[6] Krauß C, Schneider M, Bayarou K, Eckert C (2007), STEF: A Secure Ticket-Based En-route Filtering Scheme for Wireless Sensor Networks. 2nd IEEE International Conference on Availability, Reliability and Security, 310-317

[7] Ren K, Lou W, Zhang Y (2008), LEDS: Providing Location-aware End-to-end Data Security in Wireless Sensor Networks. IEEE Transactions on Mobile Computing 7(5):585-598

[8] Ishida Y (1990), Fully Distributed Diagnosis by PDP Learning Algorithm: Towards Immune Network PDP Model. Proceedings of International Joint Conference on Neural Networks, 777-782

[9] Jerne N (1973), The Immune System. Scientific American 229(1):52-60