Adaptive Threshold Determining Method for Statistical Filtering Scheme in Sensor Networks

Chung Il Sun and Tae Ho Cho,

Sungkyunkwan University, Suwon 440-740, South Korea

Summary
Many sensor network applications are dependent on the secure operation of networks, and will have serious outcome if the network is disrupted or injured. Bogus reports can be injected through the compromised nodes, which can lead to not only false alarms but also the depletion of limited energy resource. Ye et al. proposed a statistical en-route filtering scheme (SEF) to detect such bogus reports during the forwarding process. In SEF, it is important that determining a number of the message authentication code (MAC) which will be attached in event report when a real event occurs since it trades off detection power and overhead. This paper presents an adaptive threshold determining method for energy saving based on distance. Sensor field is divided into several areas by considering the distance from the base station. Each area uses the different number of the MACs so that the size of report can be reduced by the distance between the base station and the region where event occurs in hop count. Thus, we can reduce the energy consumption in transmitting.

Key words:
Sensor networks, false data filtering, secure method, statistical en-route filtering

1. Introduction

Recent advances in micro-electro-mechanical systems, electronics and wireless communications have made it practical now to develop and deploy low-cost, high-performance and low-power sensors [1, 2]. Wireless sensor network offer unprecedented capabilities to monitor the physical world and enable a variety of applications such as military surveillance, and vehicle safety monitoring [3]. Sensor networks consist of a large number of sensor nodes that have limited processing power, small storage space, narrow bandwidth, limited energy, and a few base stations that collect the sensor readings. In sensor networks, sensor nodes within their wireless transmission ranges can communicate with each other directly, while sensors outside the range have to rely on some other sensors to relay the message [4]. Sensor nodes are deployed randomly in unattended environment that may be destroyed, compromised or dead as times go by. Hence sensor nodes are vulnerable to false data injection attacks in which adversaries inject fabricated reports into the networks through compromised nodes (Fig. 1). Fabricated reports deceive the base station or drain the limited energy resource of the networks [5].

Ye et al. proposed the statistical en-route filtering scheme (SEF) [6] to filter out fabricated reports during process. In SEF, multiple sensing nodes collaboratively generate a legitimate report and endorse it by attaching to it their message authentication codes (MACs). Each MAC is generated by a node using one of its stored symmetric keys and represents its agreement on the report [7]. A represented node collects MACs as many as a threshold value (T) and makes the final report. The threshold value is determined by the user before the node deployment. It decides the number of MACs in the final report. As the report is forwarded towards the base station over multiple hops, each forwarding node verifies the correctness of the MACs attached to the report with certain probability and drops the reports if an incorrect MAC is detected [6]. Due to a characteristic of SEF mechanism, a few reports with incorrect MACs cannot be dropped during forwarding process and reach the base station. However, the base station can verify the correctness of every MAC and refuse the fabricated reports because it has all the keys in the global key pool.

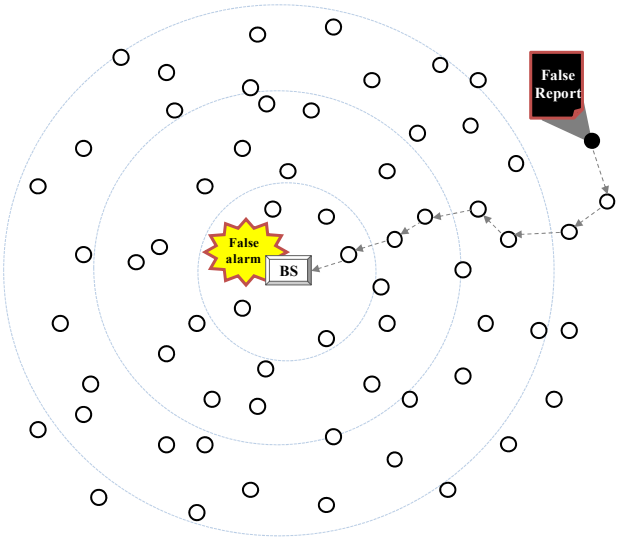

Fig. 1 false data injection attack

Manuscript received July 5, 2008.
Manuscript revised July 20, 2008.

In this paper, we propose an adaptive threshold determining method for energy saving based on distance in SEF. The network is divided into several areas by considering the network situation. Each area applies to the different threshold value that regulates the number of MACs on aggregation process. The report produced in an area near by the base station contains a small number of MACs according to the threshold value of occurrence area. Thus, the nodes can save the energy consumption for delivery.

The remainder of the paper is organized as follows: Section 2 gives a brief description of SEF. Section 3 explains the proposed method. Section 4 reviews the simulation results. Finally, the conclusion and future work are discussed in Section 5.

2. Background

2.1 Statistical En-route Filtering scheme

In SEF, the base station maintains a global key pool which is divided into n partitions. Each partition has some keys, and each key has a unique key index. Before a sensor node is deployed, the user randomly selects one of the multiple partitions, and randomly chooses a small number of keys form this partition to be loaded into every node [6]. Fig. 2 shows an example of a global key pool with n = 8 paritions, each of which has m keys. Node 1 has k keys randomly selected from one partition.

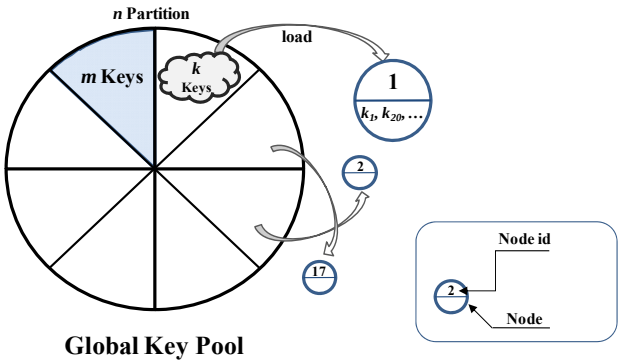

Fig. 2 a global key pool

When a real event occurs in the sensor field, one of detecting nodes is elected as the center-of-stimulus (CoS) node to generate a sensing report. Each detecting node produces a message authentication code (MAC) using one of its stored keys. The CoS collects T key indices of distinct partitions and T MACs. This set of multiple MACs acts as the proof that a report is legitimate. A report with less than T MACs or key indices, or more than one key index in the same partition, will not be forwarded. Then the CoS forwards the report toward the base station over multi hops. Each forwarding node verifies the correctness of the MACs attached in the report using its stored keys. If there is a same key index in the report, the node verifies the correctness of the MAC using its stored keys.

An adversary can inject a fabricated report with incorrect MACs through a compromised node as shown in Fig. 3(a). However, the fabricated report may be dropped since each forwarding node verifies the correctness of the MACs carried in the report with certain probability (Fig. 3(b)). The fabricated report is not filtered out by forwarding nodes then the base station serves as the final defense that catches the fabricated reports, because it has complete knowledge of the global key pool (Fig. 3(c)). SEF can detect fabricated reports by an adversary with a fixed number of compromised partitions.

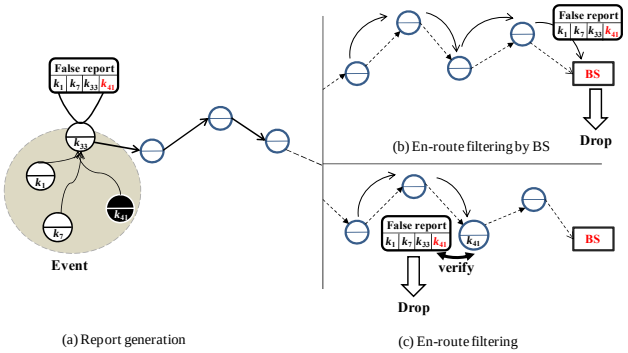

Fig. 3 a report generation and en-route filtering

3. Proposed Method

3.1 Assumption

We assume that the density of network is dense enough. We also assume that routing paths are established by flooding a control message. A control message is broadcasted by the base station upon a change of the network topology or a user's request. We further assume that the base station can know the hop count of every path in the network.

3.2 Adaptive Threshold Determining Method

In our proposal, after node deployment, routing paths are established by flooding a broadcast message. The base station can know information of the all paths if the broadcasting is finished. Then base station can find the

maximum hop count among the all paths. The maximum hop count is regarded as the size of the sensor field. The user determines the P which is security distance value by considering the size of the sensor field. The P is decided based on the maximum hop count and the threshold value by equation (1).

$$P = \frac{d_{\max}}{\frac{T}{2}+1} \quad (1)$$

Where $d_{\max}$ is the maximum hop count among the paths, and T is initially settled by the user. After decision of P, the base station divides the field into several areas using $d_{\max}/P$ (Fig. 4). The base station floods control messages to the all nodes in the network. The control messages include the information of the range of areas, P, and threshold value as shown in Table 1.

Area	***Range***	Threshold
$Area_1$	$0 < Area_1 \le P$	T_{Area2}-1
$Area_2$	$0 < Area_1 \le Area_1 + P$	T_{Area3} - 1
$Area_{n-1}$	$Area_{n-2} < Area_{n-1} \le Area_{n-2} + P$	T
$Area_n$	$Area_{n-1} < Area_n$	T

When a node receives a control message, it compares own hop count with the range of the all areas and finds the area in which it is located. Then the node modifies the new threshold value stated in the control messages in its memory. After flooding the control messages, every node belongs to one of classified areas. Farthest area from the base station applies the T which is determined by the user. The more area is closer to the base station, the more threshold value is decreased. However, the threshold value of the Area1 does not set the zero. The threshold value which equals zero means that the network does not consider filtering of the fabricated report.

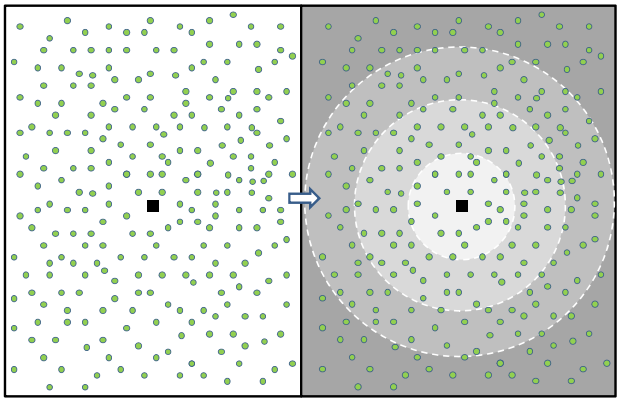

Fig. 4 Division of sensor network by the control message

Generally, every node has same threshold value that is fixed and is determined by the user. If the real event occurs nearby the base station then the fabricated report will be verified by the base station. Because a small hop count makes little chance of the report verification. In this case, the energy for forwarding report is more consumed than verifying for the report, because the large size of the report is forwarded toward the base station. To reduce the energy for the forwarding report, the size of the report that is generated at the contiguous region from the base station has to decrease. On the other hand, the fabricated report that is generated far from the base station will have large chance to verify the event report. The fabricated reports may be detected earlier before they consume a significant amount of energy.

Therefore, for energy saving, we have to determine the adaptive threshold value based on the distance between the base station and each sensor node in hop count.

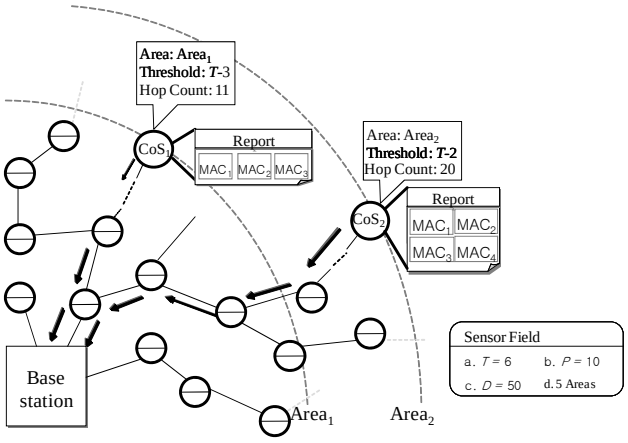

Fig. 5 MAC aggregation and forwarding the report

Fig. 5 shows that the CoS aggregates MACs and forwards the report to the base station. When the real event occurs in the $Area_1$, the CoS 1 collects 3 MACs from surrounding sensing nodes and makes a final report then forwards it toward the base station. Since the CoS 1 is located in the $Area_1$, it uses by the threshold value of the $Area_1$. The report generated by the CoS 1 is forwarded through a small multiple hops. Thus, if it is the bogus report then it cannot be detected by forwarding nodes easily. However, the size of report is smaller than other areas, can reduce the energy consumption in transmitting.

4. Simulation Results

To show the effectiveness of the proposed method, we have compared the adaptive threshold with fixed threshold value through the simulation. Each node takes 16.25 μ J to transmit/receive a byte and each MAC generation consumes 15 μ J [3]. The size of an original report is 24 bytes and the size of a MAC is 1byte. There are 1,000 keys in the global key pool which is divided into 10 partitions. An initial threshold value is 8 and a minimum threshold value is 4 by equation (1). Every node has different threshold value in the different area to generate the report using equation (1).

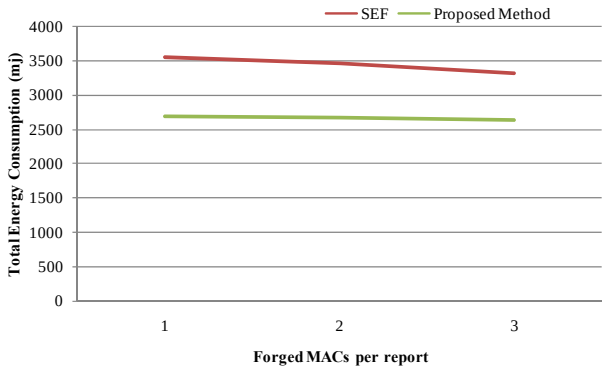

Fig. 6 Total energy consumptions per forged MACs

In Fig. 6 shows that average energy consumption caused by a fabricated report when $T = 8$, $d_{max} = 50$, and $P = 12$. The number of the fabricated MACs aggregated in the report is more than $T/2$. The T value of the $Area_1$ is 3. As show in the figure, the proposed method consumes no more energy than the fixed threshold values.

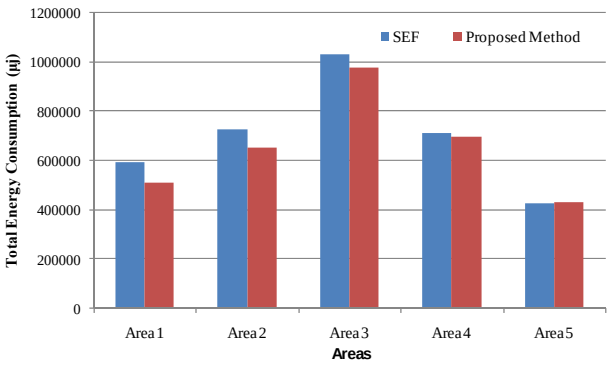

Fig.7 Total energy consumption of each area

Fig. 7 shows the total energy consumption of each area when the sensor field is divided into 5 areas. The proposed method is more efficient than the SEF. The $Area_1$ and $Area_2$ that close to the base station consume no more energy than the original SEF. That is, the proposed method can conserve energy better than the SEF.

6. Conclusion

We present an adaptive threshold determining method for energy saving. In our proposed method, the node in each different area uses the threshold value of such area considering by distance from the base station. The distance based adaptive threshold determining method can conserve energy, while it provides sufficient detection power. The effectiveness of the proposed method was shown with the simulation result. The result shows that the proposed method can lower the energy consumption than the existing method.

Acknowledgments

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement). (IITA-2008-C1090-0701-0028)

References

[1] Guorui Li, Jingsha He, and Yingfang Fu, “Analysis of an Adaptive Key Selection Scheme in Wireless Sensor Networks”, LNCS 4490, pp. 409-416, 2007.

[2] K. Akkaya and M. Younis, “A Survey on Routing protocols for Wireless Sensor Networks”, Ad hoc Netw., vol. 3, no. 3, pp. 325-349, May 2005.

[3] Yang and S. Lu, “Commutative Cipher based En-Route Filtering in Wireless Sensor Networks”, Proc. of VTC, pp. 1223-1227, Sep. 2003.
[4] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, “Directed Diffusion for Wireless Sensor Networking”, IEEE/ACM Transactions on Networking, vol. 11, no. 1, Feb. 2003.
[5] W. Zhang and G. Cao, “Group Rekeying for Filtering False Data in Sensor Network: A Predistribution and Local Collaboration-base Approach”, Proc. of INFOCOM., pp. 503-514, Mar. 2005.
[6] F. Ye, H. Luo, and S. Lu, “Statistical En-Route Filtering of Injected False Data in Sensor Networks”, *IEEE J. Sel. Area Comm.*, vol. 23, no. 4, pp. 839-850, Apr. 2005.
[7] H. Y. Lee, and T. H. Cho, “Fuzzy based Security Threshold Determining for the Statistical En-Route Filtering in Sensor Networks, Enformatika 14, pp. 157-160, Aug. 2006

Chung Il Sun received his B.S. degrees in computer engineering from Kyungwon University, Korea, in February 2007. He is currently a graduate student in the School of Information and Communication Engineering at Sungkyunkwan University. His research interests include wireless sensor networks, modeling & simulation and security in wireless sensor networks.

Tae Ho Cho received the Ph.D. degree in electrical and computer engineering from the University of Arizona, USA, in 1993, and the B.S. and M.S. degrees in electrical engineering from Sungkyunkwan University, Korea, and the University of Alabama, USA, respectively. He is currently a Professor in the School of Information and Communication Engineering, Sungkyunkwan University, Korea. His research interests are in the areas of wireless sensor network, intelligent system, modeling and simulation, enterprise resource planning.