Fuzzy based Adaptive Threshold Determining Method 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 serious problems may result if the network is disrupted or injured. An adversary using these restrictions can compromise sensors and use them to inject bogus reports into the network. Bogus reports can lead to both false alarms and the depletion of limited resources. A statistical en-route filtering scheme (SEF) was proposed to detect such bogus reports during the forwarding process. In SEF, it is important to determine the number of the message authentication code (MAC) to attach to the event report when a real event occurs. The number of MACs which attached to the report trades off detection power and overhead. We propose a threshold determining method to save energy using a fuzzy logic. Fuzzy rule-based systems are used to divide the sensor field into several areas considering the network situation. Each area uses a different number of MACs so that the size of the report can be reduced by the distance between the base station and the region where the event occurs in the hop count. Simulation demonstrates that the proposed method is resilience and energy efficient.

Key words:
Sensor networks, false data filtering, fuzzy logic, statistical en-route filtering

1. Introduction

Recent advances in wireless communications and electronics have enabled the development of low-cost, high-performance and low-power sensors [1, 2]. Wireless sensor networks 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 monitor the environment and a few sink nodes that collect the sensor reading. Sensors have limited processing power, minimal storage space, narrow bandwidth, limited energy. In typical applications for sensor networks, sensor nodes within their wireless transmission ranges can communicate with each other directly, while sensors outside the range rely on other sensors to relay the message [4, 9]. Sensors are deployed randomly in an 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 bogus reports into the networks through compromised nodes. Bogus reports can deceive the base station or drain the finite amount of energy resource in a battery powered networks (Fig. 1) [5]. To minimize damage, bogus reports should be dropped as early as possible, and the few eluded ones should be further rejected at sink nodes [6]. This leads to significant savings of energy [5].

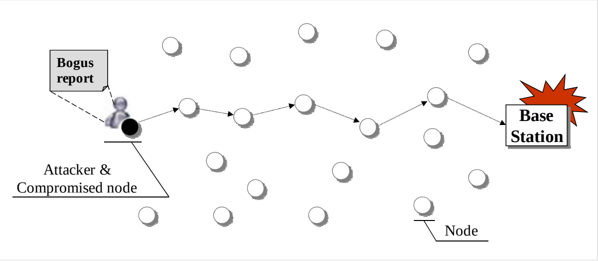


Fig. 1 False data injection attack

Ye et al. proposed a statistical en-route filtering scheme (SEF) [6] to filter out bogus reports during processing. SEF can detect bogus reports probabilistically. In SEF, multiple sensing nodes collaboratively generate a legitimate report and endorse it by attaching 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 up to a threshold value (*T*) and creates the event report. The threshold value is determined by the user before the node deployment, and the user decides the number of MACs in the event report. As the report is forwarded to the base station over multiple hops, each forwarding node verifies the correctness of the MACs attached to the report within a certain probability and drops the report if an incorrect MAC is detected [6]. Due to a characteristic of the SEF mechanism, a few reports with incorrect MACs cannot be dropped during the forwarding process, so they reach the base station. However, the base station can verify the correctness of every MAC and refuse the bogus reports because it has all the keys in the global key pool.

Manuscript received August 5, 2009
Manuscript revised August 20, 2009

In this paper, we propose an adaptive threshold determining method to save energy based on distance with SEF. We use a fuzzy system to divide the sensor network into several areas depending on the network situation. Fuzzy system is used to determine the number of MACs attached to the event reports which are created in different area. Each area applies a different threshold value that regulates the number of MACs in the aggregation process. The report produced in an area near the base station contains a small number of MACs based on the threshold value of the occurrence area. Thus, the nodes require less energy during transmission.

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 works are discussed in Section 5.

2. Background

2.1 Statistical En-route Filtering scheme

SEF [6] is the first paper that addresses false data injection attacks in the presence of compromised nodes [6]. SEF can detect bogus reports probabilistically so that it does not guarantee that a bogus report can be always detected and dropped in the forwarding process. The base station maintains a global key pool which is divided into multiple partitions. Each partition has multiple 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 from this partition to be loaded into every node [6]. SEF assumes that the same event can be detected by multiple sensor nodes. When an event occurs in the sensor field, one of the detecting nodes is elected as the center-of-stimulus (CoS) node which collects MACs for the event from its neighboring nodes and generates an event report. Each detecting node produces MAC using one of its stored keys. To create the event report, the CoS collects *T* key indices with distinct partitions and *T* MACs. This set of multiple MACs is used to prove that a report is legitimate [8]. A report with less than *T* MACs or key indices or more than one key index in the same partition will not be forwarded. The number of MACs; *T* is decided by the user and it can be changed as the network situation change. The CoS forwards the event report to the base station with multiple hops. Each forwarding node verifies the MACs attached in the report using its stored keys. If there is the same key index in the report, the node verifies the correctness of the MAC using its stored keys. An adversary can launch the false positive attack using compromised detecting node against the collaborative report generation process to intercept the reporting of illegal events. An adversary can inject a bogus report with incorrect MACs through a compromised node. However, the bogus report may be dropped since each forwarding node verifies the correctness of the MACs carried in the report within a certain probability. If the bogus report is not filtered out by forwarding nodes, the base station serves as the final defense to catch the bogus reports because it has complete knowledge of the global key pool. SEF can detect bogus reports by an adversary with a fixed number of compromised partitions.

2.2 Motivation

In SEF, the conservation of energy and the detection power of false reports are largely affected by the determination of the threshold value. Since threshold value which is decided by the user is fixed and unchanged, it can cause the consumption of energy unnecessarily when the event occur s near the base station and the event report is legitimate. The threshold value is important factor that decides the number of MACs to make event reports. The large *T* has many chances to detect incorrect MACs. However it can increase the size of event reports so that sensor nodes will consume the much energy when they transmit/receive the event report. Therefore it is important to determine the threshold value suited to the network situation. We propose the threshold determining method adaptively by a fuzzy rule-based system.

3. Fuzzy Adaptive Determining Method

3.1 Assumption

We assume that the network density is sufficient. We also assume that routing paths are established by flooding with a control message. A control message is broadcast by the base station after a change in the network topology or following a user’s request. This fashion is commonly used in most routing protocols at the initial establishment of the paths [8]. We further assume that the base station can know the hop count of every path in the network. The base station cannot be compromised and has a mechanism to authenticate a broadcast message.

3.2 Initial Area Partitioning phase

In our proposal, after node deployment, routing paths are established by flooding a broadcast message. The base station has complete knowledge of information for all paths if the broadcasting is complete. Then the base station can find the maximum hop count among all paths, which indicates the size of the sensor field. The user determines *P*, which is the security distance value, by considering the size of the sensor field. *P* is determined from Equation (1) based on the maximum hop count and the initial threshold value determined by user.

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

Where $d_{\max}$ is the maximum hop count among the paths and *T* is initially set by the user. After determining *P*, the base station divides the field into several areas using $d_{\max}/P$. The base station floods control messages to all the nodes in the network. The control messages include information on the range of areas, *P*, and initial threshold values as shown in Table 1.

Table 1: the initial threshold values according to the area

Area	***Range***	Threshold
Area_1	$1 < \text{Area}_1 \le P$	$T/2$
Area_2	$P < \text{Area}_2 \le n \cdot P$	$T/2+1$
…	…	…
Area_{n-1}	$(n-2) \cdot P < \text{Area}_{n-1} \le n \cdot P$	T
Area_n	$n \cdot P < \text{Area}_n$	T

When a node receives a control message, it compares its own hop count with the range for all the 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 the classified areas. The farthest area from the base station applies *T*, which is determined by the user. As the area close to the base station increases, the threshold value decreases. However, the threshold value of Area 1 cannot reach zero. The threshold value, which equals zero, means that the network does not consider filtering of the bogus report.

Generally, every node has the same fixed threshold value which is determined by the user. If the real event occurs near the base station, then the bogus report will be verified by the base station. A small hop count creates only a small chance for report verification. In this case, the energy required to transmit the report is greater than for the energy required to verify the report, because the large report is transmitted to the base station. To reduce the energy required to transmit the report, the size of the report that is generated in the region contiguous to the base station should be decreased. The bogus report that is generated far from the base station has a chance to be verified at many stations. The bogus reports may be detected earlier, before they consume a significant amount of energy.

3.2 Factor that determine a

After the initial area partitioning phase, the user should consider the threshold value as the network situation changes. As times go by, sensor nodes deployed in network will drain the finite energy. To save the energy, the user determines the threshold value that is suitable in each area adaptively. If sensor nodes deployed in a certain area consume the much energy then user reset the smaller threshold value than initial value. To determine an adaptive threshold value, α, we use the fuzzy rule-based system. The fuzzy system determines the adaptive threshold value by considering an energy level of each area, the distance from the base station in hop count, and the ratio of false traffic.

The fuzzy system uses three input parameters; 1) energy level of each area, 2) distance, and 3) false traffic ratio

1) Energy level of each area: The energy is the most important resource that should be considered in sensor networks. Generally, sensor nodes are limited in power and irreplaceable since these nodes have limited capacity and are unattended [3]. The energy of nodes in each area should be conserved to maintain sensing regions for a long time. If the energy level of a certain area is high, the existing threshold value will be held or increased. Therefore, we have to determine α based on the energy level of each area.

2) Distance: The distance from the base station in the hop count is also considered to determine α. If the location of the CoS is far from the base station, the chance which the forwarding nodes can verify the reports will increase. However, a short distance has a small chance to verify the event report so that the threshold value in the short distance area has to be a small. On the other hand, a long distance will have a big chance to verify and drop the bogus reports probably. Therefore, we have to determine α based on the distance.

3) Ratio of false traffic: when the bogus reports arrive at the base station, the base station can know whether the received report is false or not. Also the base station can know that where it comes from because the report contains the information of an event location. If the false traffic ratio in a certain area is high, the existing threshold value does not affect to filter out the bogus reports so that the user increases the threshold value under the area. Therefore, for energy saving, we have to determine α based on the ratio of false traffic in the network.

3.2 Fuzzy Logic Design

The membership functions of three input parameters – the ratio of false traffic (FTR), the distance from the base station in hop count(DIS), and the energy level of each area(EL) – of the fuzzy logic are shown in Fig. 2(a), (b), and (c). The labels in the fuzzy variables are presented as follows.

- FTR = {VS (Very Small), S, M, L, VL (Very Large)}
- DIS = {N (Near), M, F (Far)}
- EL = {VL (Very Low), L, E (Enough)}

The output parameter of the fuzzy logic is ALPHA = {S(Steady), INC(Increase), DEC(Decrease)}, which is represented by the membership functions as shown in Fig. 2(d).

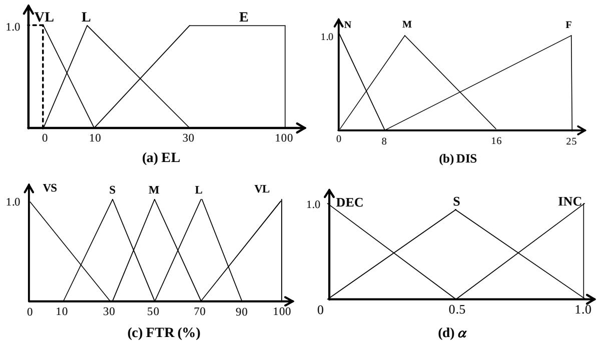


Fig. 2 Fuzzy membership functions

The fuzzy logic determines one of the following three actions using the above three input parameters.

- Steady (S): Existing threshold value does not need to change.
- Increase (INC): For detection power, each area should be updated by the threshold value which is increased
- Decrease (DEC): For energy saving, the user should make low the new threshold value.

The BS periodically or on demand determines whether the threshold value is needed to be reset using the fuzzy logic. If the output of the fuzzy logic is INC, the BS increases the existing threshold value. The BS broadcast the control message attached new threshold value to the sensor nodes then every node finds the area in which it is located and modifies the new threshold value stated in the control messages in its memory until the fuzzy logic outputs INC or DEC.

4. Simulation Results

To show the effectiveness of the proposed method, we have compared the fuzzy based adaptive threshold with a fixed threshold value through the simulation. The size of sensor filed is 575x575m^2, where 2,000 nodes are randomly distributed. Each node requires 16.25 μJ to transmit/receive a byte, and each MAC generation consumes 15 μJ [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. Each node loads 3 keys. The initial threshold value is 8 and the minimum threshold value is 4, as determined in Equation (1). Every node has a different initial threshold value in a different area to generate the report using Equation (1).

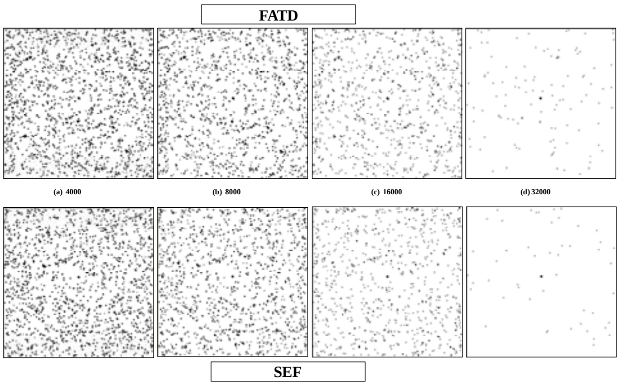
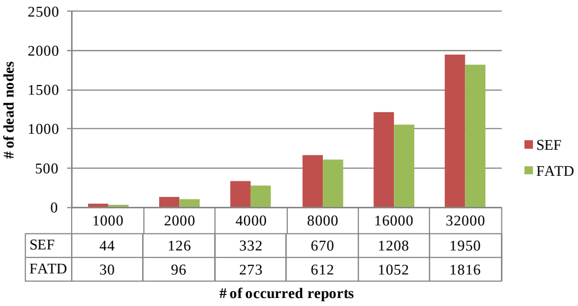


Fig. 3 State of sensor network

Fig. 3 shows the state of sensor network when the number of occurred event reports are 4000, 8000, 16000, and 32000. As show in the figure, the sensor network where applies fuzzy system to determine the adaptive threshold (FATD) can prolong the life of nodes.



	1000	2000	4000	8000	16000	32000
SEF	44	126	332	670	1208	1950
FATD	30	96	273	612	1052	1816

Fig. 4 The number of dead nodes in filed

Fig. 4 shows the number of dead nodes in sensor filed. As show in the figure, FATD is more efficient that the

SEF. According to the incidence of the report generation, the number of dead node in SEF is greater than FATD. That is, FATD can conserve the energy of nodes better than SEF.

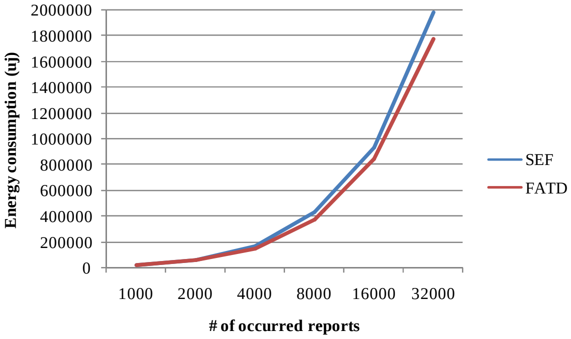


Fig. 5 Total energy consumption

Fig. 5 shows the total energy consumption when the number of occurred event reports is 1000, 2000, 4000, 8000, 16000 and 32000. In this, determines the number of MACs with the fuzzy logic consumes less energy than the fixed MACs in SEF.

6. Conclusion

We present fuzzy based adaptive threshold determining method to save energy of sensor network. In our proposed method, the node in each area creates an initial threshold value for the area depending on the distance between the area and the base station. The fuzzy based adaptive threshold determining method can conserve energy and provide sufficient detection power. We conducted a simulation to demonstrate the effectiveness of the proposed method. The results show that the proposed method requires less energy than the existing method.

Acknowledgments

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2008-313-D00827)

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

[8] C. I. Sun, and T. H. Cho, “A Path Selection Method for Improving the Detection Power of Statistical Filtering in Sensor Networks”, Journal of Information Science and Engineering, Vol. 25, No. 4, pp. 1163-1175. 2009

[9] H. Y. Lee, and T. H. Cho, “Fuzzy Security Parameter Determining Method for the Commutative Cipher Based Filtering in Sensor Networks”, Lecture Notes in Computer Science, Springer Verlag, LNCS 4706, pp. 573-583, Aug. 2007.

Chung Il Sun received his B.S. degree in computer engineering from Kyungwon University, Korea, in February 2007, and M.S degree in electrical and computer engineering from Sungkyunkwan University, Korea, in 2009. 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 and simulation and security in wireless sensor networks.

Tae Ho Cho received his Ph.D. degree in electrical and computer engineering from the University of Arizona, USA, in 1993, and 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 networks, intelligent systems, modeling and simulation and enterprise resource planning.