

A Path Selection Method for Improving the Detection Power of Statistical Filtering in Sensor Networks*

CHUNG IL SUN, HAE YOUNG LEE AND TAE HO CHO

School of Information and Communication Engineering

Sungkyunkwan University

Suwon 440-774, Korea

In many sensor network applications, sensor nodes are deployed in unattended environments, and hence are vulnerable to physical attacks, potentially compromising the node's cryptographic keys. False sensing reports can be injected through compromised nodes, which can lead to not only false alarms but also the depletion of limited energy resources in battery powered networks. Ye *et al.* [4] proposed the statistical en-route filtering scheme to detect and drop such false reports during the forwarding process. Since each node in this scheme has a limited amount of information for verification, the detection power is largely affected by the choice of routing paths. In this paper, a path selection method is proposed, to improve the detection power of the statistical filtering. Each node evaluates the detection power of each incoming path from the base station and chooses the most secure path for data delivery against false data injection attacks. The effectiveness of the proposed method against false data injection attacks is shown via simulation results.

Keywords: sensor networks, false data injection attack, statistical filtering, secure routing path selection, security

1. INTRODUCTION

Recent advances in MEMS (micro-electro-mechanical-systems) and low power highly integrated digital electronics have enabled the development of low-cost sensor networks [1, 2] consisting of small nodes with sensing, computation and wireless communications capabilities [3]. Sensor networks are expected to interact with the physical world at an unprecedented level of universality, and enable a variety of new applications [4, 5]. In many applications, the sensor nodes are deployed in open environments, and hence are vulnerable to physical attacks, potentially compromising the node's cryptographic keys [6, 7]. False sensing reports can be injected through compromised nodes, which can lead to not only false alarms but also the unwanted consumption of limited energy resources in the battery powered networks (Fig. 1) [4, 8]. To minimize damage, false reports should be dropped en-route as early as possible, and the few eluded ones should be further rejected at the base station [9].

Ye *et al.* [4] proposed the statistical en-route filtering scheme (SEF) to filter out forged reports during the forwarding process. In the SEF, multiple sensing nodes collaboratively generate a sensing report that contains multiple message authentication codes (MACs) so that each MAC is generated by a node using one of its symmetric keys,

Received August 14, 2007; revised August 8 & November 4, 2008; accepted January 8, 2009.

Communicated by Ten-Hwang Lai.

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

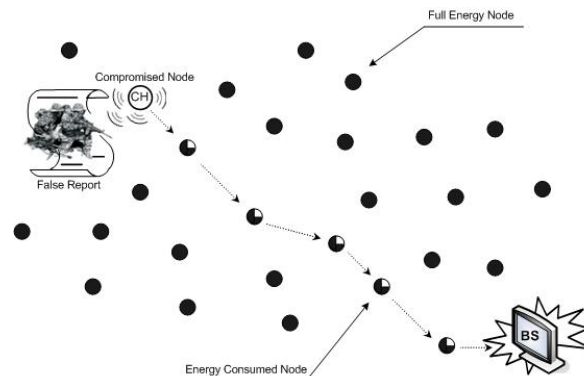


Fig. 1. False data injection.

and represents its agreement on the report [10]. As a report is forwarded towards the base station over multiple hops, each forwarding node verifies the MACs carried in the report, checking if it has any of the keys used to generate those MACs. If it does not have any of those keys, the report is forwarded without verification. Therefore, the detection power of the SEF is largely affected by the choice of routing path.

In this paper, a path selection method is proposed, to improve the detection power of the SEF where each message to establish the routing paths contains additional information about the keys of the visited nodes. By using this information, each node evaluates the detection power of each incoming path from the base station. Thus, it can choose the most secure path against false data injection attacks, which can lead to early detection of false reports and thus save energy. The proposed method is basically designed for the SEF scheme but it can be applied to other filtering schemes which exploit a global key pool approach. For example, to increase the early-detection power of false reports, the proposed method can be applied to the dynamic en-route filtering scheme (DEF) [12], the probabilistic voting-based filtering scheme (PVFS) [10], the multipath en-route filtering scheme (MEF) [15], and the fuzzy-based en-route filtering scheme (FEF) [16].

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

2. BACKGROUND AND MOTIVATION

2.1 Statistical En-route Filtering

SEF is the first scheme that addresses false data injection attacks in the presence of compromised nodes and it focuses on the detection of false event reports, which are known as false positive attacks, injected by compromised nodes. In SEF, the base station maintains a global key pool, which is divided into multiple partitions and every node loads a small number of keys from a randomly selected partition in the global key pool before it is deployed. Fig. 2 shows an example of the global key pool.

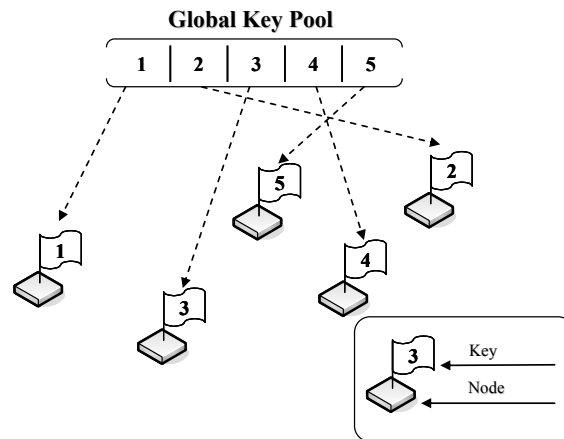


Fig. 2. Global key pool.

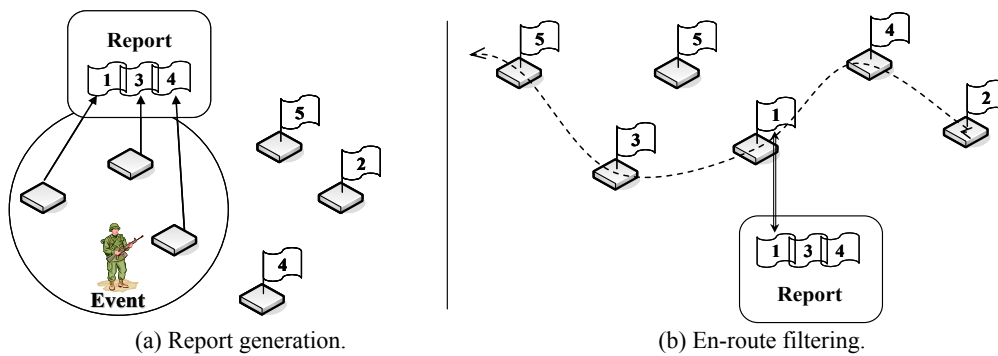


Fig. 3. Report generation and en-route filtering.

When real events occur, one of the detecting nodes is elected as the center-of-stimulus (CoS) node to generate a sensing report. The surrounding nodes, which detect the same event, produce MACs for the event, using their stored keys, and sends them to the CoS which generates a sensing report using the collected MACs. This set of multiple MACs acts as the proof that a report is legitimate [4] after which points the CoS forwards the report toward the base station (BS) over multi hops. Each forwarding node verifies the correctness of the MACs carried in the report by using its keys. When the BS receives a report, it can verify all the MACs carried in the report because it has complete knowledge of the global key pool [4]. Fig. 3 shows an example of the (a) report generation, and (b) en-route filtering in the SEF scheme.

An adversary can inject a forged report with incorrect MACs through a compromised node, as shown in Fig. 4 (a). However, the forged report may be dropped since each forwarding node verifies the correctness of the MACs carried in the report with a certain probability (Fig. 4 (b)). The probability of detecting incorrect MACs increases with the number of hops the report travels so the SEF can detect false reports forged by an adversary with a fixed number of compromised partitions.

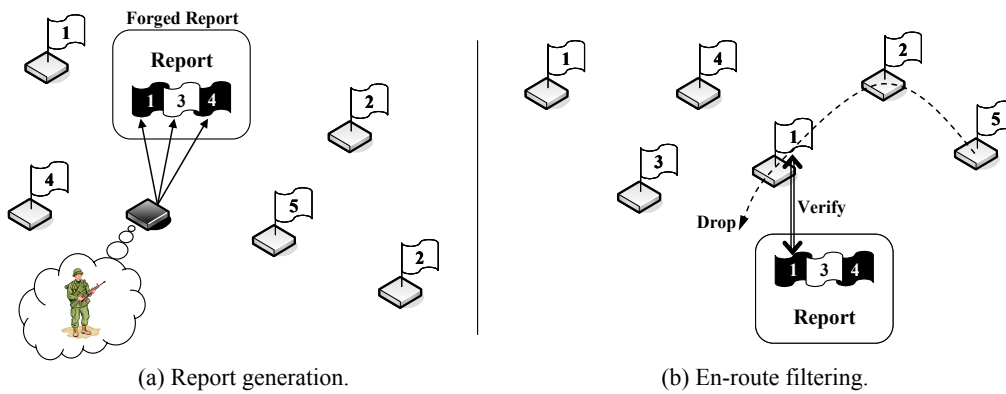


Fig. 4. Forged report filtering.

2.2 Motivation

In SEF, the detection power of false reports is largely affected by the choice of the routing paths. In the worst case, forwarding nodes may not have any of the keys used in report generation so these forwarding nodes cannot verify any false reports. Of course, the reports will not cause any false alarms since the base station can (and will) detect and reject them. However, transmitting the false reports drains the finite energy resources of the forwarding nodes so that it results in reducing the lifetime of the network. On the other hand, paths cannot be selected with the only consideration being the security issue and detour paths, chosen based on just the detection power, may reduce energy consumption for the false traffic. However, they may consume more energy in forwarding legitimate reports. Therefore, there is a requirement to choose the routing paths that can provide both the sufficient detection power needed for identifying false reports and over-all energy saving.

3. PATH SELECTION METHOD

3.1 Assumption

A sensor network is considered, composed of a large number of small sensor nodes. The network is static (*i.e.*, the topology of the network does not change), and it is assumed that the routing paths are established by flooding with a control message. This fashion is commonly used in most routing protocols (*e.g.*, directed diffusion [17] and minimum cost forwarding algorithms [18]) at the initial establishment of the paths. It is also assumed that the network uses a single-path routing protocol. To simplify the problem, it is further assumed that each node chooses a routing path based on the distance from the BS in the hop count, and the security level against false data injection attacks. After suitable paths have been established, every node forwards packets sent by its downstream (toward source nodes) nodes along the paths and packets sent by others are discarded immediately. An adversary can launch false positive attacks using compromised nodes but it is assumed that the adversary cannot compromise while the routing

paths are being set up. The adversary should launch false positive attacks only after the paths have been already-established, since every node accepts only packets forwarded by its own downstream nodes. That is, a false report should be sent to the most upstream (toward the BS) node of the compromised node. Otherwise the report will be dropped by a receiver at once. The issues of other security attacks, such as false negative attacks, are out of the scope of this paper.

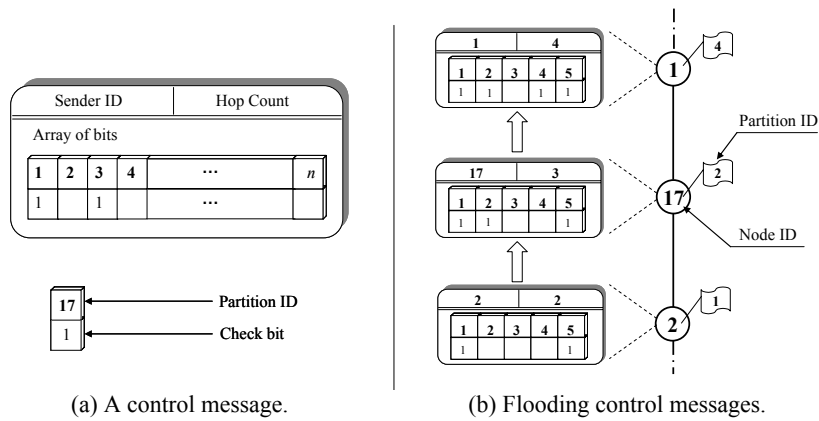
3.2 Overview

In the proposed method, every control message contains additional information about the partition IDs of visited nodes. The information is used to evaluate the security level of a path. As a control message is flooded through the network over multiple hops, each forwarding node updates the partition ID information of the message. A routing path is chosen by the node, using a qualification evaluation function, which considers the distance travelled and the security level. By varying a given security weighting factor, the user can give more priority to the security aspects or to saving energy.

3.3 Path Selection Method

In the proposed method, the BS maintains a global key pool divided into multiple partitions and each partition has a unique identification number. Each node has some keys from a randomly selected partition in the key pool before it is deployed. These keys are used to generate or verify the MACs. After node deployment, routing paths are established by flooding with a control message, which the BS broadcasts.

In most routing protocols, a control message contains the sender's ID and the number of hops that have taken place from the BS. In the proposed method, an array of bits is additionally attached into each control message and this array is used to mark the partition IDs of the visited nodes. An example form of a control message is shown in Fig. 5 (a), when the number of partitions in a global key pool is n . When a node receives a control message, it stores the sender's ID, the hop count from the BS, and the partition ID array attached in the message. The stored information is used to evaluate the detection power of the incoming paths. If the received message is the first instance of the control message, it sets the partition ID of its keys in the partition array of the message, and increases the hop count in the message. Then, it forwards the updated control message and Fig. 5 (b) shows how a partition ID array can be updated when a global key pool is divided into five partitions. Nodes 2, 17, and 1 have some keys loaded from partition 1, 2, and 4, respectively. When node 2 receives a control message, shown in Fig. 5 (b), it stores the information attached in the message. If the received message is the first instance of the message, it sets the first bit of the partition ID array in the message (since it loads some keys from partition 1). Then, the node increases the hop count of the message and forwards the updated message. When node 17 receives the message, it stores the information. If the message is the first instance, the node sets the two bits of the array, increases the hop count and forwards the updated message. Node 1 also stores the information and forwards the updated message if necessary.



(a) A control message. (b) Flooding control messages.

Fig. 5. Format of the control message & flooding with the control message.

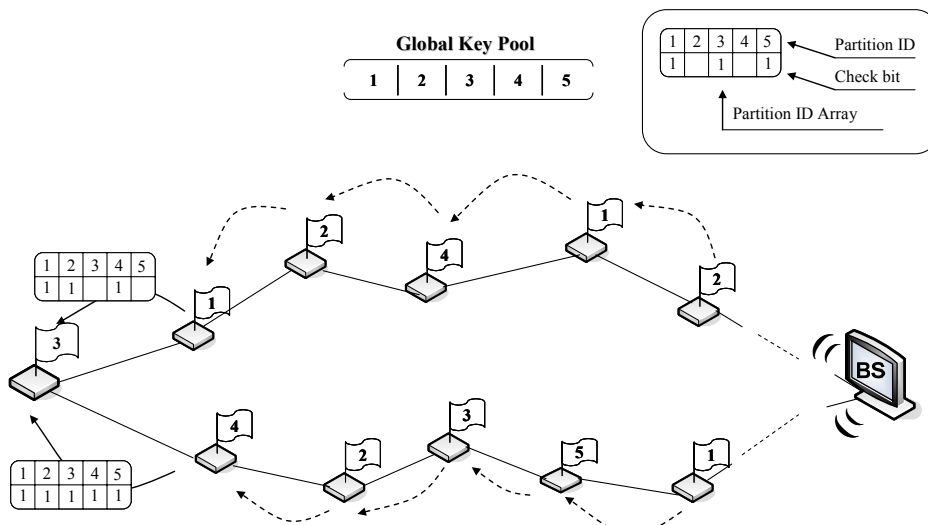


Fig. 6. Updating of partition ID array.

After the flooding using the control message, every node evaluates the detection power of each incoming path, based on the stored partition ID array. If all the bits in the array of a path are set, the path will detect the majority of false reports. That is, the path is most secure against false data injection attacks. If a very small number (or none) of bits are set, the majority of reports will not be verified by the forwarding nodes. The node may be vulnerable to false data injection attacks. Note that there is a trade-off between detection power and overhead. For a legitimate report, the former path may consume more energy than the latter path because computational overhead is incurred when a node verifies a received report. As seen in Fig. 6, the lower path may be the most secure against false data injection attacks since all the bits in the array are set. However, it may consume more energy than the upper path in report forwarding. On the other hand,

the upper path is more vulnerable than the lower but it may be more energy efficient than the lower path in report delivering.

A path is chosen by an evaluation function that decides the qualification regarding the path that is both most secure and yet energy conserving, based on the detection power and hop count of the path. An evaluation function can take the following form:

$$Q(P) = D(p) + \omega \cdot P(p). \tag{1}$$

Where p is a path, $D(p)$ is the distance of p in the hop count, ω is a security weight factor determined by the user, and $P(p)$ is the number of unset bits in the partition ID array received from p . Note that a smaller $Q(p)$ is more qualifiable than a larger one.

The security weighting factor, ω , can be between 0 and 1, as that as ω increases, each node would choose a more secure but less energy-efficient, path. On the other hand, the network does not consider the detection power if ω is 0; each node would choose the shortest path as its routing path. Suppose for example, as shown in Fig. 7, node 1 can choose one of the two paths, P_1 and P_2 as its routing path. If ω is 0.25, the qualification of the P_1 , $Q(P_1)$, is $9 + (0.25 \cdot 5) = 10.25$ and the qualification of the P_2 , $Q(P_2)$, is $10 + (0.25 \cdot 3) = 10.75$. Thus, node 1 would select P_1 , the energy-efficient but less secure path, as its routing path. On the other hand, if ω is 0.75, the qualification of P_1 , $Q(P_1)$, is $9 + (0.75 \cdot 5) = 14$ and the qualification of P_2 , $Q(P_2)$, is $10 + (0.75 \cdot 3) = 13$. Thus, node 1 would select P_2 , which has more detection power but is less energy-efficient than P_1 , as its routing path.

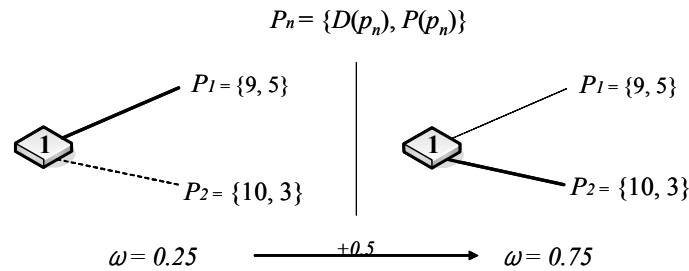


Fig. 7. Path selection in compliance with a security weighting factor.

4. SIMULATION RESULTS

To show the effectiveness of the proposed method, the original SEF scheme is compared with the new path selection method (represented by PSM) through simulation studies. The sensor network in the simulation environment consists of 2,000 nodes over a field size is $140 \times 110m^2$. The nodes are assumed to be randomly distributed in the field and the base station is located at the end of the field. Each node takes $16.56\mu J/12.5\mu J$ to transmit/receive a byte, and each MAC generation consumes $15\mu J$ [4]. The size of a MAC is 1 byte, and the report size is 12 bytes. There are 1,000 keys in the global key pool, which is divided into 20 partitions. Every node evaluates the qualification of a path using Eq. (1).

In Fig. 8 shows the ratio of the filtered false reports, according to the ω value, when the number of forged MACs in a report is 1, 4, 10 and 16. As shown in the figure, the proposed method ($\omega = 0.50, 0.75$ and 1.00), which means the average of ω , can filter out a larger number of false reports, during the forwarding process, than the original SEF ($\omega = 0.00$) since a routing path in the proposed method is chosen based on not only the distance but also the detection power.

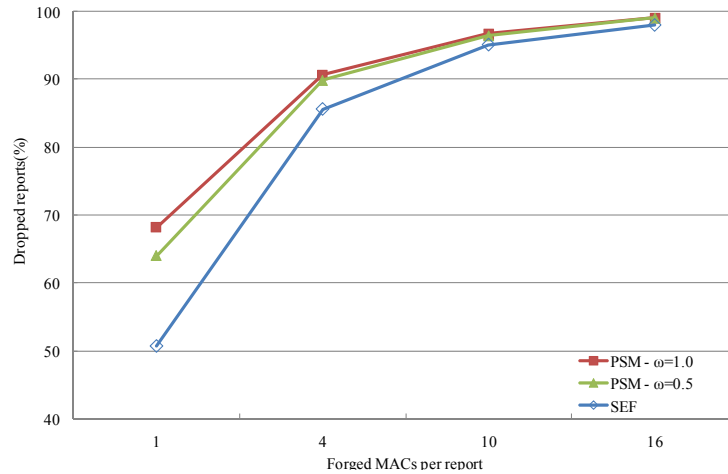


Fig. 8. Ratio of filtered false reports according to the value of ω .

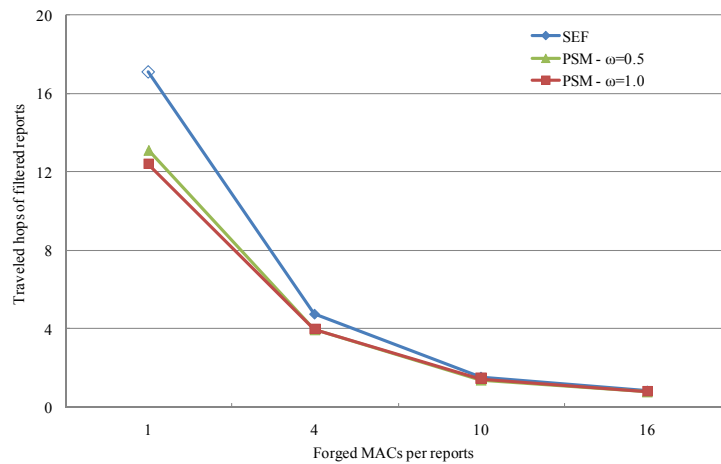


Fig. 9. Average number of hops that a filtered report traveled.

Fig. 9 shows the average number of hops that a filtered report traveled when the number of forged MACs per report is 1, 4, 10 and 16. As shown in the figure, the proposed method can detect false reports earlier than the original SEF approach, since routing paths are chosen with consideration of the detection power in the proposed method. When $\omega = 1.00$, the result is seen to be close to that of $\omega = 0.50$.

Fig. 10 shows the average energy consumption, caused by a false report, when the number of forged MACs per report is 1, 4, 10 and 16. As shown in the figure, the proposed method ($\omega = 0.5$) can conserve more energy than the original SEF approach since it can detect and drop false reports earlier than SEF, before they consume a significant amount of energy. When $\omega = 1.00$, the result closes to that of $\omega = 0.50$.

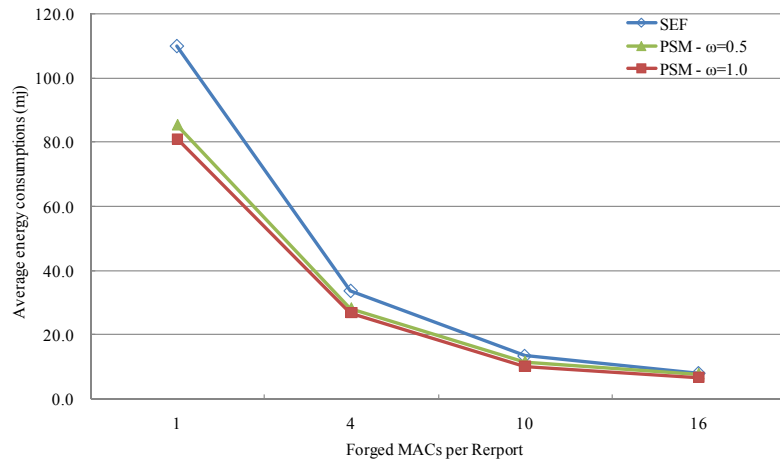


Fig. 10. Average energy consumptions per false report.

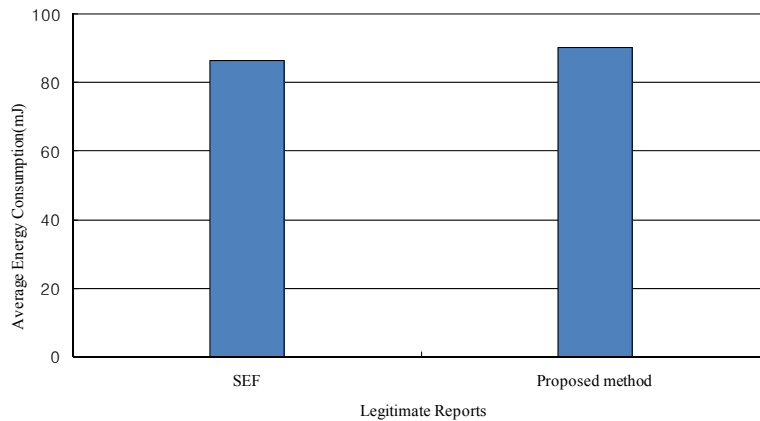


Fig. 11. Average energy consumption per legitimate report.

Fig. 11 shows the average energy consumption per legitimate report. The proposed method is less efficient than the original SEF ($\omega = 0.00$). However, the difference between them is very small. For a legitimate report, the choice of ω does not greatly affect the energy consumption characteristics. On the other hand, as ω increases, more energy can be saved from false data injection (Fig. 10). That is, the proposed method can conserve energy better than the original SEF scheme ($\omega = 0.00$).

5. RELATED WORK

During the past years, the architecture and design of sensor networks and hardware have progressed significantly [14]. Recently, the security of sensor networks has become an important research issue and several techniques have been proposed to address report authentication and fabricated report filtering with symmetric cryptography-based security solutions in wireless sensor networks. The research most related to the current work can be found in [4, 7, 9, 11, 12] as well as good overviews on the existing schemes for filtering false reports, and how the proposed scheme can take advantage of other filtering schemes as well as the SEF.

Yu *et al.* [12] proposed a dynamic en-route filtering scheme based on the SEF scheme where the filtering of the false data injection pre-distributes the symmetric keys in such a way that the keys are associated with the sensor location. This scheme can better deal with a dynamic topology of sensor networks, and outperform them in terms of energy efficiency in comparison to existing schemes. This scheme is more resilient than SEF because it is hard to compromise a number of sensor nodes in the same area to make false reports. Zhu *et al.* [7] proposed an interleaved hop-by-hop authentication (IHA) scheme to verify reports with a pair-wise symmetric keys in a deterministic and hop-by-hop fashion using a cluster-based organization so that the scheme guarantees that the false data will be detected at the base station when no more than a certain number t of nodes are compromised. Yang and Lu [11] presented a commutative cipher based en-route filtering (CCEF) scheme [9] which can verify the false reports through an encrypted session key, and un-encrypted witness key in the Query message. In a similar way as the IHA scheme, the CCEF requires a static topology of sensor networks, at least during the requesting and responding processes. Lee and Cho [13] proposed a key inheritance-based false data filtering (KIF) scheme as an enhanced interleaved authentication solution that prevents forwarding of false data. In the proposed method, an intermediate node shares some keys with all of its upstream nodes within a certain number t of hops. This guarantees that false data will be detected within t hops. Li *et al.* [10] proposed a probabilistic voting-based filtering scheme (PVFS) to detect a false negative attack. In the proposed method, injected false data attacks and injected false data on legitimate reports attack use a combination of cluster-based organization, probabilistic key assignment, and voting.

6. CONCLUSION

A path selection method for improving the detection power of the SEF scheme is presented. By using the partition ID arrays received from the incoming paths, each node can choose the safest path against false data injection attacks. A simple function to evaluate the qualification of the paths can be provided. The effectiveness of the proposed method is shown using simulation results. As stated in section 3.3, the choice of the routing path represents a trade-off between the desired security level and the overhead incurred. In the proposed method, the partition ID array does not completely reflect the possession of forwarding nodes since each node does not have all keys in a partition. Therefore, other approaches, or factors that represent such possession better, should be

considered. Future work will also study the selection of the routing paths by considering other factors such as the remaining energy of the nodes or the false traffic proportion. There can be other security attacks on sensor networks. For example, compromised detecting nodes may launch false negative attacks [4] in which these nodes insert false MACs into legitimate reports. These reports may be regarded as false ones, and thus be dropped by forwarding nodes. However, this is a different problem from false positive attacks that SEF and the proposed method address. The research of countermeasures against such attacks remains for future work. PSM assumes that the network is safe from node compromise during the routing path set-up phase. However, in dynamic networks, routing paths may change upon a change of the network topology or a user's request. Thus, some nodes may be compromised while a control message is flooding the network and this issue will be also studied further in future work.

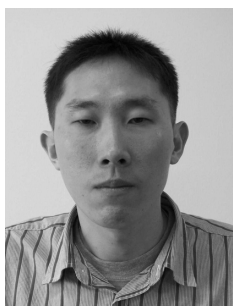
REFERENCES

1. K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, Vol. 3, 2005, pp. 325-349.
2. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, Vol. 40, 2002, pp. 102-144.
3. S. H. Chi and T. H. Cho, "Fuzzy logic based propagation limiting method for message routing in wireless sensor networks," *Lecture Notes in Computer Science*, Vol. 3983, 2006, pp. 58-64.
4. F. Ye, H. Luo, and S. Lu, "Statistical en-route filtering of injected false data in sensor networks," *IEEE Journal on Selected Areas in Communications*, Vol. 23, 2005, pp. 839-850.
5. Q. Jiang and D. Manivannan, "Routing protocols for sensor networks," in *Proceedings of the 1st Conference on Consumer Communications and Networking*, 2004, pp. 63-98.
6. B. Przydatek, D. Song, and A. Perrig, "SIA: Secure information aggregation in sensor network," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*, pp. 255-265.
7. S. Zhu, S. Setia, S. Jajodia, and P. Ning, "An interleaved hop-by-hop authentication scheme for filtering of injected false data in sensor networks," in *Proceedings of IEEE Symposium on Security and Privacy*, 2004, pp. 259-271.
8. W. Zhang and G. Cao, "Group rekeying for filtering false data in sensor network: A predistribution and local collaboration-based approach," in *Proceedings of 24th IEEE Annual Joint Conference on the Computer and Communications Societies*, 2005, pp. 503-514.
9. H. Yang and S. Lu, "Commutative cipher based en-route filtering in wireless sensor networks," in *Proceedings of IEEE Conference on Vehicular Technology*, 2003, pp. 1223-1227.
10. F. Li and J. Wu, "A probabilistic voting-based filtering scheme in wireless sensor networks," in *Proceedings of International Conference on Wireless Communications and Mobile Computing*, 2006, pp. 27-32.
11. H. Yang, F. Ye, Y. Yuan, S. Lu, and W. Arbaugh, "Toward resilient security in

- wireless sensor networks,” in *Proceedings of the 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2005, pp. 34-45.
12. Z. Yu and Y. Guan, “A dynamic en-route scheme for filtering false data in wireless sensor networks,” in *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, 2006, pp. 294-295.
 13. H. Y. Lee and T. H. Cho, “Key inheritance-based false data filtering scheme in wireless sensor networks,” *Lecture Notes in Computer Science*, Vol. 4317, 2006, pp. 116-127.
 14. V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, “Energy aware wireless microsensor networks,” *IEEE Signal Processing Magazine*, Vol. 19, 2002, pp. 40-50.
 15. M. S. Kim and T. H. Cho, “A multipath en-route filtering method for dropping in sensor networks,” *IEICE Transactions on Information and Systems*, Vol. E90-D, 2007, pp. 2108-2109.
 16. M. S. Kim and T. H. Cho, “A fuzzy-based en-route filtering scheme in sensor networks,” *Lecture Notes in Computer Science*, Vol. 4681, 2007, pp. 230-239.
 17. C. Intanagonwivat, R. Govindan, and D. Estrin, “Directed diffusion: a scalable and robust communication paradigm for sensor networks,” in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, 2000, pp. 56-57.
 18. F. Ye, A. Chen, S. Lu, and L. Zhang, “A scalable solution to minimum cost forwarding in large sensor networks,” in *Proceedings of the 10th International Conference on Computer Communications and Networks*, 2001, pp. 304-309.



Chung Il Sun (宣清日) received his B.S. degree in Computer Science from Kyungwon University, Korea, in 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, and security in wireless sensor networks.



Hae Young Lee (李海英) received his Ph.D. degree in Computer Engineering and B.S. degree in Electrical and Computer Engineering from Sungkyunkwan University, Korea, in 2009 and 2003, respectively. He is with the School of Information and Communication Engineering, Sungkyunkwan University, Korea. His research interests include wireless sensor network, intelligent system, artificial intelligence, modeling and simulation, design automation.



Tae Ho Cho (曹大昊) 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.