韓國컴퓨터情報學會 論文誌
第16卷 第10號, 2011. 10.

2011-16-10-3-3

무선 센서 네트워크에서 최소 통신비용 수행을 위한 허위 데이터 식별 프로토콜

Anuparp Boonsongsrikul*, 박 승 규*, 신 승 훈**

An Approach of False Data Identification Protocol for Minimum Communication Cost in Wireless Sensor Network

Anuparp Boonsongsrikul*, Seung-Kyu Park*, Seung-Hun Shin**

요 약

무선 센서 네트워크에서 compromised node는 데이터 병합과정에서 허위 데이터를 삽입할 수 있다. 데이터 병합의 보안성을 위한 기존 접근 방법들은 높은 연산 부하를 요구한다. 따라서 본 논문에서는 허위 데이터 삽입 공격 발생 지점을 식별하는데 소요되는 통신 부하를 최소화하는 모니터링 기반 시큐어 데이터 병합 프로토콜을 제안한다. 제안된 프로토콜은 허위 데이터 삽입이 탐지되면 모니터링을 수행하고 있던 노드들의 MAC(Message Authentication Code)을 하나의 메시지로 요약하고 이를 BS(Base Station)로 전송하는 방법을 사용하며, BS는 이를 통해 공격 노드를 식별한다. 실험 결과는 제안된 프로토콜이 MAC들의 짧은 연결과 보통 연결을 사용하는 경우, 기존 연구에 비해 각각 45% 및 36% 적은 에너지를 사용하는 것으로 나타났다.

▸ Keyword : 센서 네트워크, 데이터 병합, 에너지 소모, 보안, 허위 데이터 삽입 공격

Abstract

In wireless sensor networks, a compromised sensor node can inject false data during data aggregation. Existing solutions of securing data aggregation require high communication cost in securing data aggregation. In this paper, we propose a monitoring-based secure data aggregation

• 제1저자 : Anuparp Boonsongsrikul • 교신저자 : 박승규
• 투고일 : 2011. 08. 04, 심사일 : 2011. 08. 18, 게재확정일 : 2011. 08. 28.
* 아주대학교 정보통신공학과(Dept. of Information and Communication, Ajou University)
** 아주대학교 정보컴퓨터공학부(Div. of Information and Computer, Ajou University)

protocol that minimizes communication cost of identifying the location of false data injection attacks. The main idea is that when monitoring nodes find an injected false data, their reporting messages along with Message Authentication Codes (MACs) are summarized in a single message before sending it to the Base Station (BS). Then the BS identifies the attacking node. The simulation shows that energy consumption of the proposed protocol with short and normal concatenations of MACs are 45% and 36% lower than that of an existing protocol, respectively.

▸ Keyword : Sensor Networks, Data Aggregation, Energy Consumption, Security, False Data Injection Attacks

I. Introduction

Wireless sensor networks are vulnerable to many types of security attacks. One of them is an false data injection attack. Existing protocols require much communication cost for securing aggregation data and this results in sensor nodes consuming a significant amount of energy. In spite of high communication cost, most existing protocols can only detect the occurrence of an attack without being able to identify the location of the attacking node.

In this paper, we propose a secure data aggregation protocol that minimizes communication cost of identifying false data injection attacks. We improve the work of Boonsongsrikul[1] which aims to detect both the occurrence of an attack and the location of the attacking node while other related works only detect the occurrence of an attack.

In the work of Boonsongsrikul[1], when sensor nodes overhear an abnormal data of their parent, they send their reporting messages to the Base Station (BS) via their shortest paths. However, the problem with this work is that when many nodes are affected with an abnormal value, then those affected nodes will send reporting messages to the BS. Therefore many shortest paths will be set up. The reporting process is taken place by creating new multiple paths. This introduces significant communication cost due to increasing messages for forwarding reports by intermediate nodes.

The main difference between this work and the work of Boonsongsrikul[1] is that when a node injected false data aggregation message, its children send reporting messages to their grandparent. A grandparent of reporting nodes summarizes those reporting messages in a single report and sends it to the BS. Note that a Message Authentication Code (MAC) between each reporting node and the BS is included in the report.

The rest of this paper is organized as follows. Section 2 presents related work. Section 3 presents assumptions and an attacker model. Section 4 presents a proposed scheme. Section 5 evaluates the effectiveness of our protocol and shows the simulation results. Section 6 concludes the paper.

II. Related work

Existing solutions require high communication cost to secure aggregation data. However even with such high communication cost, most of them can only detect the occurrence of an attack. The following is a brief description of ideas and shortcomings for previous works.

He and et al.[2] proposed two different protocols to secure privacy of aggregation data. In the first one, all sensor nodes in a cluster exchange their encrypted data before sending encrypted data to their cluster head. A cluster head then sends aggregation data to the BS. In the second one, a sensor node slices its encrypted data into J pieces and sends them to J aggregators. All nodes in a network then send an aggregate value to the BS. Both protocols achieve to secure privacy of

aggregation data but require higher communication cost to secure aggregation data.

Chan and et al.[3] proposed secure hierarchical data aggregation with a hash tree. A node sends an aggregation data to its parent. The BS broadcasts a final aggregate to all nodes. Whether there is an attack or not, every node has to send an confirmation message which is hierarchically aggregated by its parent to the BS. Yang and et al.[4] proposed secure hop-by-hop data aggregation. The protocol divides sensor nodes into multiple subtrees. The BS then finds a suspicious aggregate from a set of aggregates. All nodes in a suspicious subtree resend aggregation data to the BS. Works of [3] and [4] can detect the occurrence of an attack but require many nodes and messages during the process of detection. This induces communication cost to increase.

Boonsongsrikul and et al.[1] proposed a secure data aggregation protocol to identify false data injection attacks. This protocol is divided into three phases: 1) the query dissemination phase, where the base station initiates the aggregation; 2) the aggregation phase, where all nodes perform the aggregation; and 3) the attestation phase, where suspecting nodes send verification messages to the BS to find suspicious nodes and verify them. A node who detected an abnormal value of its parent sends its reporting message to the BS via a shortest path. The BS uses the reporting nodes' IDs and the parent's ID in the reporting messages to build an attestation tree. The leaf nodes' parent of the attestation tree will be the most suspicious node who could have injected false data. However, the problem with this work is that when many nodes are affected with an abnormal value, then those affected nodes will send reporting messages to the BS by creating multiple reporting paths. The communication cost by consequence is increased due to the increased report messages by intermediate nodes.

III. Assumptions and an attacker model

1.Network topology

We use a K-ary tree such as in [5] and assume that among children of an internal node, their transmission ranges can reach each other. In other words, children can monitor all incoming and outgoing data of their parent. To achieve this condition, sensor nodes are densely deployed in a network.

2. Unique keys and pair-wise keys

We assume that node i shares a different unique key with the BS. So the BS can authenticate a message of node i. Such keys are pre-installed before the network is deployed, hence they do not require any run-time establishment. A pairwise key is used to compute a message authentication code between a node and the BS. We call it BMAC [1] and [6].

We also assume that node i shares a pairwise key with node j. So node j can authenticate a message of node i and vice versa. A node also shares a pairwise key with its grandparent and shares another pairwise key with its parent. A pairwise key is used to compute a message authentication code between a sender and a receiver. We call it AMAC [1] and [6].

3. Attacker Model

We use the same attacker model in [3] and [4]. A compromised aggregator injects aggregation data of their non-compromised children. Its goals are either tricking the BS into receiving forged data or draining energy of other sensor nodes who send aggregation or reporting messages when sensor nodes detect false data. We do not consider a denial-of-service (DoS) attack where its goal is preventing the BS from getting any aggregation data.

IV. Proposed protocol

In this section, we propose a secure data aggregation protocol that minimizes communication cost in detecting false data injection attacks.

Table 1. A format of messages

Message	Format
aggregate	{op, id, pid, D, AMAC}
report	{op, id, gid, sid, sD, AMAC, BMAC}
attest	{op, id, sid, sD, BMAC, $BMAC_{con}$}

Table 2. Notation of each field in messages

Notation	Meaning
op	2-bit specifying the message type
id	node identifier (sender ID)
gid	grandparent ID
pid	parent ID
sid	suspicious ID
D	an aggregated value
sD	a suspicious value
Ki,j	Key between node i and j
AMAC	MAC of data A using a pairwise key Ki,j between node i and j where A is op, id, pid and D
BMAC	MAC of data B using a unique key Ki,BS between node i and the BS where B is op, id, gid, sid and sD
$BMAC_{con}$	A normal concatenation of MACs type B (BMACs)
$BMAC_{FAN}$	A short concatenation of MACs type B or a short $BMAC_{con}$ by means of FAN[7]

The proposed protocol divides an aggregation session into three phases which are similar to Boonsongsrikul[1]. When an aggregator Ai sends an aggregation message to itsparent PAi, children of Ai also monitor aggregation data. If there is any inconsistency between aggregation data computed by Ai and one computed by the children then those children send reporting messages to their grandparent PAi. If aggregation data sent by Ai and all of its children of Ai are the same, then PA

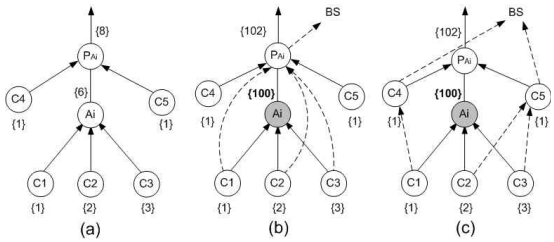

Fig.1. Models of proposed schemes in detecting false data aggregation for a SUM function

Table 3. Securing data aggregation protocol

Input: Aggregator Ai, j children {C1, C2, ..., Cj}
Output: Inconsistency between aggregation data computed by Ai and aggregation data monitored by {C1, C2, ..., Cj}, a single report is sent to the BS.
1: Ai receives and combines aggregation data of its children and sends its aggregation data AggDAi to its parent PAi
2: {C1, C2, ..., Cj} monitor incoming aggregated data of PAi
3: If there is inconsistency between aggregation data computed by {C1, C2, ..., Cj} and AggDAi , then {C1, C2, ..., Cj} send reporting messages to PAi
4: If aggregation data computed by {C1, C2, ..., Cj} and AggDAi are the same, then PAi summarizes reporting messages into an attestation message and sends it along with $BMAC_{con}$ to the BS
5: The BS then verifies an attestation message using a unique key sheared with each sensor node. If an aggregation report matches $BMAC_{con}$ computed by the BS, then Ai is a compromised node.

summarizes those reporting messages into a single message called "attestation message". Next, PAi sends an attestation message identifying that Ai is a compromised node to the BS.

Let us give three approaches in an aggregation function of SUM. First, let node C1, C2 and C3 send value 1, 2 and 3, respectively to their parent Ai. After verifying AMACs (refer to table 1 and 2), node Ai has an aggregated value which is equal to 6 as shown in Fig 1(a). Second, node C1, C2 and C3 send value 1, 2 and 3 to their parent Ai but instead of sending value 6, node Ai (gray) sends value 100 (injected false data). Therefore, node C1, C2 and C3 send reporting messages to their grandparent as shown in Fig.1(b). This approach is our proposed protocol for identifying the attacking

node. Third, if node Ai (gray) injects an abnormal value, C1, C2 and C3 send reporting messages to their BS via (different) shortest paths as shown in Fig.1(c). This approach is proposed by Boonsongsrikul[1]. Dashed lines represent sending reporting messages to the BS. Solid lines represent sending aggregated values to a parent.

Our protocol aims to minimize the number of reporting messages. Instead of sending reporting messages to the BS independently by each node, only one summarized attestation message will be sent to the BS by grandparent PAi. BMACs of monitoring nodes can be thought of as witnesses to identify a compromised node. A grandparent PAi concatenates those BMACs in a format of $BMAC_{con}$ {idc1, BMACc1, idc2, BMACc2, ... idi, BMACi} where idi, BMACi are id and BMAC of node i, respectively. To minimize the size of a concatenation of BMACs, we use the technique in reducing the MAC size proposed by Fan [7]. This can reduce the required storage space to 30% of $BMAC_{con}$. We call this technique $BMAC_{FAN}$ in our paper.

Table 4. Energy per bit and energy per message

s (b = bit)	δ (μJ/b)	ETx (mJ)	ERx (mJ)
141	2.44	37.75	34.43
282	2.12	60.03	59.74
348	2.06	72.60	71.84

When an attestation message arrives at the BS, then the BS verifies the attestation message using a unique key shared with a node who sent the BMAC. In the scenario in this paper, there are four MACs: one BMAC of a grandparent node and three BMACs of grandchildren. Since the BS knows IDs of nodes that sent the BMACs, the BS uses four different keys to compute such four BMACs. If they match, then sid in an attestation message is the ID of a compromised node. After the compromised node is identified, the BS will flood instruction to revoke compromised keys and propagate new ones. The BS's instructions can be authenticated by hash chains as proposed by Perrig [8]. The research on the issue of revoking compromised keys is left in the future work.

V. Modelling attack detection and sensor network simulation

To model attack detection and simulate energy consumption, we divide this section into three parts. First, we setup the network environment and the location of an attack in a network. Second, we calculate the energy per bit for sending attestation messages. Third, we simulate energy consumption based on a number of BMACs in an attestation message.

1. Network environment

We set the network environment as follows. The transmission range of a sensor node is 15 m, the number of nodes is 350 and an area is 200 × 20 m^2. A compromised aggregator is randomly located in a network so as to evaluate communication costs for identifying attacking node. Since the energy is consumed proportionally to the amount of communications for sending and forwarding attestation messages to the BS, we will replace the cost model of communication by that of energy consumption from sensor nodes. Its location is limited to maximum of 100 hops from the BS. Three BMACs of child nodes participate in an attestation message.

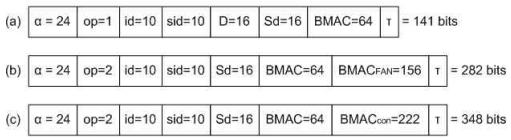

Fig. 2. (a) an attestation message in Boonsongsrikul [1], (b) an attestation message with $BMAC_{FAN}$, (c) an attestation message with $BMAC_{con}$.

2. The model of energy per bit

Since the sizes of attestation messages in each scheme are different, the energy consumption for sending or receiving a bit is different as well. Therefore a bit-level energy model for communication is required.

We use energy model of Sankarasubramaniam [9] to compute the energy per bit, δ.

$$\delta = k_1 + k_1\frac{\alpha+\tau}{l} + \frac{k_2 + E_{dec}}{l} \quad (1)$$

where k1 = 1.85 μJ/bit is the consumed energy in the communication of a bit at data rate 20 Kbps, k2 = 24.86 μJ is the start-up energy consumption, α,l, and τ are a header, a payload and a trailer field, respectively in the data link layer ($s = \alpha + l + \tau$) and E_{dec} is the decoding energy per packet. For simplicity, let us assume the following parameters. α is 24 bits. No error control is used (τ, E_{dec} = 0). The size of l is the size of a reporting message to the BS. There are three monitoring nodes (3 BMACs) in which the node ID is 10 bits, a data value 16 bits, and BMAC 64 bits.

In Boonsongsrikul [1], the size of an attestation message requires 141 bits (Fig 2(a)). For $BMAC_{FAN}$ and $BMAC_{con}$ in this paper, the sizes of attestation messages require 282 bits (Fig 2(b)) and 348 bits (Fig.2(c)), respectively where parameters op, id, sid, D, Sd etc. are explained in Table 1.

We model the energy per bit for each scheme as shown in the Table 4. The amount of the energy per bit for different message sizes given in Fig. 2(a), 2(b) and 2(c) are calculated by equation (1) as 2.44, 2.12 and 2.06 μJ/b, respectively.

Note that when we increase the number of BMACs in an attestation message, the energy per bit δ is decreased and gets stable at a given number of bits. This results can be confirmed by the work of Shih [10].

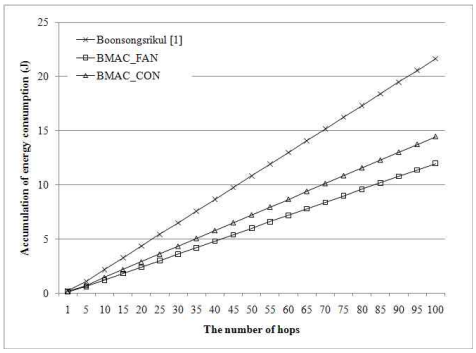

Fig. 3. Accumulation of energy consumption

3. Energy consumption for attack detection

In the proposed protocol, we use the energy model of Heinzelman [11] to establish a energy model for sending an attestation message E_{Tx} which is represented as the following equation,

$$s.(\delta + \theta.d^q), \quad (2)$$

where s is the message size and δ (μJ/b) is the energy required to communicate one bit of information across a single hop. The θ = 100 pJ/b/m is the coefficient for a distance-dependent term. The q = 2 is the exponent for the distance-dependent term, and d is the transmission distance. The energy in receiving a message of a node E_{Rx} is

$$s.\delta \quad (3)$$

The energy consumptions of transmission E_{Tx} in equation (2) and reception E_{Rx} in equation (3) are presented in Table 4. Note that E_{Tx} is an average energy for transmitting an attestation message.

In the case without any false data injection attacks, the energy consumption during sending data aggregation are same for both Boonsongsrikul [1] and the work in this paper. Thus, the evaluation took place for the case that a compromised aggregator injects false data and the attacking node is identified. The energy

consumptions are compared for sending attestation messages to the BS between the model of [1] and that of this proposed work. The results are averaged over ten simulated topologies.

Comparing the energy consumption individually with respect to each message, the proposed scheme consumes more than that of [1] due to the longer size of a combined message. Evaluating overall energy consumption for all messages, the proposed scheme consumes much less due to the fact that there is only one attestation message while the work of [1] Boonsongsrikul (2010) requires as much as that multiplied by the number of attestation messages which is represented by BMACs. The accumulation of energy consumption in the Fig.3 is the cumulated energy consumption of all intermediate nodes who forward an attestation message to the BS. Let a Rate of Energy Consumption, REC denote a ratio of accumulation of energy consumption to the total number of hops that an attestation message travelled to the BS.

As illustrated in Fig. 3, the more the number of hops we have in the network, the more reduction effect of REC we have in the proposed protocol with respect to the work of Boonsongsrikul [1]. At 100 hops, REC of Boonsongsrikul [1], REC of our protocol using $BMAC_{con}$ and REC of work using $BMAC_{FAN}$ are 0.22, 0.12 and 0.14 J/hop, respectively. By comparing our protocol using $BMAC_{FAN}$ and Boonsongsrikul [1], energy consumption of our protocol is 45% lower. By comparing our protocol using $BMAC_{con}$ and Boonsongsrikul [1], energy consumption of our protocol is 36% lower.

VI. Conclusion

In wireless sensor networks, compromised sensor nodes can distort data by injecting false data. Previously known techniques on false data detection do not aim to identify false data injection attacks so that the attacker has chances to repeatedly injects false data. As a result, sensor nodes waste energy for sending false data. The main reason goes to the fact that the mechanism can only detect the occurrence of attacks without being able to identify attacking node.

In this paper, we enhance the mechanism of identifying false data injection attacks in Boonsongsrikul [1]. The proposed scheme also identifies the compromised node using a sufficient number of BMACs. The number of attestation messages is significantly reduced by the proposed scheme for the intermediate nodes, which saves the energy accordingly. Simulation shows that by comparing our protocol using $BMAC_{FAN}$ with that of Boonsongsrikul [1], energy consumption of our protocol is 45% lower. In case comparing protocol using $BMAC_{con}$, energy consumption proposed in this paper $BMAC_{con}$ is 36% lower than that of Boonsongsrikul [1].

References

[1] A. Boonsongsrikul K. Lhee and M. Hong, "Securing Data Aggre gation against False Data Injection in Wireless Sensor Networks," The 12th International Conference on Advanced Communication Technology (ICACT'2010), pp. 29-34, 2010.

[2] W. He, X. Liu and H. Nguyen, K. Nahrstedt and T.T Abdelzaher, "PDA: Privacy-Preserving Data Aggregation in Wireless Sensor Networks," Proceedings of the 26th IEEE International Conference on Computer Communications. (IEEE INFOCOM 2007), pp. 2045-2053, 2007.

[3] H. Chan, A. Perrig and D. Song, "Secure Hierarchical In- Net work Aggregation in Sensor Networks," Proceedings of conference on Computer and Communications Security(CCS'06), 2006.

[4] Y. Yang, X. Wang, S. Zhu, and G. Cao, "SDAP: A Secure Hop-by-Hop Data Aggregation Protocol for Sensor Networks," Proceedings of The 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc'06), pp. 356-367, 2006.

[5] A. Boonsongsrikul, S. K. Park and S. H. Shin, "A Sextant Cluster Based Monitoring on Secure Data Aggregation and Filtering False Data in Wireless Sensor Networks," Journal of Korea Society of Computer and Information, Accepted.

[6] A. Boonsongsrikul, K. Lhee and S. K. Park, "Monitoring-Based Secure Data Aggregation Protocol against a Compr omised Aggregator in Wireless Sensor Networks," Journal of Korea Information Processing Society, Accepted.

[7] Y. Fan, H. Luo, L. Songwu and Z. Lixia, "Statistical En-Route Filtering of Injected False Data in Sensor Networks," IEEE Journal on selected areas in communications, vol. 23, No. 4, pp. 839-850, 2005.

[8] A. Perrig, R. Szewczyk, J.D. Tygar, V. WEN and D. Culler, "SPINS: security protocols for sensor networks," Journal of Wireless Networks, vol. 8, Issue 5, pp. 521 - 534, 2002.

[9] Y. Sankarasubramaniam, I.F Akyildiz and S.W McLaughlin, "Energy Effi ciency based Packet Size Optimization in Wireless Sensor Networks," Proceedings of the 1st IEEE International Workshop on Sensor Network Protocols and Applications, pp. 1-8 , 2003.

[10] E. Shih, B. H. Calhoun, H. C. Seong, A.P. Chandrakasan,, "Energy-efficient link layer for wireless microsensor networks," Proceedings IEEE Computer Society Workshop on VLSI 2001. Emerging Technologies for VLSI Systems, pp. 16-21, 2001.

[11] W. R. Heinzelman A. Ch and H. Balakrishnan, "Energy efficient communication protocol for wireless microsensor networks." Proceedings of the 33rd Hawaii International Conference on System Sciences, pp. 3005 - 3014, 2000.

저 자 소 개

Anuparp Boonsongsrikul

1998 : Mahanakorn기술대학 교통신공학과 공학사.

2002 : Kasetsart대학교 전자공학과 공학석사.

현 재 : 아주대학교 정보통신공학과 박사과정

관심분야: 센서 네트워크,에드혹 네트워크, VANET, IC 디자인 등

Email : anuparp@ajou.ac.kr

박 승 규

1974 : 서울대학교응용수학과 학사

1976 : 한국과학원(KAIST) 전산학과 석사

1982 : Institut National Polytechn- ique de Grenoble 전산학과 박사

1976~992 : KIST, KIET, ETRI 선임/책임연구원

1992~현재 : 아주대학교 정보 및 컴퓨터공학부 교수

관심분야 : 임베디드 테스팅, 자가 컴퓨팅/치료 시스템, 차세대 컴퓨터 구조 등

Email : sparky@ajou.ac.kr

신 승 훈
2000 : 아주대학교 정보 및 컴퓨터공학부 공학사
2002 : 아주대학교 정보통신공학과 공학석사
2011 : 아주대학교 정보통신공학과 공학박사
현재 : 아주대학교 정보컴퓨터공학부 특임교원
관심분야 : 소프트웨어 테스트 자동화, 멀티미디어 서비스 정책 등
Email : sihnsh@ajou.ac.kr