

PrivaSense: Providing Privacy Protection for Sensor Networks*

Yi Ouyang, Zhengyi Le
Computer Science Department
Dartmouth College
Hanover, NH

{ouyang, zyle}@cs.dartmouth.edu

James Ford, Fillia Makedon
Computer Science and Engineering Department
University of Texas at Arlington
Arlington, TX

{jcford, makedon}@uta.edu

Categories and Subject Descriptors

C.2.2 [Network Protocols]: Routing protocols

General Terms

Design, Security, Performance.

Keywords

Sensor Networks, Privacy, System Design.

1 Introduction

Sensor networks are used in a variety of applications such as battlefield reconnaissance, environmental monitoring, and traffic monitoring. Security and privacy become important concerns when people are participants in sensor network applications [1]. In this research, we focus on providing a framework in which the privacy of people participating in sensor networks applications is protected.

Sensor networks are highly vulnerable to attacks due to their open wireless medium. An eavesdropper can monitor the communications among sensor nodes and extract sensitive information through traffic analysis; such information includes the location of a message's source and the routing path of messages [2]. Consequently, the identity information of sensor nodes can be compromised as well. The identities of the sensor nodes are highly private information if sensor nodes are carried by people.

We propose a framework called PrivaSense in which fake traffic is generated to disguise real event messages and confuse the attacker, and cryptographic methods are used to mask the real identities of sensor nodes. The framework includes two components: probabilistic message hiding and one-way hash based anonymization.

2 PrivaSense

In this section, we will describe the two components of the PrivaSense framework: Probabilistic Message Hiding and One-way Hash based Anonymization. Using this framework, the privacy of sensor nodes, as well as the people being monitored, is protected against the most powerful class of attacker, one which has access to all the traffic information in the network.

* Supported in part by the National Science Foundation under award Number ITR 0312629 and IDM 0308229.

2.1 Probabilistic message hiding against laptop-class attacker

Assuming attackers can monitor all the communications in the network, existing privacy preserving routing protocols cannot hide the source location. We propose a probabilistic message hiding method to solve this problem. In our method, real event messages are hidden among fake messages. We call these fake messages, which have the same format as real event messages, *maintenance messages*. An attacker cannot distinguish between a maintenance message and an event message; however, continually sending fake messages will deplete the battery power of sensor nodes. Thus, we use two mechanisms to reduce the energy cost:

(1) Every node sends maintenance messages at a pseudo random time interval, which is generated by a Pseudo Random Number Generator (PRNG). Because of the random delays, the number of maintenance messages being sent is less. Since the event messages are disguised as maintenance messages, the time for these real event messages to arrive the base station is also delayed. To minimize the delivery time of event messages, we propose a greedy algorithm, in which every node forwards the event messages to the neighbor that has minimum waiting time. Every node can exploit the predictability of PRNGs of its neighbors and compute their future delays.

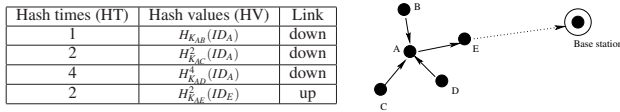
(2) Every node probabilistically sends maintenance messages. To further save energy, we analyze the possibility of letting sensor nodes probabilistically send maintenance messages while preserving the privacy of the source. In a dense sensor network, in which there are many sensor nodes in a small area, maintenance messages sent by fewer nodes can still confuse attackers as long as the attackers can always hear a message at any location in a given period. Based on this observation, we analyze the relation between privacy of the source and the probability of nodes sending maintenance messages. Proper probability value can be selected depending on the density of the sensor network.

2.2 One-way hash based anonymization

In this section, a basic ID randomization method based on one-way keyed hash chain is proposed to protect the anonymity of sensor nodes. The basic idea is for each individual sensor node to use a sequence of hash values as its IDs. The first value of this hash chain will be the original ID of the node; at any time, the sensor node can apply a simple

hash function to its current ID to generate a new ID. Using a simple hash function to evolve a sensor node's ID lets the base station and each sensor node both be aware of the node's current ID without using much storage.

We will describe how a sensor nodes s_i sends a message to the base station through an intermediate node s_j . After determining the routing information and creating pairwise shared keys with its neighbors, a sensor node s_i will create a table R_i that includes the keyed hash values of neighboring nodes' IDs. Table R_i is a set of tuples $(HT, HV = H_{K_{ij}}^{HT}(x), Link)$, where $x = ID_j$ if s_j is s_i 's uplink-node, $x = ID_i$ if s_j is s_i 's downlink-node, $H_K^{HT}(x)$ means using K as a key to hash x for HT times, HT denotes the times of hashes, and $Link$ denotes uplink or downlink. An example of R is shown in Figure 1(a) for node A in Figure 1(b).



(a) An example of table R (b) Example of node A creating table R

Figure 1. Table R of a node

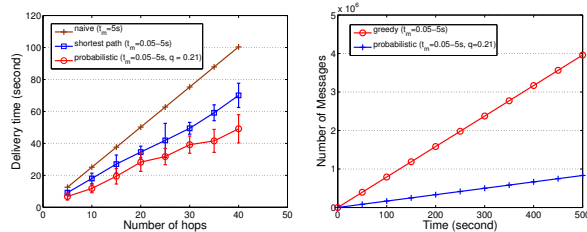
The operations of node s_i are described in Procedure 1. The operations of node s_j forwarding the message are described in Procedure 2.

Procedure 1 Operations of node s_i sending a message

- 1: $t = 1$
 - 2: **while** true **do**
 - 3: s_i collects data D_t .
 - 4: s_i sends $M = H_{K_{ij}}^{HT_j}(ID_j) || H_{K_i}^{HT_i}(ID_i) || t || D_t$ to s_j .
 - 5: $HT_j = HT_j + 1$, $HV_j = H_{K_{ij}}(HV_j)$ (Refresh table R_i for entry s_j)
 - 6: $t = t + 1$
 - 7: **end while**
-

3 Evaluation

We used simulation to evaluate the performance of our algorithms. In the simulation, there are 20,000 sensor nodes uniformly random distributed in a 5,000m by 5,000m area. Figure 2(a) shows a comparison of delivery time between naive, shortest path, and probabilistic algorithms. Figure 2(b) shows a comparison of energy cost between the greedy algorithm and our probabilistic algorithm.



(a) Comparison on delivery time (b) Comparison on energy cost
Figure 2. Performance comparison between naive, shortest path, greedy, and probabilistic algorithm

Procedure 2 Operations of node s_j forwarding a message from s_i

- 1: s_j receives $M = H_{K_{ij}}^{HT_j}(ID_j) || H_{K_i}^{HT_i}(ID_i) || t || D_t$.
 - 2: s_j checks its table R_j to find a match of $H_{K_{ij}}^{HT_j}(ID_j)$ and HT_j .
 - 3: **if** there is a match **then**
 - 4: $HT_i = HT_i + 1$, $HV_i = H_{K_{ij}}(HV_i)$ (s_j refreshes table R_j for entry s_i)
 - 5: s_j checks its table R_j for an uplink s_k
 - 6: s_j changes M to $M' = H_{K_{jk}}^{HT_k}(ID_k) || H_{K_i}^{HT_i}(ID_i) || t || D_t$, and sends M' to uplink-node s_k
 - 7: $HT_k = HT_k + 1$, $HV_k = H_{K_{jk}}(HV_k)$ (s_j refreshes table R_j for entry s_k)
 - 8: **else**
 - 9: s_j discards the message and returns to original state.
 - 10: **end if**
-

Using the one-way hash based anonymization component, when a sensor node sends or forwards a message, the receiver's ID is hidden in the first keyed hash value $H_{K_{ij}}^{HT_j}(ID_j)$. The sender can send a message to one of its neighbor nodes without specifying the identity of this node. When the sensor node needs to send another message, it will use the keyed hash value $H_{K_{ij}}^{HT_j+1}(ID_j)$ to specify the same receiver. The two keyed hash values look random to an adversary. In this way, the adversary cannot determine the identity of both parties in these communications and whether these two messages are between the same sender and receiver. Thus, when forwarding messages, a sensor node's anonymity is protected.

The ID of the original sender included in the messages for each sensor node is also changed with each use by using a keyed hash function that operates on the original ID again and again. Since the base station and each individual sensor node know their shared key and the node's original ID, the base station can evolve IDs in parallel to determine the original ID of the sender of any message. For a third party such as an adversary or an intermediate forwarding sensor node, it is not possible to determine this without knowing the secret key being used. Thus, the anonymity of the original senders of messages is also protected.

4 Conclusions

In the PrivaSense framework, the probabilistic message hiding method can reduce the energy cost while maintaining privacy for sensor nodes in the network. It exploits the fact that pseudo random numbers can be predicted to minimize the delivery time, and a small number of decoy messages can be enough to confuse attackers in a dense sensor network. One-way hash based anonymization provides another layer of protection when the contents of message are compromised and randomized IDs prevent attackers' traffic analysis.

5 References

- [1] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson. People-centric urban sensing. In *Proc. of WICON '06*, page 18, 2006.
- [2] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk. Enhancing source-location privacy in sensor network routing. In *Proc. of ICDCS '05*, pages 599–608, 2005.