Available online at www.sciencedirect.com
ScienceDirect
Computer Communications 30 (2006) 55–65
computer communications
www.elsevier.com/locate/comcom

Resilience of sink filtering scheme in wireless sensor networks

Miao Ma *

Institute for Infocomm Research (I^2R), 21 Heng Mui Keng Terrace, Singapore 119613, Singapore

Received 5 May 2006; received in revised form 20 July 2006; accepted 21 July 2006
Available online 22 August 2006

Abstract

One of severe security threats in wireless sensor network is node compromise. A compromised node can easily inject false data reports on the events that do not occur. The existing approaches in which each forwarding sensor along a path probabilistically filters out injected false data may not be adequate because such protection may break down when more than a threshold number of nodes are compromised. To solve this problem, we present a sink filtering scheme in clusters of heterogeneous sensor networks. In addition to basic sensors, some powerful data gathering sensors termed as cluster heads (CHs) are added. Each aggregation report generated by a CH must carry multiple keyed message authentication codes (MACs); each MAC is generated by a basic sensor that senses the event. The sink node checks the validity of the carried MACs in an aggregation report and filters out the forged report. We analyze the resilience and overhead of the sink filtering scheme. Both analytical and simulation results show that the scheme is resilient to an increasing number of compromised nodes, with graceful performance degradation. Particularly, we adopt Poisson Approximation to investigate the performance tradeoff between resilience and overall cost, and give some suggestions on how to choose the parameters. The scheme is also scalable and efficient in communication, computation and storage.

Keywords: Wireless sensor networks; Security; Node compromise; False data injection; Resiliency

1. Introduction

Sensor networks are expected to play an important role in the upcoming age of pervasive computing, and are being deployed for a wide variety of applications, for example, military sensing and tracking, environment and habitat monitoring, industrial sensing, traffic control, etc. For the mission-critical applications, security is a major concern since in most cases the sensors work in an unattended or hostile environment. A major threat in wireless sensor network is *false data injection attack*, i.e., the compromised sensors forge the events that do not occur. Such type of attack not only leads to false alarm, but also wastes the sensors' energy.

To defend against false data injection attack, Zhu et al. [1] presented an interleaved hop-by-hop authentication (IHA) scheme and Ye et al. [2] proposed a statistical en-route filtering (SEF) mechanism. The methodology of both IHA and SEF schemes is that, each forwarding sensor along a path probabilistically filters out the injected false data. However, for both IHA and SEF schemes, the whole system will break down if an attacker has compromised T nodes. In other words, after compromising T nodes, the attacker can forge events happening at arbitrary locations without being detection.

Recently, Yang et al. [3] proposed a location-based resilient security (LBRS) scheme, which eliminates the threshold breakdown problem (in IHA and SEF schemes) by exploiting a location-based approach as the fundamental mechanism towards resilient security. The location-binding property constraints the scope for which individual keys can be misused, thus limiting the damages caused by a collection of compromised nodes; however, LBRS assumes that once deployed every node can obtain its geographic location via a location scheme. We comment that such an assumption may not always be practical, because the overhead

* Tel.: +65 68743116; fax: +65 67755014.
E-mail address: miaom@i2r.a-star.edu.sg

0140-3664/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2006.07.015

incurred may be huge if every sensor needs to obtain its geographic location.

As an alternative to LBRS scheme, in this paper, we propose a sink filtering scheme in clusters of heterogeneous sensor networks. In addition to basic sensors, some powerful data gathering sensors termed as cluster heads (CHs) are added. Each aggregation report generated by a CH must carry multiple keyed message authentication codes (MACs); each MAC is generated by a basic sensor that senses the event. The sink node checks the validity of the carried MACs in an aggregation report and filters out the forged report. We analyze the resilience and overhead of the scheme. Both analytical and simulation results show that the scheme is resilient to an increasing number of compromised nodes, without threshold breakdown problem. We also adopt Poisson Approximation to investigate the performance tradeoff between resilience and overall cost. Suggestions on how to choose the parameters are also given. In addition, the scheme is scalable and efficient in communication, computation and storage. An earlier version of this paper can be found in [4].

The rest of this paper is organized as follows. In Section 2, we describe the assumptions and the network model. The sink filtering scheme in clusters of heterogeneous sensor networks is presented in Section 3. Section 4 studies the resilience behavior within a cluster for the proposed scheme. Section 5 analyzes the resilience of the scheme and investigates the tradeoff between resilience and overall cost. Section 6 evaluates the overhead of the scheme. Finally, Section 7 concludes the paper.

2. Assumptions and network model

We consider a heterogeneous sensor network, where two types of sensors are deployed: basic sensor and cluster head (CH). A basic sensor is simple, inexpensive and power-limited, while a CH has more capabilities on processing and communication, richer power supply, and is more compromise-resilient.

We consider a target deployment area as a two-dimensional square region with size A^2. The sink is located at the center $(0,0)$, and the network is $\{(x,y)|\ |x| \leqslant \frac{A}{2},\ |y| \leqslant \frac{A}{2}\}$, shown in Fig. 1. The basic sensors are uniformly distributed across the entire deployment area. The sensing range of a basic sensor is r_s; the communication range of a basic sensor is r_c. We define the deployment density as the average number of basic sensors within a basic sensor's sensing range, denoted by n. The total number of basic sensors N is determined by the deployment area A^2, the deployment density n, and the sensing range r_s. That is, $N = [nA^2/\pi r_s^2]$. Every basic sensor has a unique identification (ID).

The deployment area is divided into C equal size grids (i.e., clusters), with each grid's size as a^2. The choice of C has an upper bound, i.e., the area of each cluster at least covers a basic sensor's sensing area. We use C_b to denote the upper bound of C, and we have $C_b = \frac{A^2}{4r_s^2}$. Without loss

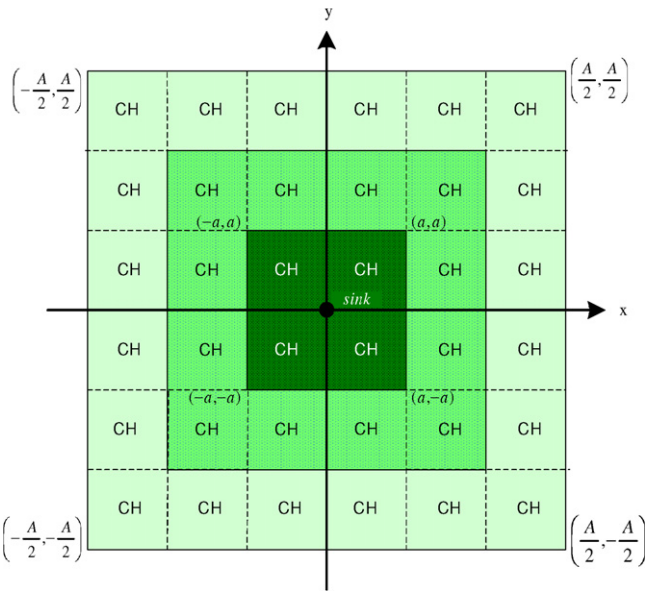

Fig. 1. The deployment area A^2. It is divided into C equal size grids (i.e., clusters), with each grid's size as a^2. We define the clusters within the area $\{(x,y)|a(l-1) \leqslant |x| \leqslant al, a(l-1) \leqslant |y| \leqslant al\}$ as of lth level clusters $(1 \leqslant l \leqslant L)$. In this picture, $C = 36$, $L = 3$.

of generality, we assume that each CH is deployed at the center of each grid. Every CH also has a unique ID. Each basic sensor is assigned to the nearest CH. For simplicity, we assume that there are s (i.e., $\frac{N}{C}$) basic sensors in each cluster. Our numerical results will demonstrate that this assumption has a negligible impact on our analysis and can be moved out. We also assume that these s basic sensors are uniformly distributed within a grid.

In order to prevent a malicious node claiming to be a CH node and confusing others, we differentiate the IDs of the CHs from the IDs of the basic nodes. Our methodology is the same as the approach in [5] which is used to differentiate a mobile station from a regular node. In our scheme, we set the ID for a CH node as a pseudo-random number output from a hash function and 8 bytes long; and set the ID for a basic sensor as an integer between 1 and 65,536 (i.e., assuming that there are at most 65,536 basic sensors in a network) and 2 bytes long. The ID of a CH node will unlikely fall into the range [1–65,536] since the probability that all the other 6 bytes of a CH ID output from a hash function are zeros is negligible.

The key management for the clusters of heterogeneous sensor network is as follows: (1) before deployment, each sensor (either a basic sensor or a CH) shares a secret key with the sink; (2) we assume the neighborhood relationship among CHs is known in advance. Before deployment, each CH simply pre-loads eight pairwise keys with its eight immediate neighboring CHs, respectively. The CHs, therefore, organize themselves into a static ad hoc network. (3) Upon deployment, each basic sensor establishes a pairwise key with its one-hop neighboring basic sensors; the one-hop pairwise key establishment scheme in LEAP [6] is adopted to achieve this goal. We denote the number of