Ad Hoc Networks 7 (2009) 1501–1514

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Privacy preservation in wireless sensor networks: A state-of-the-art survey

Na Li [a,*], Nan Zhang [b], Sajal K. Das [a], Bhavani Thuraisingham [c]

[a] Department of Computer Science and Engineering, The University of Texas at Arlington, Box 19015, 416 Yates St., Room 300, Nedderman Hall, Arlington, TX 76019-0015, United States
[b] Department of Computer Science, The George Washington University, 801 22nd Street NW, Suite 704, Washington DC 20052, United States
[c] Department of Computer Science, Erik Jonsson School of Engineering & Computer Science, The University of Texas at Dallas, 800 W. Campbell Road, MS EC31, Richardson, TX 75080, United States

ARTICLE INFO

Available online 22 April 2009

Keywords:
Wireless sensor network
Privacy

ABSTRACT

Much of the existing work on wireless sensor networks (WSNs) has focused on addressing the power and computational resource constraints of WSNs by the design of specific routing, MAC, and cross-layer protocols. Recently, there have been heightened privacy concerns over the data collected by and transmitted through WSNs. The wireless transmission required by a WSN, and the self-organizing nature of its architecture, makes privacy protection for WSNs an especially challenging problem. This paper provides a state-of-the-art survey of privacy-preserving techniques for WSNs. In particular, we review two main categories of privacy-preserving techniques for protecting two types of private information, data-oriented and context-oriented privacy, respectively. We also discuss a number of important open challenges for future research. Our hope is that this paper sheds some light on a fruitful direction of future research for privacy preservation in WSNs.

1. Introduction

In recent years, wireless sensor networks (WSNs) have drawn considerable attention from the research community on issues ranging from theoretical research to practical applications. Special characteristics of WSNs, such as resource constraints on energy and computational power, have been well defined and widely studied [3]. What has received less attention, however, is the critical *privacy concern* on information being collected, transmitted, and analyzed in a WSN. Such private information of concern may include *payload data* collected by sensors and transmitted through the network to a centralized data processing server. For example, a patient's blood pressure, sugar level and other vital signs are usually of critical privacy concern when monitored by a medical WSN which transmits the data to a remote hospital or doctor's office. Privacy concerns may also arise beyond data content and may focus on *context information* such as the location of a sensor initiating data communication. Note that an alert communication originating from a patient's heart monitor in the medical WSN is enough for an adversary to infer that the patient suffers from heart problem. Effective countermeasure against the disclosure of both *data* and *context-oriented* private information is an indispensable prerequisite for the broad application of WSNs to real-world applications.

Privacy protection has been extensively studied in various fields related to WSN such as wired and wireless networking, databases and data mining. Nonetheless, the following inherent features of WSNs introduce unique challenges for privacy preservation in WSNs, and prevent the existing techniques from being directly transplanted:

- *Uncontrollable environment*: Sensors may have to be deployed to an environment uncontrollable by the defender, such as a battlefield, enabling an adversary to launch physical attacks to capture sensor nodes or deploy counterfeit ones. As a result, an adversary may retrieve private

[*] Corresponding author. Tel.: +1 8172727409.
E-mail addresses: na.li@mavs.uta.edu (N. Li), nzhang10@gwu.edu (N. Zhang), das@exchange.uta.edu (S.K. Das), bhavani.thuraisingham@utdallas.edu (B. Thuraisingham).

1570-8705/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2009.04.009

keys used for secure communication and decrypt any communication eavesdropped by the adversary.

- *Sensor-node resource constraints*: A battery-powered sensor node generally has severe constraints on its ability to store, process, and transmit the sensed data. As a result, the computational complexity and resource consumption of public-key ciphers is usually considered unsuitable for WSNs. This introduces additional challenges for privacy preservation.
- *Topological constraints*: The limited communication range of sensor nodes in a WSN requires multiple hops in order to transmit data from the source to the base station. Such a multi-hop scheme demands different nodes to take diverse traffic loads. In particular, a node closer to the base station (i.e., data collecting and processing server) has to relay data from nodes further away from base station in addition to transmitting its own generated data, leading to higher transmission rate. Such an unbalanced network traffic pattern brings significant challenges to the protection of context-oriented privacy information. Particularly, if an adversary holds the ability of global traffic analysis, observing the traffic patterns of different nodes over the whole network, it can easily identify the sink and compromise context privacy, or even manipulate the sink node to impede the proper functioning of the WSN.

The unique challenges for privacy preservation in WSNs call for the development of effective privacy-preserving techniques. In this paper, we provide a state-of-the-art survey of existing privacy-preserving techniques in WSNs. We review two main categories of privacy-preserving techniques for protecting two types of private information, data-oriented and context-oriented privacy, respectively. In the category of data privacy, we mainly discuss how to enable the aggregation of sensed data without violating the privacy of the data being collected and guarantee the privacy of data query initiated by users of the network. For context-based privacy, we analyze the protection of location privacy and temporal private information. In addition, we build a table to compare different techniques in terms of their effectiveness in practical applications. Last but not the least, we discuss some interesting and challenging open issues on this topic, which are expected to shed light on a fruitful direction of future research on privacy preservation in WSNs.

The rest of the paper is organized as follows. We review privacy-preserving techniques in related fields in Section 2. In Section 3, we introduce our taxonomy of privacy-preserving techniques in WSNs. Sections 4 and 5 address techniques for data and context-oriented privacy protection, respectively. In Section 6, we evaluate and compare the performance of different privacy-preserving techniques. Section 7 outlines the open challenges for future research, followed by final remarks in Section 8.

2. Related work

Research on issues related to WSNs requires multi-disciplinary studies spanning networking, databases, distributed computing, etc. To properly understand the challenges of privacy preservation in WSNs and the techniques necessary to address such challenges, it is important to first examine the privacy issues and privacy-preserving techniques in such related fields as databases, data mining and wireless networks, which we briefly review as follows.

In the field of database and data mining, privacy concerns may arise from three types of systems [36]: The first is an information sharing system which involves two or more mutually untrusted parties. The objective is to guarantee that no private information beyond the minimum necessary is disclosed during information sharing. Cryptographic secure multi-party computation techniques are usually used for this type of systems [1,11,35,38]. The second is a data collection system where one centralized data collector/analyzer collects and mines data from multiple distributed data providers. Random perturbation techniques [2,15,33,34] are usually applied to protecting privacy in these systems. The third type is a data publishing system, the objective of which is to publish data to support data analytical application without compromising the anonymity of individual data owners. *k*-anonymity [27] and *l*-diversity based algorithms [19,20] are proposed for privacy protection in these systems.

Privacy issues have also been extensively studied in the domain of generic networking. Location privacy is of particular concern with the pervasive development of advanced wireless device, like PDA, and with the advent of location-based service (LBS). In an LBS system, a user holding a wireless device queries the LBS server to obtain the nearest restaurant or hospital to the user. Nonetheless, the user would not willingly disclose his/her real location. To address such location privacy concerns, ANONYMIZER, a trusted-third-party based framework, was proposed [13]. With this framework, a user first sends his/her location to the centralized anonymizer which then queries the LBS server with not the user's real location but a cloaking region which covers not only the user but also a number of other users. This technique prevents the LBS server from distinguishing one user from many others. However, it is unlikely to be practical for two reasons: First, it is not reasonable to assume the existence of a trustable third party. Second, even if such a trusted third party exists, it creates a single point-of-failure for the system because if the third party is compromised, the privacy preservation over the whole system will completely collapse. To remove the requirement of a trusted third party, a private-information-retrieval based technique was proposed [12]. Nonetheless, this technique also suffers from significant computation and communication overhead. Besides the direct disclosure of user's location from query payload, traffic flow information may also breach location privacy. In particular, the server may pinpoint the location of a user based on the user's IP address. In order to provide traffic flow confidentiality, Tor [10] is designed, as the second generation onion routing [22], and becomes a popular anonymous communication network which consists of thousands of Tor routers to relay user traffic to or from LBS server.