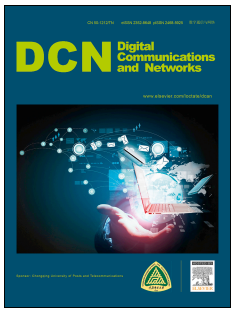
Accepted Manuscript



Achieving privacy-preserving big data aggregation with fault tolerance in smart grid

Zhitao Guan, Guanlin Si

PII: S2352-8648(17)30066-4

DOI: 10.1016/j.dcan.2017.08.005

Reference: DCAN 99

To appear in: *Digital Communications and Networks*

Received Date: 15 February 2017

Revised Date: 3 August 2017

Accepted Date: 10 August 2017

Please cite this article as: Z. Guan, G. Si, Achieving privacy-preserving big data aggregation with fault tolerance in smart grid, *Digital Communications and Networks* (2017), doi: 10.1016/j.dcan.2017.08.005.

Achieving Privacy-Preserving Big Data Aggregation with Fault Tolerance in Smart Grid

Zhitao Guan, Guanlin Si

School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206 China

*Zhitao Guan (Corresponding author) and Guanlin Si are with the School of Control and Computer Engineering, North China Electric Power University, Beijing, China. E-mail: guanzhitao@126.com, m18811612766@163.com.

Abstract

In smart grid, a huge amount of data are collected for various applications, such as load monitoring and demand response. These data are used for analyzing the power state and creating optimal dispatching strategy. However, these big energy data in terms of volume, velocity and variety raise consumer's privacy concerns. For instance, in order to optimize the energy utilization and support the demand response, numerous smart meters are installed at consumer's home to collect energy consumption data at a fine granularity, but these fine-grained data may disclose the appliances consumption patterns and then discover consumer's behaviors at home. In this paper, we propose a privacy-preserving data aggregation scheme based on secret sharing with fault tolerance in smart grid, which ensures that control center gets the integrated data without compromising user's privacy. Meanwhile, we also consider fault tolerance and the resistance to differential attack during the data aggregation. At last, we analyze the security analysis and performance evaluation of our scheme compared with the other similar schemes. The analysis shows that our scheme can meet the security requirement, and it also has better performance than that of the other popular methods.

Keywords: Big data, Smart grid, Privacy-preserving, Fault tolerance

1. Introduction

Fig. 1: System Model

As a new generation of energy network, smart grid is considered a useful way to solve the severe environmental issues and resource crisis. It is the product of the combination of energy network and information technology. Differing from the unidirectional centralized grid, the control mode of the smart grid is more flexible and reliable. It supports bidirectional power flow between the users and grid. User in smart grid is not only a consumer but also a generator. In smart grid, large quantities of data are collected to support basic services [1]. For example, to create power plan or dynamic price, control center needs to collect and analyze real-time data from various applications by adopting the smart meter installed at the user's house. What's more, electric vehicle drivers need to upload their location message to control center for power dispatching.

Although big data collected from users is necessary for the basic service, it is usually sensitive to user's privacy [2]. For instance, smart meters are adopted to collect the real-time data from users to control center, but these data may disclose user's family behaviors. Thus, if a thief gets the real-time data from user's smart meter, he may gain entry to user's house when he notices that there is nobody home. Besides, user's location-privacy may be disclosed during the interaction between electric vehicles and smart grid, which may help the adversary to catch user's track [3] [4]. If user's sensitive data isn't preserved very well, the implement of smart grid will meet resistance. Therefore, the privacy-leaking in smart grid becomes an extremely important problem.

For the privacy-preservation in smart grid, there are various solutions. As we know, traditional privacy-preserving strategies can be divided into two aspects. One is to hide the user's identity; the other one is to protect the user's sensitive data [5]. As the properties of big data in smart grid are reflected by volume, velocity and variety [6], approaches for privacy-preservation need to consider more on communication overhead and computational cost [7] [8].