Accepted Manuscript

Privacy-preserving deduplication of encrypted data with dynamic ownership management in fog computing

Dongyoung Koo, Junbeom Hur

PII: S0167-739X(17)30130-9
DOI: http://dx.doi.org/10.1016/j.future.2017.01.024
Reference: FUTURE 3306

To appear in: *Future Generation Computer Systems*

Received date: 16 May 2016
Revised date: 27 December 2016
Accepted date: 21 January 2017

Please cite this article as: D. Koo, J. Hur, Privacy-preserving deduplication of encrypted data with dynamic ownership management in fog computing, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.01.024

Privacy-preserving deduplication of encrypted data with dynamic ownership management in fog computing

Dongyoung Koo[a], Junbeom Hur[a,*]

[a]*Department of Computer Science and Engineering, College of Informatics, Korea University, Seoul, South Korea*

Abstract

The explosion in the volume of data generated by end-point devices, arising from IoT proliferation, has lead to the adoption of data outsourcing to dedicated data centers. However, centralized data centers such as cloud storage cannot afford to manage large stores of data in a timely manner. To allow low latency access to large amounts of data, a new computing paradigm, called fog computing, has been introduced. In a fog computing environment, privacy issues surrounding outsourced data become more critical due to its complicated innards of the system. In addition, efficient resource management is another important criterion considering the application of pay-per-use in commercial fog storage. As an extension of cloud storage, most fog storage service providers will choose to adopt data deduplication techniques to minimize resource dissipation. At the same time, data owners may update or remove outsourced data stored in the remote storage to reduce expenses. In this paper, we propose the first privacy-preserving deduplication protocol capable of efficient ownership management in fog computing. It achieves fine-grained access control by introducing user-level key management and update mechanisms. Data-invariant user-level private keys enable data owners to maintain a constant number of keys regardless of the number of outsourced data files. The update of user-level public keys for valid data owners at the remote storage dramatically reduces communication overhead. Security and performance analyses demonstrate the efficiency of the proposed scheme in terms of communication and key management in fog storage.

Keywords: Data deduplication, fog computing, data privacy, data ownership management, efficiency

1. Introduction

Fog computing, an evolutionary framework of future generation computing, is a combination of the Internet of Things (IoT) and cloud computing. Due to the rise of IoT devices with limited computing resources, cloud-based solutions have been extensively researched. However, forecasts based on the recent growth of the IoT market [1, 2] indicate that centralized clouds will be unlikely to be able to provide satisfactory services to users in the near future. As an extension of the cloud computing paradigm from the core to the edge of the network, fog computing can be seen as a layered structure of services [3–5]. While the central cloud provides a wide range of computing services, it also manages decentralized heterogeneous fog devices. Individual fog devices located near IoT devices provide faster cloud services to end users based on their own computation, storage, and network capabilities. Therefore, fog computing is a promising next-generation computing paradigm, with three attractive attributes: (1) low latency, (2) enhanced user experience (*i.e.*, high quality service), and (3) context awareness based on locational proximity [6, 7].

Centralized cloud storage is unable to handle enormous volumes of data in a timely manner given a finite network bandwidth. Distributed storage, namely fog device, is incapable of providing permanent and global computing services to users owing to its limited resources and restricted field of vision, respectively. Therefore, efficient resource management (especially that of storage space and network bandwidth) can be seen as one of the most important goals of commercial online storage services. Under the fog computing environment, utilization of these resources in a harmonious way between the central cloud and fog devices would be one of the most desirable approaches. Deduplication is able to utilize space efficiently by storing only a single copy of duplicate data and providing owners with a link to it. Compared to cloud storage, fog devices located near the premises of end users with temporal storage can provide a faster data outsourcing service to data owners. At the same time, the central cloud can efficiently utilize storage space by receiving and maintaining only unique data from fog devices.

Despite the compelling benefits of deduplication, privacy issues surrounding outsourced data have also received close attention. Because data owners lose control of their

*Corresponding author: Junbeom Hur, Information System Security (ISS) Laboratory, Department of Computer Science and Engineering, College of Informatics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, South Korea. Tel.: +82 2 3290 4603.

Email addresses: dykoo@korea.ac.kr (Dongyoung Koo), jbhur@korea.ac.kr (Junbeom Hur)