Chapter 9

Secure and Privacy Preserving Keyword Search over the Large Scale Cloud Data

Wei Zhang
Hunan University, China

Jie Wu
Temple University, USA

Yaping Lin
Hunan University, China

ABSTRACT

Cloud computing has attracted a lot of interests from both the academics and the industries, since it provides efficient resource management, economical cost, and fast deployment. However, concerns on security and privacy become the main obstacle for the large scale application of cloud computing. Encryption would be an alternative way to relief the concern. However, data encryption makes efficient data utilization a challenging problem. To address this problem, secure and privacy preserving keyword search over large scale cloud data is proposed and widely developed. In this paper, we make a thorough survey on the secure and privacy preserving keyword search over large scale cloud data. We investigate existing research arts category by category, where the category is classified according to the search functionality. In each category, we first elaborate on the key idea of existing research works, then we conclude some open and interesting problems.

INTRODUCTION

Cloud computing has attracted a lot of interests from both the academics and the industries, since it provides efficient resource management, economical cost, and fast deployment (Armbrust, et al, 2010). For both personal and enterprise users, the cloud computing creates a lot of opportunities for them to enjoy innovation, collaboration, and convenience. Due to the huge potential economical benefits of cloud computing, a lot of companies have deployed their cloud centers. For example, the Elastic Compute Cloud (EC2) of Amazon, the App Engine of Google, the Azure of Microsoft, and Blue Cloud of IBM.

DOI: 10.4018/978-1-5225-0105-3.ch009
Although the cloud computing has a lot of benefits, both individual and enterprise users are not willing to outsource their sensitive data (e.g., personal health records, financial records) to the cloud server. Because once these data are outsourced to the cloud server, the corresponding data owners will lose direct control over these data. The Cloud Service Provider (CSP) would promise that they can preserve the security of these data by using techniques like fireware, virtualization, and Intrusion Detection System (IDS). However, since the CSP takes full control of these data, these techniques cannot prevent employers of the CSP from revealing sensitive data. Encryption would be an alternative way to solve the problem. However, data encryption makes the traditional plaintext based search schemes impractical. A probable solution is downloading all these encrypted files and decrypting them locally to find the desired files. However, this is obviously unrealistic, since it would cause unacceptable communication and computation cost for the end users, whose communication and computation capabilities are often constrained. Therefore, devising a secure and privacy search scheme over encrypted cloud data would be grateful.

To address this problem, secure and privacy preserving keyword search over large scale cloud data is proposed and widely developed. A secure search system often includes three entities, i.e., data owner, cloud server, and data users. The data owner outsources encrypted files and indexes to the cloud server. Authorized data users generate secret trapdoors and submit them to the cloud server. The cloud server further returns the search results without knowing sensitive data. We can categorize existing secure search schemes based on different criterions.

First, based on the search functionality, existing researches include: secure conjunctive keyword search, secure ranked keyword search, secure fuzzy keyword search, and privacy preserving similarity keyword search.

Second, based on the adopted encryption method, these researches can be categorized into: symmetric encryption based schemes and asymmetric encryption based schemes. The symmetric encryption based schemes often achieve high efficiency while the asymmetric encryption based schemes achieve strong security.

Third, based on the threat model, existing research works consider two different models, one assumes the cloud server to be “curious but honest”, i.e., the cloud server will follow the proposed schemes, but they will try to reveal the sensitive data of both the data owner and data users. The other one assumes the cloud server to be “dishonest” (Zhang, et al, 2015), i.e., the cloud server would probably return false retrieval result, therefore, the corresponding research works seek to verify the retrieval results.

Forth, based on the number of data owner involved in the search system, existing research works are divided into single owner model and multi-owner model. In comparison, the multi-owner model would be more adapted to be deployed in reality.

Fifth, based on the number of cloud server, there are two kinds of research works, i.e., secure keyword search on the centralized cloud server, and secure keyword search among distributed cloud servers.

Finally, based on the design goals, different researches have different concerns: preserving data confidentiality and data privacy, achieving low computation and communication cost, hiding data access pattern (preventing the cloud server from knowing which data are actually returned), enabling scalability, and allowing data updating.

In this chapter, we make a thorough survey on secure and privacy preserving keyword search over large scale cloud data. We investigate research works category by category, where the category is classified according to the search functionality. In each category, we first elaborate on the key idea of existing research arts, then we conclude some open and interesting problems.

www.igi-global.com/e-resources/library-recommendation/?id=96

Related Content

Programming the Blockchain
(2017). Decentralized Computing Using Blockchain Technologies and Smart Contracts: Emerging Research and Opportunities (pp. 64-71).
www.igi-global.com/chapter/programming-the-blockchain/176869?camid=4v1a

Smart Contracts and Smart Properties Over Blockchains
(2017). Decentralized Computing Using Blockchain Technologies and Smart Contracts: Emerging Research and Opportunities (pp. 44-63).
www.igi-global.com/chapter/smart-contracts-and-smart-properties-over-blockchains/176868?camid=4v1a

Advanced Topics in Blockchains
(2017). Decentralized Computing Using Blockchain Technologies and Smart Contracts: Emerging Research and Opportunities (pp. 28-43).
www.igi-global.com/chapter/advanced-topics-in-blockchains/176867?camid=4v1a

Addressing Security Issues of the Internet of Things Using Physically Unclonable Functions
www.igi-global.com/chapter/addressing-security-issues-of-the-internet-of-things-using-physically-unclonable-functions/222273?camid=4v1a