Information Systems 48 (2015) 196–212

Contents lists available at ScienceDirect

Information Systems

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

A secure data sharing and query processing framework via federation of cloud computing

Bharath K. Samanthula, Yousef Elmehdwi, Gerry Howser, Sanjay Madria*

Department of Computer Science, Missouri University of Science and Technology, 500 West 15th Street, Rolla, MO 65401, United States

ARTICLE INFO

Available online 31 August 2013

Keywords:
Privacy
Cloud computing
Homomorphic encryption
Proxy re-encryption

ABSTRACT

Due to cost-efficiency and less hands-on management, data owners are outsourcing their data to the cloud which can provide access to the data as a service. However, by outsourcing their data to the cloud, the data owners lose control over their data as the cloud provider becomes a third party service provider. At first, encrypting the data by the owner and then exporting it to the cloud seems to be a good approach. However, there is a potential efficiency problem with the outsourced encrypted data when the data owner revokes some of the users' access privileges. An existing solution to this problem is based on symmetric key encryption scheme but it is not secure when a revoked user rejoins the system with different access privileges to the same data record. In this paper, we propose an efficient and Secure Data Sharing (SDS) framework using homomorphic encryption and proxy re-encryption schemes that prevents the leakage of unauthorized data when a revoked user rejoins the system. We also modify our underlying SDS framework and present a new solution based on the data distribution technique to prevent the information leakage in the case of collusion between a revoked user and the cloud service provider. A comparison of the proposed solution with existing methods is provided in detail. Furthermore, we demonstrate how the existing work can be utilized in our proposed framework to support secure query processing. We provide a detailed security as well as experimental analysis of the proposed framework on Amazon EC2 and highlight its practical value.

1. Introduction

Cloud computing [41,42,4,64] is a mean by which highly scalable and technology enabled services can be easily consumed over the Internet on an as-needed-basis. This innovative paradigm has generated a significant interest in both the marketplace and the academic world, resulting in a number of notable commercial and individual cloud computing services, e.g., from Amazon, Google, Microsoft, Yahoo, and Salesforce. Top database vendors such as IBM and Oracle are adding cloud support to their databases.

Cloud computing is clearly one of today's most enticing technologies, at least in part due to its cost-efficiency and flexibility. Several security issues in the cloud [50] are impeding the vision of cloud computing as a new IT procurement model. Security concerns preventing companies from taking advantage of the cloud can be categorized into three categories [11]:

- *Traditional security*: It involves concerns related to computer and network intrusions. Some of these attacks include VM-level attacks [43], cloud provider vulnerabilities [54], phishing cloud provider such as the Salesforce phishing incident [45], authentication and authorization [58], and forensics in the cloud [34].

* Corresponding author. Tel.: +1 573 341 4856.
E-mail addresses: bspq8@mst.edu (B.K. Samanthula), ymez76@mst.edu (Y. Elmehdwi), gwhrkb@mst.edu (G. Howser), madrias@mst.edu (S. Madria).

0306-4379/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.is.2013.08.004

Cloud providers respond to these concerns by arguing that their security measures and processes are more mature and tested than those of the smaller company. Nevertheless, data owners have their own privacy concerns; therefore, they may not fully trust the cloud provider.

- *Availability*: It involves concerns centering on critical applications and data availability [52]. Server uptime issues, single points of failure and attack, and the inability of an enterprise to insure that the cloud provider is faithfully running a hosted application and giving the valid results make companies nervous. A real-life incident for cloud outage is the seven-hour downtime of Amazon S3 on July 20, 2008 [17].
- *Third-party data control*: For the data owner who outsources data to the cloud, the cloud acts as a semi-trusted third-party. Data placed in the cloud can reside anywhere. Outsourcing data to the cloud should not lead to *de facto* outsourcing of data control to the cloud [11].

As mentioned above, one of the main security concerns in cloud computing is how the data is being used by a third party Cloud Service Provider (CSP). On one hand, the legal implications of the data (especially in cloud computing data location matters as the laws governing the data differ across geographic boundaries) and applications being held by a third party are complex and are not well understood [43,50]. On the other hand, there is also a potential lack of control and transparency when a third party holds the data. Some of the resulting security concerns include due diligence, auditability, contractual obligations, cloud provider espionage, data lock-in, and the transitive nature of the data control [11]. Furthermore, data owners may not have sufficient information about the underlying data privacy and security mechanisms used by the CSP in order to determine how well their data is protected on the cloud. An interesting question is, can the data owners ensure the security of their own data by some means? Along this direction, trusted computing and applied cryptographic techniques may offer new tools to solve these problems [46,2].

In general, encryption is a useful tool for protecting the confidentiality of sensitive data so that even if a database is compromised by an intruder, the data remains protected even after the database has been successfully attacked or stolen. Provided that the encryption is done properly and the decryption keys are not also accessible to the attacker, encryption can provide protection for the sensitive data stored in the database, reduce the legal liability of the data owners, and reduce the cost to society of fraud and identity theft. However, with the data being in encrypted form, there remain issues such as preventing user access to unauthorized fields, efficiently revoking users' privileges without re-encrypting massive amounts of data and re-distributing the new keys to the authorized users, handling collusion between users and CSP, and issuing changes to a user's access privileges. In addition, various issues related to secure query processing, which is a common task in Database-as-a-Service (DBaaS) applications, over encrypted data in the cloud need to be addressed [1].

In this paper, we investigate efficient methods for handling user access rights, revoking those rights efficiently, and issuing either same or different access rights to a returning user. We also address the issues of security from a "curious" cloud, collusion among users, and collusion between a user and the Cloud Service Provider. Furthermore, we demonstrate how the existing work on query processing over encrypted data can be extended to our framework in a privacy-preserving manner. In addition, we also report results of experimental evaluation using EC2 cloud of our proposed scheme using different performance metrics.

1.1. Problem statement

In our problem setting, refer to Fig. 1, we include three different players: the data owner (also referred to as Alice), the cloud, and a set of data consumers (also referred to as authorized users). The data owner Alice initially encrypts her data locally to ensure privacy, outsources the data and stores it on the cloud for easy user access. To facilitate a fine-grained access control, a set of attributes is associated with each data record which helps to control user access to a specific set of data fields for each authorized user, e.g., Bob. The data owner Alice then issues a decryption key for each authorized user according to his/her access rights. More details are provided in Section 4.

In this paper, we propose a scheme to achieve fine-grained data sharing and access control over the outsourced data in the cloud. That is, if Bob is an authorized (which is decided by Alice) to access a set of attributes in the data record d, then he can access only those set of attributes on an on-demand basis from the cloud. However, Bob cannot access d if he is not an authorized user.

In contrast to the scheme proposed by Yang [59], which is based on attribute-based encryption and proxy re-encryption, we propose a new Secure Data Sharing (SDS) framework using homomorphic encryption and proxy re-encryption as the underlying sub-routines. We emphasize that there are potential issues in the similar fine-grained

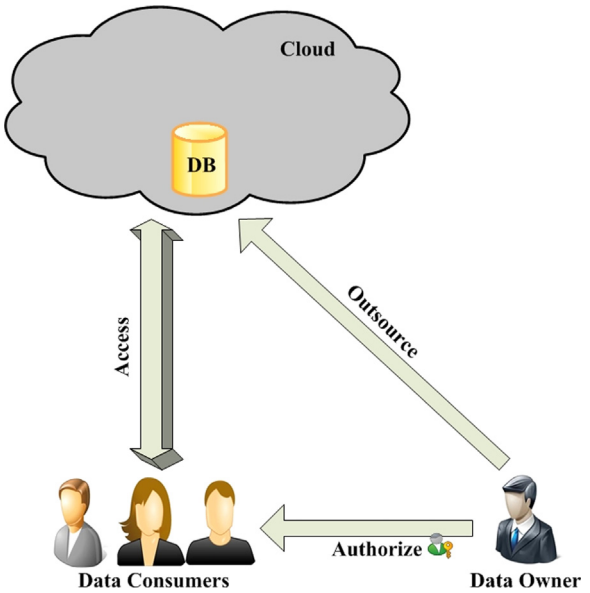

Fig. 1. System model.