COMPUTERS & SECURITY 42 (2014) 151–164

Available online at www.sciencedirect.com

ScienceDirect

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

Computers & Security

Achieving an effective, scalable and privacy-preserving data sharing service in cloud computing

Xin Dong [a], Jiadi Yu [a,*], Yuan Luo [a], Yingying Chen [b], Guangtao Xue [a], Minglu Li [a]

[a] Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, PR China
[b] Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA

ARTICLE INFO

Article history:
Received 7 May 2013
Received in revised form
30 October 2013
Accepted 18 December 2013

Keywords:
Data security
Data sharing
Privacy-preserving
Attribute-based encryption
Scalability
Cloud computing

ABSTRACT

Data sharing in the cloud, fueled by favorable trends in cloud technology, is emerging as a promising technique for allowing users to conveniently access data. However, the growing number of enterprises and customers who stores their data in cloud servers is increasingly challenging users' privacy and the security of data. This paper focuses on providing a dependable and secure cloud data sharing service that allows users dynamic access to their data. In order to achieve this, we propose an effective, scalable and flexible privacy-preserving data policy with semantic security, by utilizing ciphertext policy attribute-based encryption (CP-ABE) combined with identity-based encryption (IBE) techniques. In addition to ensuring robust data sharing security, our policy succeeds in preserving the privacy of cloud users and supports efficient and secure dynamic operations including, but not limited to, file creation, user revocation and modification of user attributes. Security analysis indicates that the proposed policy is secure under the generic bilinear group model in the random oracle model and enforces fine-grained access control, full collusion resistance and backward secrecy. Furthermore, performance analysis and experimental results show that the overheads are as light as possible.

1. Introduction

Cloud computing (Armbrust et al., 2009) is currently emerging as a technology in which cloud service providers (CSP) offer efficient data storage and computing facilities to a global client base. The only requirement for a user is a connected terminal. By employing a combination of virtualization techniques, service-oriented computing and other emerging technologies, cloud computing can be categorized into three types of "X as a service (XaaS)" pay-as-you-go services: the Platform as a Service (PaaS) model, e.g. Microsoft Azure (Mic), where users can deploy their own applications and tools to the cloud; Infrastructure as a Service (IaaS), e.g. Amazon EC2 (Ama), where users can utilize cloud services provided by the CSP to deploy arbitrary software; and Software as a Service (SaaS), e.g. Google App Engine (Goo), where users use applications provided by the CSP that run on the cloud infrastructure.

Storing data in the cloud offers users the convenience of access without requiring direct knowledge of the deployment and management of the hardware or infrastructure. Although cloud computing is much more powerful than personal

* Corresponding author. Tel./fax: +86 21 3420 5856.
E-mail address: jiadiyu@sjtu.edu.cn (J. Yu).
0167-4048/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cose.2013.12.002

computing, it brings new privacy and security challenges, as users relinquish control by outsourcing their data they no longer having physical possession of it. By having full access to cloud services, users' data are exposed to a variety of threats and malicious attacks and cases of security breaches occur frequently (Arrington, 2006). For example, some clouds may be unfaithful to data confidentiality for monetary reasons; confidential information may be disclosed to business competitors; or the CSP may conceal data loss to maintain their reputation (Shah et al., 2007). In summary, although cloud computing is economically attractive to consumers and enterprises by offering users large-scale data sharing, it does not guarantee users privacy and data security.

Data owners demand high levels of security and confidentiality when they outsource their data to a cloud; although they usually encrypt their data when storing it in a cloud server, they still want control over it, for example, if they frequently update it (Erway et al., 2009; Ateniese et al., 2008). Direct employment of traditional cryptographic primitives cannot achieve the data security required. Thus, a considerable amount of work has recently been directed towards ensuring the privacy and security of remotely stored shared data using a variety of systems and security models (Yu et al., 2010a; Wang et al., 2010). These have mainly focused on preserving users' privacy while realizing desired security goals, without introducing excessively high levels of complexity to the users at the decryption stage. To solve these issues, researchers have either utilized key-policy attribute-based encryption (KP-ABE) for secure access control or employed hierarchical identity-based encryption (HIBE) for data security. Yu et al. (2010a) were the first team to achieve secure data access control with provable security in cloud computing using KP-ABE. However, by revealing some of the users' attributes to cloud, these systems were unable to fully preserve users' privacy. Conversely, the HIBE-based scheme (Wang et al., 2010) utilizes hierarchical encryption to ensure data security in a cloud, but this introduces too many private keys for each user to be managed efficiently. In summary, these schemes either have privacy flaws or provide security at the expense of performance; therefore, the challenge of achieving the dual goals of privacy-preserving with effective cloud data sharing remains unresolved.

To realize an effective, scalable and privacy-preserving data sharing service in cloud computing, the following challenges need to be met: firstly, data owners should be able to assign other cloud users with different access privileges to their data; secondly, the cloud needs to be able to support dynamic requests so that data owners can add or revoke access privileges to other users allowing them to create or delete their data; thirdly, the users' privacy must be protected against the cloud so that they can conceal their private information while accessing the cloud; finally, users should be able to access shared data in the cloud through connected technologies with low computing ability, such as smartphones and tablets. To date, solving these important areas in cloud computing remains elusive.

In this paper, we propose an effective, scalable and flexible privacy-preserving data sharing scheme in the cloud, that ensures both semantic security and effective availability of user data. To preserve privacy and guarantee data confidentiality against the cloud, the scheme employs a cryptographic primitive, named cipher-text policy attribute-based encryption (CP-ABE) and combines it with an identity-based encryption (IBE) technique; each data file is described by a set of meaningful attributes, allowing each user to be assigned an access structure that defines the scope of data files they can have access to. To enforce these access structures, this scheme defines a public-private key pair for each attribute. For each user' secret key, it is a combination of user's ID (i.e., user's public key) and the attribute's secret key, thereby ensuring that each attribute presents a different key to each user. Data files are encrypted by public key components and access matrices converted from the access structure; user secret keys are defined to reflect their access privileges so that a user can only decrypt a ciphertext if they have the matched attributes to satisfy the ciphertext. To resolve the challenging issues of collusion resistance, our scheme provides users with a public key fitted to their secret keys; we use user's ID (public key) to "tie" together the attributes belonging to this user so that they cannot be successfully combined with another's user's attributes. To protect user privacy, our scheme does not need to update user secret key so that it prevents cloud access user access structure. To reduce the key management issue, the data owner simply assigns secret keys to users via the cloud.

Compared to previous schemes, our proposed scheme provides the benefits of security and efficiency: 1) the cloud can learn nothing about a user's privacy or access structure, as such the scheme is fully collusion resistant; 2) all extended operations, including user revocation, can only affect the current file or user without involving key updates. Therefore, the main contributions of this paper can be summarized as follows:

1. Our scheme proposes effective, scalable encryption for a cloud data sharing service that simultaneously achieves full privacy-preserving, collusion resistance and data confidentiality.
2. We prove that the proposed scheme provides semantic security for data sharing in cloud computing through the random oracle under the generic bilinear group model (Boneh et al., 2005). Furthermore, our scheme simultaneously enforces fine-grandness, backward secrecy and access privilege confidentiality.
3. The performance analysis indicates that our scheme only incurs a small overhead compared to existing schemes; meanwhile, the experimental results demonstrate that the overheads are as light as possible.

The remainder of this paper is organized as follows: Section 2 discusses related works; Section 3 introduces the system model, adversary model, security requirements and our design goal; Section 4 provides the details of our scheme; Section 5 shows how our scheme can support file creation/deletion, user addition/revocation and modification of user attributes; Sections 6 and 7 analyze the security and performance of our scheme, respectively; finally, Section 8 provides the concluding remarks of the paper.

2. Related work

The concept of identity-based encryption (IBE) was proposed by Shamir (1985); however, a full IBE scheme was not