Monitoring the integrity of Dynamic Data stored in Cloud Computing

Muneshwara M.S¹, Arvind Tejas Chandral²

¹, ² Department of CSE, S J C Institute of Technology, Chickaballapur-562101, Karnataka, India.

Abstract

Cloud computing is an internet based computing which enables sharing of services. Many users place their data in the cloud, so correctness of data and security is a prime concern. This work studies the problem of ensuring the integrity and security of data storage in Cloud Computing. Security in cloud.

To ensure the correctness of data, we consider the task of allowing a third party auditor (TPA), on behalf of the Cloud client[1], to monitor the integrity of the data stored in the cloud by utilizing public key and private key based.

There is a probable risk while exchanging confidential data in the cloud computing environment. To preserve data integrity and confidentiality existing solutions adopt cryptographic methods and disclose data decryption keys only to authorized users. These solutions inevitably introduce a processing overhead on the data owner for key distribution and data management. These issues are addressed in this work by defining and enforcing access policies based on data attributes, and allowing the data owner to delegate most of the computation tasks involved in data access control to cloud servers [6] without affecting the integrity of data. The proposed work is accomplished by exploiting and uniquely combining some techniques of attribute-based encryption (ABE) [14] and proxy re-encryption (PRE) [16]. This proposed scheme is implementing for simplifying the development efforts and provides direct support for security.

Keywords- Attribute-based encryption, Cloud Computing, Data dynamics, Proxy re-encryption, Monitoring, SaaS, TPA

I. INTRODUCTION

Cloud computing is a general term for anything that involves delivering hosted services over the internet. These services are broadly divided into three categories: Infrastructure-as-a-Service(IaaS), Platform-as-a-Service(PaaS) and Software-as-a-Service(SaaS). A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour; it is elastic - a user can have as much or as little of a service as they want at any given time and the service is fully managed by the cloud service provider(the consumer needs nothing but a personal computer and Internet access). The advantage of cloud is cost savings. The prime disadvantage is security. Cloud computing is used by many software industries nowadays. Since the security is not provided in cloud, many companies adopt their unique security structure [18]. For eg) Amazon has its own security structure. Introducing a new and uniform security structure for all types of cloud is the problem we are going to tackle in this paper. Since the data placed in the cloud is accessible to everyone, security is not guaranteed. To ensure security, cryptographic techniques cannot be directly adopted. Sometimes the cloud service provider may hide the data corruptions to maintain the reputation. To avoid this problem, we introduce an effective third party auditor to monitor the user’s outsourced data when needed.

TPA is the third party auditor who will monitor the data of data owner or client. TPA eliminates the involvement of the client through the auditing of whether his data stored in the cloud are indeed intact. The released audit report helps the owner or client to evaluate the risk of their subscribed cloud data services and also it will be beneficial for the cloud.

As promising as it is, this paradigm also give rise to many new challenges for data security and access control. Data security, as it exists in many other applications, is among these challenges that would raise great concerns from users when they store sensitive information on cloud servers. These concerns originate from the fact that cloud servers are usually operated by commercial providers which are very likely to be outside of the trusted domain of the users. Data confidential against cloud servers is hence frequently desired when users outsource data for storage in the cloud.

As a significant research area for system protection, data access control has been evolving in the past thirty years and various techniques have been developed to effectively implement access control, which allows flexibility in specifying differential access rights of individual users. Traditional access control architectures [6] usually assume the data owner and the servers storing the data are in the same trusted domain, where the servers are fully entrusted as reference monitor responsible for defining and enforcing access control policies. This assumption however no longer holds in cloud computing since the data owner and cloud servers are very likely to be in two different domains. On one hand, cloud servers are not entitled to access the outsourced data content for data confidentiality; on the other hand, the data resources are not physically under the full control of the owner. For the purpose of helping the data owner enjoy access control of data stored on untrusted cloud servers, a feasible solution...
would be encrypting data through certain cryptographic primitive(s), and disclosing decryption keys only to authorized users. Unauthorized users, including cloud servers, are not able to decrypt since they do not have the data decryption keys. This general method actually has been widely adopted by existing works which aim at securing data storage on untrusted servers [10]. One critical issue with this branch of approaches is how to achieve the desired security goals without introducing a high complexity on key management and data encryption.

Another biggest concern with cloud data storage is that of data integrity verification at untrusted servers. For example, the storage service provider, which experiences intricate failures occasionally, may decide to hide the data errors from the clients for the benefit of their own. What is more serious is that for saving money and storage space the service provider might neglect to keep or deliberately delete rarely accessed data files which belong to an ordinary client.

Previously many cryptographic methods were adopted to preserve data integrity and confidentiality. They are classified as:

- Secret Key Cryptography (SKC): Uses a single key for both encryption and decryption.
- Public Key Cryptography (PKC): Uses one key for encryption and another for decryption.
- Hash Functions: Uses a mathematical transformation to irreversibly "encrypt" information.

At present, these methods alone were proven inefficient for providing security. This inefficiency is a result of advancement in technology. The challenges in cloud computing paradigm are:

- Storing large data files correctly and maintaining them.
- Verify correctness of the remote data even without the existence of local copies.
- Data security. These concerns originate from the fact that cloud servers are usually operated by commercial providers which are very likely to be outside of the trusted domain of the users.
- Data confidential against cloud servers when users outsource data for storage in the cloud.

These challenges, if not well resolved, may impede its fast growth. In this work the challenges are addressed by [16] using Key-policy attribute-based encryption (KP-ABE) cryptographic primitive and Proxy re-encryption. KP-ABE is proposed to resolve the problem of data access control in one-to-many communications, while proxy re-encryption is a technique which is mainly used to protect the data from its disclosure in the cloud storage server. This technique helps the cloud servers to provide the requested data to the client without knowing the underlying content.

II. RELATED WORK

A. Uniqueness of Cloud Computing

Cloud computing is cost-effective. Here, cost is greatly reduced as initial expense and recurring expenses are much lower than traditional computing. Maintenance cost is reduced as a third party maintains everything from running the cloud to storing data. Cloud is characterized by features such as platform, location and device independency, which make it easily adoptable for all sizes of businesses, in particular small and mid-sized. However, owing to redundancy of computer system networks and storage system cloud may not be reliable for data, but it scores well as far as security is concerned. In cloud computing, security is tremendously improved because of a superior technology security system, which is now easily available and affordable. Yet another important characteristic of cloud is scalability, which is achieved through server virtualization. Some of [19] the most important five key characteristics are,

1. On-demand nature Service
2. Open Network Access
3. Resource Pooling
4. Calculated Service
5. Selection of Provider

B. Definitions

Various terminologies related to this paper have been explained in this section [1][2].

- **Cloud Service Provider (CSP):** A CSP is the most important part of cloud architecture. CSP acts as an interface between the cloud resources and cloud prescribes. It provides all the services (SaaS, PaaS, IaaS) available for a client in "pay as you use" manner.
- **Cloud Storage Server (CSS):** A CSS is an entity, which is managed by Cloud Service Provider (CSP). It has significant storage space and computation resource to maintain client’s data.
- **Third Party Auditor (TPA):** A TPA is an entity which maintains records for clients and data owner. Records include login, file access and logout information.
- **Software as a Service (SaaS):** This is one of the services provided by cloud, in which a user no longer owns the software that is utilized but instead uses it when required.
- **Key-policy attribute-based encryption (KP-ABE):** KP-ABE is a public key cryptography primitive for one-to-many communications. In this, data are associated with attributes for each of which a public key component is defined.
- **Proxy Re-Encryption (PRE):** PRE is also a cryptography primitive that allows a proxy to transform a cipher text computed under A’s public key into one that can be opened by B’s secret key.
C. Purpose & AIM

The purpose of this paper is to provide security to the confidential data that the data owner store on the cloud storage server, to preserve the data integrity while it is transmitted from data owner to CSS and downloaded by the authorized client and this application assigns unique access structure to each client to access the data stored on the cloud monitored by TPA. The application is desktop based, and also user can access the CSS from any location using the data owner and client interface. The application manages and provides security to all the confidential data that the data owner stores and this application also stores information regarding each interaction that data owner and the authorized clients perform on the CSS. This project also aims to reduce the burden on the data owner begin online to distribute the cryptographic keys to the authorized clients. The application facilitates data storage and data integrity. The application GUI is intuitive and user friendly.

For the purpose of helping the data owner to secure data stored on untrusted cloud servers, the existing solution is to encrypt data through certain cryptographic primitive(s), and disclosing decryption keys only to authorized users. Unauthorized users, including cloud servers, are not able to decrypt since they do not have the data decryption keys. This general method actually has been widely adopted by existing works which aim at securing data storage on untrusted servers. In the existing technique there was a possible threat while exchanging the confidential data as data is stored on the untrusted cloud servers. One critical issue with this approach is how to achieve the desired security goals without introducing a high complexity on data owner of key management and data encryption using numerous secret keys. In the existing approach there exists heavy computation overhead on the data owner and there is no facility access to the data stored on the cloud server. And also in the *existing technique the TPA can only keep track* of the different interactions on the data stored in the cloud server [3]. By this information a data owner can only know, the culprits (unauthorized users) who made the access and what data have been modified or deleted. But this technique cannot prevent the cloud provider or other culprits from modifying or deleting the contents on cloud. With this motivation, an application is developed to reduce the burden on the data owner by assigning most of the computation intensive tasks on the cloud server and also it won’t allow the provider or any others to know the underlying contents stored in the cloud server. Any platform is used and the deployment is remote desktop based. The remoting concept is used provide an abstract approach to interprocess communication that separates the remotes object from a specific client or server application domain and from a specific mechanism of communication.

III. CRITICS ON EXISTING WORKS

Ateniese et al. are the first to consider public auditability in their defined “provable data possession” (PDP) model for ensuring possession of files on untrusted storages. In their scheme, they utilize RSA based homomorphic tags for auditing outsourced data, thus public auditability is achieved.

**Disadvantage**

Ateniese et al. do not consider the case of dynamic data storage, and the direct extension of their scheme from static data storage to dynamic case may suffer design and security problems.

Ateniese et al. propose a dynamic version of the prior PDP scheme. However, the system imposes a priori bound on the number of queries and does not support fully dynamic data operations.

**Disadvantage**

It only allows very basic block operations with limited functionality, and block insertions cannot be supported.

Wang et al. consider dynamic data storage in a distributed scenario, and the proposed challenge-response protocol can both determine the data correctness and locate possible errors.

**Disadvantage**

They only consider partial support for dynamic data operation.

Juels et al. describe a “proof of retrievability” (PoR) model, where spot-checking and error-correcting codes are used to ensure both “possession” and “retrievability” of data files on archive service systems. Specifically, some special blocks called “sentinels” are randomly embedded into the data file F for detection purpose, and F is further encrypted to protect the positions of these special blocks.

**Disadvantage**

The number of queries a client can perform is also a fixed priori, and the introduction of pre-computed “sentinels” prevents the development of realizing dynamic data updates. In addition, public auditability is not supported in their scheme.

Shaicham et al. design an improved PoR scheme with full proofs of security. They use publicly verifiable homomorphic authenticators built from BLS signatures, based on which the proofs can be aggregated into a small authenticator value, and public retrievability is achieved.

IV. PROBLEM STATEMENT

The figure 1. Shows the three different network entities for cloud data storage. ] A number of problems exist [3] to providing secured data storage and data integrity which are addressed by the following
The data owner can only know who are all the culprits (unauthorized users) who made the access and what data have been modified or deleted.

The existing technique cannot prevent the cloud provider or other culprits from modifying or deleting the contents on cloud.

To keep sensitive user data confidential against untrusted servers, existing solutions usually apply cryptographic methods by disclosing data decryption keys only to authorized users. However, in doing so, these solutions inevitably introduce a heavy computation overhead on the data owner for key distribution and data management when data access control is desired, and thus do not scale well.

The data owner should always online to provide decryption keys to authorized clients.

V. SYSTEM ANALYSIS

The verification schemes with public audit ability any TPA in possession of the public key can act as a verifier. We assume that TPA is unbiased while the server is untrusted. For application purposes, the clients may interact with the cloud servers via CSP to access or retrieve their pre-stored data. More importantly, in practical scenarios, the client may frequently perform block-level operations on the data files. The most general forms of these operations we consider in this paper are modification, insertion, and deletion.

A. Existing System

In the present world the integrity and confidentiality of data storage in Cloud Computing is maintained by delegating the task of allowing a third party auditor (TPA) on behalf of the cloud client, to verify the integrity of the dynamic data stored in the cloud. The introduction of TPA eliminates the involvement of client through the auditing of whether his data stored in the cloud is indeed intact, which can be important in achieving economies of scale for Cloud Computing. The support for data dynamics via the most general forms of data operation, such as block modification, insertion and deletion, is also a significant step toward practicality, since services in Cloud Computing are not limited to archive or back up data only.

To ensure cloud data storage security, it is critical to enable a third party auditor (TPA) to evaluate the service quality from [5] an objective and independent perspective. Public verifiability also allows clients to delegate the integrity verification tasks to TPA while they themselves can be unreliable or not be able to commit necessary computation resources performing continuous verifications.

But in this technique the TPA can only keep track of the different interactions on the data stored in the cloud. By this information a data owner can only know, the culprits (unauthorized users) who made the access and what data have been modified or deleted. But this technique cannot prevent the cloud provider or other culprits from modifying or deleting the contents on cloud.

In other techniques to keep sensitive user data confidential against untrusted servers, existing solutions usually apply cryptographic methods by disclosing data decryption keys only to authorized users. However, in doing so, these solutions inevitably introduce a heavy computation overhead on the data owner for key distribution and data management when data access control is desired, and thus do not scale well. Here the data owner should be online always to provide decryption keys to authorized users.

At present, a number of challenges exist in providing secured data storage and data integrity which are addressed in this work. The challenges faced at present are:

- The data owner can only know, who are all the culprits (unauthorized users) who made the access and what data have been modified or deleted.
- The existing technique cannot prevent the cloud provider or other culprits from modifying or deleting the contents on cloud.
- To keep sensitive user data confidential against untrusted servers, existing solutions usually apply cryptographic methods by disclosing data decryption keys only to authorized users. However, in doing so, these solutions inevitably introduce a heavy computation overhead on the data owner for key distribution and data management when data access control is desired, and thus do not scale well.
- The data owner should be online always to provide decryption keys to authorized clients.

B. Proposed System

There are many challenges in Cloud Computing, if not well resolved, may impede its fast growth. In this work the challenges of the existing techniques are addressed by using some techniques of Key-policy attribute-based encryption (KP-ABE) and Proxy re-encryption. KP-ABE is proposed to resolve the problem of data access control in one-to-many communications, while proxy re-encryption is a technique which is mainly used to protect the data from its
disclosure in the cloud storage server. This technique helps the cloud servers to provide the requested data to the client without knowing the underlying content. The .Net Framework used facilitates ease of use, reduces development efforts and supports improved security and communication features platform remoting allows objects to interact with each other across application domains.

The figure 2 shows the major network entities in the proposed cloud architecture are [1]:

- **Data owner**: An entity, which has large data files to be stored in the cloud and relies on the cloud for data maintenance and computation, can be either individual consumers or organizations.

- **Cloud Storage Server (CSS)**: An entity, which is managed by Cloud Service Provider (CSP). It has significant storage space and computation resource to maintain client’s data.

- **Effective Third Party Auditor (ETPA)**: A TPA is an entity which maintains records for clients and data owner. Records include login, file access and logout information.

![Figure 2. Network Entities of Proposed Cloud architecture](image)

**Figure 2. Network Entities of Proposed Cloud architecture**

**VI. SYSTEM DESIGN**

**A. System Architecture**

Large systems are always decomposed into sub-systems that provide some related set of services. The initial design process of identifying these sub-systems and establishing a framework for sub-system control and communication is called Architecture design and the output of this design process is a description of the software architecture.

The architectural design process is concerned with establishing a basic structural framework for a system. It involves identifying the major components of the system and communications between these components.

The system architecture shows the blocks required for the paper. Figure 3 shows the existing system architecture.

![Figure 3 Cloud Computing System Architecture](image)

**Figure 3 Cloud Computing System Architecture**

The data owner is the one who encrypts the document (text, image and word document) that he wants to upload on the cloud server. After which he performs the AES encryption to encrypt the plain document, proxy re-encryption and attribute based encryption to provide unique access structure to each client. Then he categorizes the data into different categories by linking different proxy key generated using the proxy re-encryption technique. The data owner is the one who has the priority of editing and also deleting the contents stored on the cloud server. Then he performs the attribute based encryption to provide unique access structure to each client. After all these steps the data owner uploads the encrypted file, re-encryption key and access structure on to the cloud server. Only authorized client can log in using his private key. After successful he gets a unique access structure that has been assigned by the data owner. Clients request the cloud server for the re-encryption key and files that he wants to download. Finally client decrypts the re-encryption key and then decrypts the encrypted file using this key.

**VII. SYSTEM IMPLEMENTATION**

**A. Establishing client and TPA**

An entity, which has large data files to be stored in the cloud and relies on the cloud for data maintenance and computation, can be either individual consumers or organizations. TPA which has expertise and capabilities that clients do not have, is trusted to assess and expose risk of cloud storage services on behalf of the clients upon request.

**B. Algorithm for overall implementation:**

Step 1: Initially the cloud server is started.

Step 2: Then the data owner login using the data upload interface.

Step 3: He encrypts the data.

Step 4: Perform PRE (Proxy Re-Encryption), using client’s public key.

Step 5: Further he performs KP-ABE (Key Policy Attribute Based Encryption).
Step 6: He then generates different categories using the proxy key generated during the PRE. Upload the data.
Step 7: Authorized clients who are authenticated by the data owner login using their private key.
Step 8: Each client gets access structure of separate files based on access policy defined by the data owner to that particular client.
Step 9: Download required files.

C. Results

The figure 4 shows the monitored information of the data owner or the cloud clients/users by the TPA.

![Figure 4. Monitoring of Cloud Clients by TPA](image)

The figure 5 shows the Data owner uploading the Information of the encrypted format with the help of Public Key, Private Key and Attribute Key.

![Figure 5. Data Owner Uploading data to the Cloud](image)

The figure 6 shows the Authenticated Cloud client’s can access the uploaded data by the data owner with the help of its private key only when the data owner permits.

![Figure 6. Cloud Users/ Clients downloading the data](image)

VIII. CONCLUSION & FUTURE ENHANCEMENT

This work studies the problem of ensuring the integrity of data storage in Cloud Computing. In particular, we consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to monitor the integrity of the dynamic data stored in the cloud. We have achieved data confidentiality and preserved data integrity with the help of KP-ABE and proxy re-encryption techniques.

We utilize and uniquely combine the public key based homomorphic authenticator with random masking to achieve the privacy-preserving public cloud data auditing system. This scheme is the first to support scalable and efficient public auditing in the Cloud Computing. The technique of AES algorithm where TPA can perform the auditing tasks simultaneously. The data in the cloud does not remain static. Unlike most prior works, the new scheme further supports secure and efficient dynamic operations on data stored in the cloud, including: data update, delete and append. Unique access structure enhances security. Implementation can be simplifies the development efforts and provides direct support for security. Extensive security and performance analysis show that the proposed scheme is highly efficient and provably secure.

In Future, the Author intends to enhance the system by developing a module to write files to the storage server. In addition the author is interested in handling different types of multimedia, as part of future enrichment. The author is interested in making this application web based.
APPENDIX

B. KP-ABE for constructing Public AUDIT

The public auditing system can be constructed in four phases by using the Key Policy Attribute-Based Encryption (KP-ABE) algorithm

Step1: Setup:
This algorithm takes as input a security parameter and the attribute universe U = {1, 2...N} of cardinality N. It returns the public key PK as well as a system master key MK. While PK is publicly known to all the parties in the system, MK is kept as a secret by the authority party.

Step2: Encryption
This algorithm takes a message M, the public key PK, and a set of attribute I as input. It outputs the cipher text E.

Step3: Key Generation
This algorithm takes as input an access tree T, the master key MK, and the public key PK. It outputs a user secret key SK.

Step4: Decryption
This algorithm takes as input the cipher text E encrypted under the attribute set I, the user’s secret key SK for access tree T, and the public key PK. Finally, it outputs the message M if and only if I satisfies T.

A. AES ALGORITHM For Encryption and Decryption

AES algorithm is used for generating the cipher text for a given plain text because the input as well as intermediate data can be considered as a matrix with four rows and four columns called state. Each element of the matrix is composed of eight bits, therefore enabling efficient implementation of AES on 8 bit platforms also.

The AES algorithm has four basic transformations.

1) Sub Byte Transformation - a nonlinear transformation applied to the elements of the matrix. This first step in each round is a simple substitution, implemented as a Look Up Table (LUT). It operates independently on each byte of state using S-box.

2) Shift Rows Transformation - a cyclical shift operation with constant offsets, applied to the rows of the matrix. This second step in each round is permutation of rows by left circular shift; the first (leftmost, high order) i elements of row i are shifted around to the end (rightmost, low order).

3) Mix Columns Transformation - the third step is a resource intensive transformation on the columns of state under which the four elements of each column are multiplied by a polynomial, essentially diffusing each element of the column over all four elements of that column.

4) AddRoundKey Transformation - performs modulo 2 (XOR) operations with the round key, which is obtained from the initial key by a key expansion procedure. The encryption flow starts with the addition of the initial key to the plaintext. Then the iteration continues for (Nr - 1) rounds (Nr being the total number of rounds). In last round the Mix Column step is bypassed.

AES Algorithm

KeyExpansion(byte key[4*Nk], word w[Nb*(Nk+1)], Nk)
Cipher(byte in[4*Nb], byte out[4*Nb], word w[Nb*(Nk+1)])
begin
byte state[4,Nb]
state = in
AddRoundKey(state, w[0, Nb-1])
for round = 1 step 1 to Nr-1
SubBytes(state)
ShiftRows(state)
MixColumns(state)
AddRoundKey(state, w[round*Nb, (round+1)*Nb-1])
end for
SubBytes(state)
ShiftRows(state)
AddRoundKey(state, w[Nr*Nb, (Nr+1)*Nb-1])
out = state
end

ACKNOWLEDGMENT

The authors would like to thank the editor, mystifying reviewers for their priceless suggestions that appreciably improved the eminence of this paper. Finally thankful to our Head-of-Departments and our dear faculties, Department of CSE and ISE, SJC Institute of Technology, Chickballapur, Karnataka, India.

I heartily thank Principal, BMSIT, Bangalore and, HOD, Department of CSE, BMSIT, Bangalore, Karnataka, India for their continuous encouragement and motivation.

Last but not least to all of them who helped us to get a warm success in completion of this paper.

REFERENCES


[14] Shucheng Yu1, Kui Ren2, Wenjing Lou1, and Jin Li2,"Defending Against Key Abuse Attacks in KP-ABE Enabled Broadcast Systems”


Mr. Muneshwara Ms is currently pursuing his final semester Master of Technology in Computer Science and Engineering at S J C Institute of Technology, Chickballapur, affiliated to Visvesvaraya Technological University, Belgaum, Karnataka, India.

Mr. Arvind Tejas Chandra is an Associate professor in the Department of Computer Science and Engineering at S J C Institute of Technology, affiliated to Visvesvaraya Technological University, Belgaum, Karnataka, India.