JCSE International Journal of Computer Sciences and Engineering Open Access

Survey Paper Volume-6, Special Issue-3, April 2018 E-ISSN: 2347-2693

A Survey on Sharing Data in Cloud

S. Nivetha[1], R. Sowmya[2], B. Monica Jenefer[3*]

Department of CSE, Meenakshi Sundararajan Engineering College, Anna University, Chennai, India

**Corresponding Author: hod.cse@msec.edu.in, Tel.: +91 9940667342*

Available online at: www.ijcseonline.org

Abstract— Cloud is an online storage application mainly used to store documents, media and various files. However, due to large and easy access of cloud there arise various security issues such as data stealing and authentication issues when trying to share data on public platform. Fine-grained sharing of encrypted data is achieved through Attribute-based encryption. This is done with a methodology to provide data confidentiality and integrity using key generation, encryption and decryption. There are various techniques used to share data in cloud and each of these techniques employ different procedures or steps to achieve the end result. One among those techniques are public key encryption techniques which is a form of asymmetric key encryption where both public and private keys are used to keep the data or document secure. A security model is built through verifiable decryption algorithm. This is achieved by introducing a verification key from the output of the encryption algorithm. Finally, we present an approach to securely share data in an efficient manner through the ABE scheme.

Keywords— cloud computing, attribute-based encryption, Key Exchange, Data Sharing.

I. INTRODUCTION

Cloud computing have grown vastly over the years and lately many IT industries rely on safe and secure form of transfer of data to either individuals or to a group of people. The benefit organizations can gain from data sharing is higher productivity. With social networking services gaining popularity it needs to focus on sharing data. Google Docs is one such cloud platform which provides data sharing capabilities as groups of students, or teams working on a project can share documents and can collaborate with each other effectively. There is an assumption that data servers can be trusted to keep the data secure. However, this assumption is no longer true today since services are increasingly storing data across many servers that are shared with other data owners. The Cloud is susceptible to many privacy and security attacks. The biggest obstacle hindering the progress and the wide adoption of the Cloud is the privacy and security issues associated with it.

An example of this is cloud data storage where cloud service providers are not in the same trusted domains as end users, and hardware platforms are not under the direct control of data owners. To mitigate user's privacy concerns about their data, a common solution is to store data in encrypted form so that it will remain private, even if data servers or storage devices are not trusted or compromised.

The encrypted data, however, must be amenable to sharing and access control. Data encryption using symmetric or public key cryptography is not amenable to scalable access control. A promising approach to address this issue is attribute-based encryption (ABE), ABE schemes can be divided into two categories: Cipher text- Policy ABE (CP-ABE) and Key-Policy ABE (KP-ABE), depending on the access policy is embedded into the ciphertext or the user's private key. Here, Both CP-ABE and KP-ABE can prevent any unauthorized users from accessing data, even if the user stores data in an untrusted server.

The rest of this paper is organized as follows. Section 2 presents some fundamental concepts Public Key Encryption. Section 3 gives the details of domain based storage protection. Section 4 gives the verification of performance evaluation through high level petri nets. Section 6 presents privacy preserving techniques employed over the years to share data across database; section 7 involves the key exposure techniques and key aggregate cryptosystems. Finally, Section 8 concludes our work.

II. PUBLIC KEY ENCRYPTION

This type of encryption is also called asymmetric cryptography. Here it uses both public and private keys to encrypt and decrypt data.

A. Identity-Based Encryption

An Identity Based Encryption scheme is a public-key cryptographic algorithm, which consists of 3 elements: key generation, encryption and decryption where any string is a valid public key.

In [3] Jianghong et al. proposed a revocable storage identity based encryption. This was mainly introduced to overcome various disadvantages faced by data sharing over cloud. Some of the issues faced are confidentiality of data and they are compromised sometimes as the cloud server are not reliable they may outsource data for their own benefits. The main advantage posed by this paper was to not make the data available for access once the user is denied access or his authority has been expired.

The major drawback in this approach is it uses a decrypt then re-encrypts which involves user's secret key information which makes the data open to new attacks and vulnerable and furthermore the paper does not discuss on authenticity and availability of shared data

B. Attribute Based Encryption

Attribute based encryption is a public-key encryption. Here the attributes are made to be dependent on the secret key and ciphertext. It reduces the number of key used and thus make encryption and decryption process faster.

In [6] Ruixuan et al. proposes an algorithm called lightweight data sharing scheme with an addition to ciphertext policy attribute based encryption is used to offer efficient access control and proxy servers are used for encrypting and decryption operations this reduces computational overhead. Here it uses lazy re-encryption to reduce the overhead caused by revocation. The major disadvantage posed by this paper is data integrity is not verified and ciphertext retrieval over existing data sharing schemes are not figured

In [8] Joseph et al. The proposed concept integrates attribute-based encryption with fine-grained two-factor access to remove the use of public key exposure and the storage of secret key. The ciphertext are of Attribute based encryption form. The main difficulty in this framework is it needs to revoke the old device and to use new security device to do decryption properly and in order to revoke the old device it needs the cloud to update the old ciphertext before sending them to user.

In [12] Baodong et al. The paper pointed some of the disadvantages posed by ABE where there is relatively large ciphertext size and high decryption cost, and this problem is especially acute for resource limited devices such as mobile devices. Specifically, in an ABE scheme, the size of the ciphertext and the cost of decryption grow with the complexity of the access structures/policies. Therefore, it converts an ABE cipher text into a El Gamal-type cipher text using a public transformation key provided by a user so that the user can decrypt the latter much more efficiently than the former. Major drawback is it helps to ensure the data owner's data being stored in the cloud is valid or not that is data integrity is not given importance

III. DOMAIN BASED STORAGE PROTECTION

DBSP provides both data confidentiality and integrity protection mechanism for IaaS environments. It also provides transparent storage isolation between IaaS clients which relies on trusted computing principles.

In [2] Nicole et al. In this paper DBSP (domain-based storage protection) is presented as a virtual disk encryption mechanism where encryption of data is done directly on the compute host, while the key material necessary for re-generating encryption keys is stored in the volume metadata. It would provide 3 advantages;1. This would eliminate bottle neck problems and congestion.2. Response time would be fast when compared to the traditional approach. 3. Easy to analyse and provide heap memory space based on the usage. For example, both server and virtual machines should be connected to a same Internet connection to upload and access the user files. The paper has covered only a fraction of attacks posed on a Iaas environment it does not focus much on strengthening the trust model.

IV. HIGH LEVEL PETRI NETS

Petri Nets provide graphical representation of the system and can be applied to variety of systems. A HLPN is a 7-tuple $N = (P, T, F, \varphi, R_n, L, M_0)$, where P is set of places; T refers to the set of transitions such that $P \cap T = \emptyset$; Flow relations are defined by F]; φ maps places P to the data types. Rules for transitions are defined by R_n; L is a label on F and M_0 represents the initial marking. In the above definition, the structure of the Petri Net is given by P, T, and F

In [4] Mazhar et al. And [5] saif et al. Both the paper proposed using a high level petri nets for formal modeling and verification of the methodology The HLPN define mathematical properties for the system and simulate the system to analyze behavior. Then the verification of both the HLPN model using Satisfiability Modulo Theories Library (SMT-Lib) and Z3 solver is done. To verify the model, Petri Net model is first translated into SMT along with specified properties. Subsequently, Z3 solver is used to determine whether or not the properties hold.

V. DEDUPLICATION

Deduplication is the process of avoiding repeated storage of a single file so as to reduce the storage overhead. It is also called as single instance storage as it does not allow redundant copies of a file to be stored in the storage device. This helps in enhancing the storage utilization and it can be used in network data transfers to control the number of bytes that must be sent.

In [1] Yifeng et al. Propose a functionality that can be added to the cloud service which completely eliminates the storage of redundant copies of a particular file. The cloud service may encounter such a problem because of hosting images, videos or even text from different sources. This functionality is just an addition to the existing framework to perform secure deduplication. Different from the existing work which ensures a defending system for offline brute-force attacks, this paper formulates the encrypted cloud media center with comprehensive protection. This framework resists data leakage and also the off-line brute-force attack.

The system architecture consists of an encrypted cloud media center and an agency server. This agency server properly executes encryption of data in a way that will help achieve efficient deduplication. The main disadvantage is this is done over only predictable videos and data.

VI. PRIVACY PRESERVING

Privacy preserving is just a term used for accessing documents only for the user's purpose. It involves two users who own a confidential database. It runs a data mining technique on the union of their database without revealing sensitive information.

In [9] Bernardo et al. propose a secure framework for privacy preserving outsourced storage and retrieval of dynamically updated image repositories. This paper focuses on image data as the visual data is nowadays the major part of the global internet traffic. The proposed system based on Image Encryption Scheme (IES) with Content-Based Image Retrieval (CBIR) properties. By using this image is separated into its color based and texture based data and each portion is encrypted using different algorithms. Mostly the texture information is more useful while compared to color information in object recognition. Therefore, in this paper texture information is encrypted using semantically secure algorithms and color information is encrypted using a comparatively less secure algorithm.

The system model architecture consists of Cloud Infrastructure, Key Distribution Service and end users. The main advantage of this paper is that by using this framework ensuring privacy is executed without overloading the client overhead (for encryption purposes). The main disadvantage of this framework is that it cannot be extended to other data sets.

VII. KEY EXPOSURE AND AGGREGATE CRYPTOSYSTEM

The issue mainly faced by cloud is key being leaked. In order to overcome this, key exposure framework is used wherein even if the key is leaked the data shared will be secure.

In [10] Ghassan et al. has used framework which has All-Or-Nothing encryption technique. This technique can be easily broken by brute-force attack so this paper proposed a technique- Bastion which ensures confidentiality even when key is made public. Here ciphertext blocks are stored in multi-cloud storage system. This scheme improves the performance of existing primitives and incurs negligible overhead. The drawback that can be found is that until now this framework has not implemented leakage resilient feature.

Key Aggregate Cryptosystem (KAC)is a system that was built to help the process of efficiently broadcasting a single key to multiple users using which those users can decrypt multiple data classes.

In [11] Sikhar et al. has gone a step ahead to overcome the security proofs drawback by building a CPA and CCA secure KAC for cloud based data sharing environments. The paper proposes an efficiently implementable version of the basic key-aggregate cryptosystem using asymmetric bilinear pairings. This idea is useful as this proposes an efficient feature for basic KAC scheme which is fully collusion resistant with low overhead ciphertexts and aggregate keys.

Table 1. Comparative analysis

Techniques	**Parameters**			
	confidentiality	***authenticity***	***Integrity***	***Storage Utilization***
Public key Encryption *IBE*	✓		✓	
ABE	✓	✓	✓	
DBSP	✓		✓	✓
Deduplication			✓	✓
Privacy Preserving	✓	✓		
Key Exposure & Key Aggregate Cryptosystem	✓	✓		

VIII. CONCLUSION

The data sharing has always been a problem in cloud as there can be no way of identifying data stealing or other attacks. This paper helps in identifying the pros and cons of using various techniques that have been employed for secure transaction or data sharing over cloud server. The public key encryption that are widely used is identity based encryption which is then evolved to be an attribute based encryption but here integrity is not verified or evaluated. However, in both domain based storage protection and key exposure techniques both confidentiality and integrity are maintained. Some of the other features that are added to the cloud for data sharing is deduplication, privacy preserving which maintains the cloud free of duplicate data and to make the data available only for the user's purpose. As a result, we build a secure data sharing system across cloud which helps to overcome the obstacles of sharing or storing data on the long run and there are many opportunities provided to improve.

REFRENCES

[1] Yifeng Zheng, Xingliang Yuan, Xinyu Wang, Jinghua Jiang, Cong Wang, Xiaolin Gui, "*Towards Encrypted Cloud Media Center with Secure Deduplication*" IEEE Transactions on Multimedia, Vol.**19**, Issue.**2**, **2017**.

[2] Nicolae Paladi, Christian Gehrmann, Antonis Michalas, "*Providing User Security Guarantees in Public Infrastructure Clouds*" IEEE Transactions on Cloud Computing, Vol.**5**, Issue.**3**, **2017.**

[3] Jianghong Wei, Wenfen Liu, Xuexian Hu, "*Secure Data Sharing in Cloud Computing Using Revocable-Storage Identity-Based Encryption*" IEEE Transactions on Cloud Computing, Vol.**PP**, Issue.**99**, **2015**. doi: 10.1109/TCC.2016.2545668

[4] Mazhar Ali,Revathi Dhamotharan,Eraj Khan,Samee U.Khan,Athanasios V.Vasilakos,Keqin Li and Albert Y.Zomaya, "*SeDaSC: secure data sharing in clouds*" IEEE Systems Journal, Vol.**11**, Issue.**2**, **2017**.

[5] Mazhar Ali,Saif U.R.Malik and Samee U.Khan, "*DaSCE: Data Security for Cloud Environment with Semi-Trusted Third Party*" IEEE Transactions on Cloud Computing Vol.**5**, Issue.**4**, **2017.**

[6] Ruixuan Li; Chenglin Shen; Heng He; Zhiyong Xu; Cheng-Zhong Xu, "*A Lightweight Secure Data Sharing Scheme For Mobile Cloud Computing*" IEEE Transactions on Cloud Computing, Vol.**PP**, Issue.**99, 2017.** doi: 10.1109/TCC.2017.2649685

[7] Jian Shen , Tianqi Zhou, Xiaofeng Chen, Jin Li, and Willy Susilo, "*Anonymous And Traceable Group Data Sharing In Cloud Computing*" IEEE transactions on information forensics and security, Vol.**13**, Issue.**4**, **2018.**

[8] Joseph K. Liu, Man Ho Au,Xinyi Huang, Rongxing Lu, and Jin Li, "*Fine-Grained Two-Factor Access Control For Web-Based Cloud Computing Services*" IEEE Transactions on Information Forensics and Security, Vol.**11**, Issue.**3**, **2016.**

[9] Bernardo Ferreira, Joao Rodrigues, Joao Leitao, Henrique Domingos, "*Practical Privacy-Preserving Content-Based Retrieval In Cloud Image Repositories*" IEEE Transactions on Cloud Computing, Vol.**PP**, Issue.**99, 2017.** doi: 10.1109/TCC.2017.2669999

[10] Ghassan O. Karame, Claudio Soriente, Krzysztof Lichota, Srdjan Capkun, "*Securing Cloud Data Under Key Exposure*" IEEE Transactions on Cloud Computing, Vol.**PP**, Issue.**99, 2017.** doi: 10.1109/TCC.2017.2670559

[11] Sikhar Patranabis, Yash Shrivastava, Debdeep Mukhopadhyay, "*Provably Secure Key-Aggregate Cryptosystems With Broadcast Aggregate Keys For Online Data Sharing On The Cloud*" IEEE Transactions on Computers, Vol.**66**, Issue.**5**, **2017.**

[12] Baodong Qin, Robert H. Deng, Shengli Liu, Siqi Ma, "*Attribute-Based Encryption With Efficient Verifiable Outsourced Decryption*" IEEE Transactions On Information Forensics And Security, Vol.**10**, Issue.**7**, **2015.**

Authors Profile

B. MonicaJenefer received her Bachelor of Engineering in Electrical & Electronics Engineering from Government College of Engineering, Tirunelveli affiliated to Manonmaniam Sundaranar University in 2000 and Master of Engineering in Computer Science and Engineering from College of Engineering Guindy, Anna University, Chennai in 2002. She is presently working as Head of the Department & Associate Professor at Meenakshi Sundararajan Engineering College for past 15 years. She is pursuing research at Sathyabama University and her area of research is biomedical image processing and computational modelling. She has published papers in many national/international conferences and journals. She has been involved in various capacities of research work in her career like consultant projects and student project coordination. She received the Cambridge International Certificate for Teachers and Trainers in the year 2011.

Ms S Nivetha is currently pursuing Bachelor of Engineering from Anna university Chennai since 2014 in Department of computer science.She has interest in Data Analytics and Big Data.

Ms R Sowmya is currently pursuing Bachelor of Engineering from Anna university Chennai since 2014 in Department of computer science. She has interest in Network Security.