Data Storage Security and Privacy Implementation based on Cryptographic Technique

Miss. Sayali M. Jawake Dr. S. S. Sherekar Dr. V. M. Thakare

Abstract **:- Cloud computing is the delivery of computing and storage capacity as a service to users. Cloud storage, as a subservice of infrastructure as a service (IaaS) in cloud computing, is a model of networked online storage where data is stored in virtualized pools of storage. As fast development and application of cloud computing and cloud storage, users concern more and more about security and privacy issues involved in these techniques. From industrial and academic viewpoints currently, cryptography is considered as a key technology to solve security and privacy problems. By reviewing the definition of cloud storage, and subsequently review the existing secure cloud storage based on cryptographic techniques, we can analyze and indicate what type of cryptographic techniques is mainly adopted in existing cloud storages and what role the cryptographic techniques play. In this paper we propose a new encryption algorithm, which we call Reverse Encryption Algorithm (REA).REA algorithm consists of one bit shift operation but in proposed system our approach is to implement the crossover operation like XOR. This new encryption algorithm (REA) is simple and is fast enough for most applications. REA encryption algorithm provides maximum security and limits the added time cost for encryption and decryption to as to not degrade the performance of a system.**

Keywords **– Cloud Computing, REA**

I. INTRODUCTION

Recently, cloud computing prevails over the whole world. One of the most important and popular services of cloud computing is cloud storage service, such as Amazon's S3 and Microsoft's Azure storage services. A lot of sensitive data will be stored into the cloud, so that the security of the cloud should be guaranteed. Due to data privacy, it is necessary for users to encrypt their sensitive data before storing them into the cloud. However, there exist some shortcomings in the setting of traditional encryption. The great flexibility and economic saving of cloud computing are motivating all kinds of users, such as customers, enterprises, and even government organizations, to adopt cloud. Cloud computing is an emerging paradigm, but its security and privacy risks has been attracting significant attentions of cloud users and cloud providers. One of the important reasons is that cloud users have to trust the security mechanisms and configuration of the cloud provider and the cloud provider itself. In the community of industry and academy, cryptographic technique is currently treated as one of the key techniques to solve security and privacy problems existed in cloud computing environment. It is well known that cloud storage is a specific sub-offering within IaaS of cloud computing. With cloud storage technology, private data of users is stored on multiple third-party providers, rather than on the dedicated providers used in traditional networked data storage. The providers supply data storage service through the Internet to users themselves and others. In cloud computing, users can utilize powerful computing resources and obtain ample storage spaces. This is called databaseas- a-service (DaaS) , software-as-a-service (SaaS) or infrastructure-as-a-service (IaaS) . Although it brings users great convenience, security issues in cloud computing are the primary concerns of users. Armbrust et al.gave a view of cloud computing and listed top ten obstacles and opportunities for cloud computing. The first obstacle is availability/business continuity where users in cloud computing are concerned about whether the computing services have adequate availability. The approach to overcome this obstacle is to use multiple cloud servers. The second obstacle is data lock-in, where users cannot extract their data from one site to run on another site. The potential solution for this obstacle can be hybridized: cloud computing and standardized APIs. The third obstacle is data confidentiality and audibility. This obstacle could be overcome by applying encryption and firewall techniques. In these obstacles, the first three affect the adoption of cloud computing, the next five affect the growth of cloud computing, and the last two are policy and business obstacles. Therefore, to improve the adoption of cloud computing, a desirable scheme should provide the following properties. First, it can provide confidentiality to sensitive data. Second, multiple service providers can co-exist. Third, it can make services available across domains. This obstacle can be overcome by providing open source or changing the licensing structures. In data storage systems, users can store their data on external proxy servers to reduce maintenance cost and enhance access and availability. To protect the confidentiality of the outsourced files, the owner encrypts them prior to outsourcing them to an untrusted proxy server. The proxy server can perform some functions on the ciphertexts, such that an authorized user can access the desired sensitive files.

II. BACKGROUND

Predicate encryption is a new cryptographic primitive that provides fine-grained control over the accesses to encrypted data. In the setting of predicate encryption, any messages can be encrypted with a set of attributes. Current privacy preserving search schemes over encrypted cloud storage services do not provide effective revocation for search privileges. Aiming at symmetric predicate encryptions and cloud storage requirements, we propose controllable privacy preserving search in cloud storage. This scheme is based on whose efficiency is much better than other. controllable privacy preserving search scheme has two new features. One is revocable delegated search which makes it possible for the secret key owner to control the lifetime of the delegation. The other is undecryptable delegated search. Due to this feature, a delegated person cannot decrypt the returned matched

ciphertexts even though he has the delegated privilege of search.

Cloud computing is a distributed system where multiple domains co-exist together. It is desirable that users in different domains can share sensitive data with others. Therefore, a sound identity-based data storage scheme in cloud computing should support not only intra-domain queries but also the inter-domain queries. So identity-based data storage scheme is proposed which supports both intra-domain and inter-domain queries. In proposed scheme, the re-encryption key is computed by the data owner independently without the help of the PKG. For one query, the requester can only access one file of the owner, while the requester and the proxy server can cooperatively access all the files of the owner in previous schemes as the access permission is not bound to the ciphertext in these schemes. Furthermore, identity based data storage scheme is secure against collusion attacks and is selective-identity secure in the standard model. Therefore, identity based data storage scheme can improve the adoption of cloud computing as it can overcome the availability/business continuity where users in cloud computing are concerned about whether the computing services have adequate availability, data lock-in, where users cannot extract their data from one site to run on another site as well as data confidentiality and audibility.

- secure cloud storage designed using cryptographic techniques

In proposed methodology, survey of existing application of cryptographic techniques with their main features will be done which results into secure cloud storage. simply, main focus is on cloud storages that are designed using cryptographic techniques but not in the framework of cryptography theory refer to as Class A and cloud storages that are designed using cryptographic techniques and also in the framework of cryptography theory refer to as Class B. For Class-B, cloud storage schemes were proven secure in the framework of provable-security theory in cryptography field.

III. Previous Work Done

When user from cloud server downloads all encrypted data, first step the user decrypt that data and searches the required information. It will be very inefficient and inconvenient if decrypted data are huge or the client is mobile user .This is obvious not feasible so in order to overcome this drawback predicate encryption schemes [C.Blundo,2009] were introduced. Like traditional encryptions, there are two categories in predicate encryption. [E. Shen,, 2009] are symmetric predicate encryption schemes. [D. Boneh, 2007] are asymmetric predicate encryption schemes. Generally speaking (but not absolutely), the symmetric predicate encryption can be used in secure cloud storage and the asymmetric predicate encryption can be used in encrypted mail search, biometric matching and credit card payment gateways.

Ivan and Dodis [A. Ivan,2007] implemented two identity-based proxy encryption schemes where the master secret key held by the PKG is split into two parts. One is for the user and the other is for the proxy server. Then, the user can cooperate with the proxy server to decrypt a ciphertext. Unfortunately, these schemes are not secure against collusion attacks [G. Ateniese,2006] as the user and the proxy server can collaborate to compute the master secret key. Green and Ateniese [M. Green,2007] introduced the concept of identity based proxy re-encryption (IBPRE). Chu and Tzeng [C.K. Chu,2007] implemented two IBPRE schemes in the standard model. Unfortunately, these schemes [C.K. Chu,2007] are not secure against the collusion attacks. If the designated decrypter can compromise the proxy server, he can obtain the secret key of the original decrypter. However, all above identity based proxy re-encryption schemes cannot be employed in cloud computing as they cannot support inter domain queries and resists collusion attack. So to overcome these problems identity based data storage scheme is implemented which supports both intra domain and inter domain queries.

The basic requirements for cloud storage systems include mass storage and low expense. However, users are reluctant to move important and sensitive data to cloud unless security and privacy issues can be well solved. To deal with this problem, lots of secure cloud storage architectures have been designed in recent years, and most of them are based on cryptographic techniques [Popa R A,2010]. As Security and privacy risks are the biggest concern when users want to apply cloud computing and cloud storage, there are various security problems involved in the cloud computing,[Subashini S, Kavitha V,2011] have presented security issues in SaaS, PaaS, IaaS of cloud computing including data security, network security, data segregation etc. To protect data in cloud computing and cloud storage, a standard approach currently is to apply cryptographic techniques into sensitive data [Tang Y, 2010]. Cryptographic techniques play an important role in the security protection of cloud storage, and in return the demand on secure cloud storage can promote the research of cryptography.

IV. Existing Methodologies

1) Identity based data storage scheme supporting inter domain and intra domain queries

2) Survey of results of various cryptographic techniques into secure cloud storage, categorized into two classes i.e. class A and class B.

V. Analysis and Discussions

- **Existing methodologies**
- **Preliminaries:**

By $s \xleftarrow{R} S$, we denote s is selected from S at random. If S is a finite set, by $s \xleftarrow{R} S$, we denote s is selected uniformly from S. By F(X) →y.we denote y is obtained by running the algorithm F on input x.

- **Identity based data storage:**

There are four entities in an identity-based data storage scheme: the private key generator (PKG), the data owner, the proxy server (PS) and the requester. The PKG validates the users' identities and issues secret keys to them. The data owner encrypts his files and outsources them to the proxy server. He

validates the requesters and issues access permissions to the proxy server. The proxy server stores the ciphertexts and can transfer them to ciphertexts for the requester when he obtains corresponding reencryption keys from the owner. The requester can decrypt the re-encrypted ciphertext. An identity-based data storage scheme supporting intra-domain and inter-domain queries consists of the following algorithms:

Setup(1^l)→(*params*,(MSK_1,PK_1),(MSK_2,PK_2))

This algorithm takes as inputs a security parameter 1ℓ and outputs the public parameters params, master secret–public key pairs (MSK_1, PK_1) and (MSK_2, PK_2) for PKG_1 in domain $\mathfrak{D}_1$ and PKG_2 in domain $\mathfrak{D}_2$, respectively.

KeyGen(params, ID, MSKi) → SK_{ID}. This algorithm takes as inputs the public parameters *params*, an identity *ID* in the domain Di and the master secret key MSKi, and outputs a secret key SK_{ID} for the identity ID, where i ∈{1, 2}.

Encryption (*params, ID,M*) → CT. This algorithm takes as inputs the public parameters params, the identity ID and the message M, and outputs the ciphertext CT = Encryption (params, ID,M). It sends the ciphertext CT to the proxy server PSi in the domain Di, where i = {1, 2}. Query (ID′, $SK_{ID'}$, CT) → Θ. The requester R with identity ID' queries the proxy server on the ciphertext CT. This algorithm takes as input the requester's identity ID', secret key $SK_{ID'}$ and the ciphertext CT, and outputs an authentication information Θ. The requester sends Θ to the proxy server PSi.

Permission (params, ID′, CT, SK_{ID}) → AK. The owner validates the requester by verifying the authentication information Θ. If the requester is legal, this algorithm takes as inputs the public parameters params, the requester's identity ID′, the intended ciphertext CT and the owner's secret key SK_{ID}, and outputs an access key (re-encryption key) AK. It sends AK to the proxy server PSi.

Re-encryption (params, ID′, AK, CT) →CT′. This algorithm takes as inputs the public parameters params, the requester's identity ID′,the access key AK and the ciphertext CT, and outputs the re-encrypted ciphertext CT′ = Encryption(params, ID′,M). Decryption. There are two algorithms. One is for the owner and the other is for the requester.

1. Decryption1 (*params, SK_{ID}, CT*) → M. This algorithm takes as inputs the public parameters params, the owner's secret key SK_{ID} and the ciphertext CT, and outputs the message M.

2. Decryption2 (params, $SK_{ID'}$, CT′) → M. This algorithm takes as inputs the public parameters params, the requester's secret key $SK_{ID'}$ and the re-encrypted ciphertext CT′, and outputs the message M.

$$\Pr\left[\text{Decryption}_1(params, SK_{ID}, CT) \to M \;\middle|\; \begin{array}{l}\text{Setup}(1^\ell) \to (params, MSK, PK);\\ \text{KeyGen}(params, ID, MSK) \to SK_{ID};\\ \text{Encryption}(params, ID, M) \to CT\end{array}\right] = 1$$

and

$$\Pr\left[\text{Decryption}_2(params, SK_{ID'}, CT') \to M \;\middle|\; \begin{array}{l}\text{Setup}(1^\ell) \to (params, MSK, PK);\\ \text{KeyGen}(params, ID, MSK) \to SK_{ID};\\ \text{KeyGen}(params, ID', MSK) \to SK_{ID'};\\ \text{Permission}(params, ID', CT, SK_{ID}) \to AK;\\ \text{Re-encryption}(params, ID', AK, CT) \to CT'\end{array}\right] = 1$$

Box I.

Definition 1. We say an identity-based data storage scheme supporting intra-domain and inter-domain queries is correct if the conditions given in Box I are satisfied, where the probability is taken over the random coins which all the algorithms in the scheme consumes.

Identity based data storage scheme in cloud computing:

An identity-based data storage scheme supporting intra-domain and inter-domain queries has been implemented to prove its security. In proposed scheme, the access permission can be determined by the data owner independently without the need of the PKG. Especially, the access permission is bound to the requested ciphertext. Furthermore, this scheme is secure against the collusion attacks.

Overview:

Main aim is to design an identity-based data storage scheme where multiple domains can co-exist and users can share files across domains. For simplicity, suppose that there are two domains: D1 and D2. At first, the private key generator PKGi in the domain Di generates his secret-public pair (ξi, (gi, hi)) where i € {1, 2}. Then, users in the domain Di authenticate themselves to the PKGi and obtain secret keys from PKGi. Prior to outsourcing the data, the data owner encrypts it under his identity ID. Then, the owner sends the ciphertext to the proxy server PS. PS validates the ciphertext. If it is computed correctly, PS stores it for the owner; otherwise, he rejects the ciphertext. Suppose that the PS can detect which domain the requester is from and the owner can know which file the requester wants to access from the partial ciphertext. If the requester wants to access a file, he can use his secret key $K_{ID'}$ to compute an authentication information (Q, F, $K_{ID'}$,3) and sends it to the PS. If the requester and the owner are in the same domain, the PS sends (ID′, Q, F , $K_{ID'}$,3, C2) to the owner, where C2 is the partial ciphertext. If the requester and the owner are in different domains, the PS sends (ID′, Q, F , $K_{ID'}$,3, (gi, hi)) to the owner. To resist the illegal requesters outside the cloud access the file, the owner validates the requester by verifying (Q, F , $K_{ID'}$,3). If the authentication is successful, the owner creates an access key (P1, P2, P3, K_{ID},2) and sends it to the PS. PS reencrypts the ciphertext using the access key and sends the reencrypted ciphertext to the requester. At the end, the requester can use his secret key to decrypt the re-encrypted ciphertext. In the inter-domain query, suppose that the owner is in the domain Di and the requester is in the domain D3–i, where i € {1, 2}. Actually, in proposed scheme, the owner in Di can use his secret key to generate an access key1 for the requester in D_{3-i}. Furthermore, the proxy server PSi can use the access key to transfer a ciphertext for the owner to a ciphertext for the requester.

Cloud Storage for Class-A

1. Kamara et al.'s scheme:

Kamara et al. designed secure cloud storage architectures for consumer scenarios by using non-standard cryptographic techniques, such as attribute decryption, searchable encryption, etc. Cryptographic cloud storage in consumer scenario, as shown in Fig.

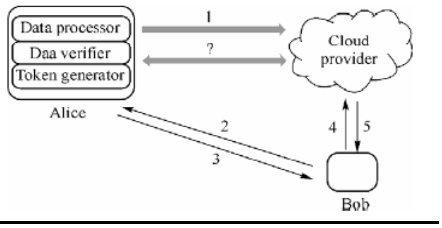

Fig1: Cryptographic cloud storage in consumer scenario

How to implement it using cryptographic techniques: When user wants to store data in cloud storage, the data processor firstly indexes data and encrypts it with a symmetric encryption scheme (e.g., advanced encryption standard (AES)) under a unique key. Then, the data processor respectively encrypts the index using a searchable encryption scheme and encrypts the unique key with an attribute-based encryption scheme under an appropriate policy. Finally, the data processor encodes the encrypted data and index such that the data verifier can later verify their integrity using a proof of storage

2. Kumbhare et al.'s scheme:

Kumbhare et al presented architecture, Cryptonite as shown in Fig. 5, for a secure data repository service design on top of a public cloud infrastructure. They use digital signature which is deployed in file manager and audit manager for the purposes of integrity verification as well as auditing purpose, broadcast encryption which is deployed in Cryptonite client library to distribute the file encryption and file signature keys, and searchable encryption which is deployed in the secure index manager to allow searching within an encrypted file without decrypting the entire file or revealing the contents of the file to the searching entity.

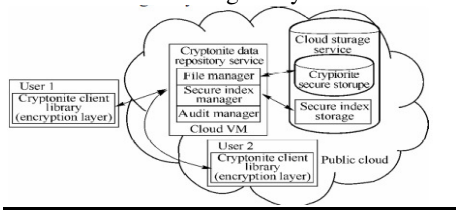

Fig 2: Architecture diagram of the Cryptonite secure data repository and client libraries

B) Cloud storages of Class-B

1. Kamara et al.'s scheme:

By using symmetric searchable encryption (SSE), search authenticators and proofs of storage (PoS), Kamara et al implemented a cloud storage system CS2 from the cryptographic standpoint. For cryptography researcher, this work is very suitable for them as a reference to begin study cloud storage. CS2 can ensure confidentiality, integrity, and verifiability without sacrificing utility. In this system, SSE is used to encrypt data by a client such that it can later generate search tokens for a storage provider, and search authenticator is used by the cloud provider to prove to the client that it returned all and only the correct files. Moreover, PoS is used to guarantee the integrity of CS2.

2 .Chow et al.'s scheme

Similar to Kamara et al.'s work, Chow et al also implemented secure cloud storage with properties of data provenance and support of dynamic users from the cryptographic standpoint. The authors not only implemented a cryptographic system model for secure cloud storage, but give a security model based on cryptography theory which includes confidentiality, anonymity, and traceability. The security of Chow et al.'s scheme can be proven under some number-theoretical assumptions in the framework of provable-security theory in cryptography field. Moreover, Chow et al.'s scheme is a pairing-based cryptographic cloud storage and designed using VLR (verifier-local revocation) group signature scheme and a variant of identity-based broadcast encryption.

VI. PROPOSED METHODOLOGY

As Data storage is one of the most prominent cloud applications, with which individuals can store data online and companies can backup local data to the cloud, there are many chances that these sensitive data may be corrupted. So to provide the security many schemes and algorithms are used like RSA, AES etc. Out of which REA is most secure and efficient one. REA is a new encryption algorithm restating its functions over other similar encryption algorithm. REA algorithm limits the added time cost for encryption and decryption so as to not degrade the performance of a database system.

Steps of the encryption algorithm of the Reverse Encryption Algorithm REA (Algorithm 1). The following are the steps:

Step1: Input the text and the key.
Step2: Add the key to the text.
Step3: Convert the previous text to ASCII code.
Step4: Convert the previous ASCII code to binary
Step5: Reverse the previous binary data.
Step6: Gather each 8 bits from the previous binary data and obtain the ASCII code from it.
Step7: Divide the previous ASCII code by 4.
Step8: Obtain the ASCII code of the previous result divide and put it as one character.
Step9: Obtain the remainder of the previous divide and put it as a second character.
Step10: Return encrypted text.

VII. POSSIBLE OUTCOME AND RESULTS

As compared to the other encryption techniques like RDPC, IBPRE and privacy preserving public auditing scheme, REA i.e. Reverse Engineering algorithms is the most convenient and efficient method which provide data storage security. This new encryption algorithm REA limits the added time cost for encryption and decryption so as to not degrade the performance of the database system. The proposed encryption algorithm REA represents a significant improvement over the encrypted databases.

VIII. CONCLUSION

There is a lot of very important data in the cloud, which need to be protected from attack. Cryptographic support is an important mechanism of securing them. People, however, must tradeoff performance to ensure the security because the operation of encryption and decryption greatly degrades the performance. For the query types that require extra query processing over encrypted database, the cost differentials of query processing between non-encrypted and encrypted database increase linearly in the size of relations. To solve such a problem, the proposed encryption algorithm REA can implement over the encrypted database.

FUTURE SCOPE

In the future work, we can extend the proposed encryption algorithm REA in order to apply it to other kind of databases such as distributed DBMSs and object oriented DBMSs.

REFERENCES

[1] C. Blundo, V. Iovino, G. Persiano, Private-key hidden vector encryption with key confidentiality, in: The 8th International Conference on Cryptology and Network Security, CANS 2009, in: LNCS, vol. 5888, Springer Verlag, 2009, pp. 259–277

[2] T. Okamoto, K. Takashima, Hierarchical predicate encryption for innerproducts, in: Advances in Cryptology, ASIACRYPT 2009, in: LNCS, vol. 5912, Springer Verlag, 2009, pp. 214–231.

[3] Chun-I Fan, Shi-Yuan Huang, "Controllable privacy preserving search based on symmetric predicate encryption in cloud storage" , Elsevier Science Direct Future Generation computer System, VOL 29, NO. 7, PP 1716-1724., September 2013.

[4] Jinguang Han, Willy Susilo, Yi Mu, "Identity-based data storage in cloud computing", Elsevier ScienceDirect Future Generation computer System, VOL 29, NO. 3, PP .673-681, March 2013

[5] PENG Yong, ZHAO Wei, XIE FENG, DAI Zhong-hua, GAO Yang, CHEN Dong-qing, "Secure cloud storage based on cryptographic technique s," Journal of china Universities of post and Telecommunication, VOL 19, NO. 2, PP 182-189, October 2012.

[6] C.K. Chu, W.G. Tzeng, Identity-based proxy re-encryption without random oracles, in: J.A. Garay, A.K. Lenstra, M. Mambo, R. Peralta (Eds.), Proceedings: Information Security Conference, ISC 2007, in: Lecture Notes in Computer Science, vol. 4779, Springer-Verlag, Valparaso, Chile, 2007, pp. 189–202.