International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 6, June 2020||

Sensitive Information Hiding Of Data in Cloud Storage Using RC4 Algorithm

Athulya V S[1], Sankaranarayanan P N[2]

P.G. Student, Department of Computer Engineering, Government Engineering College, Idukki, Kerala, India[1]

Assistant Professor, Department of Computer Engineering, Government Engineering College, Idukki, Kerala, India[2]

ABSTRACT:Cloud storage is used generally to store data and realize the data sharing with others. Data that are stored to the cloud have to be secured in order to avoid the data from exploiting. Encrypting the whole shared file can be used for information hiding but will make this shared file unable to be used by others. So sensitive information hiding technique is used to hide only the sensitive information. Blinding and Sanitizing are the two methods used here. A RC4 algorithm is used here in order to reduce the complexity in the user side. At first the sensitive data is extract from the document and these data are hidden by using blinding algorithm and the output is given to the sanitizer. The sanitizer again unify the messy code and the blinded file is uploaded in to the Cloud. The data can be downloaded from the cloud by another users request. The user will request the cloud. Then a notification will be given to the sanitizer then the sanitizer will decrypt the blinded data of the organisation, the sensitive data will not be visible to the sanitizer. If only the user request the Doctor only the sensitive data will be displayed to the user. A comparison of computational overhead between the RC4 and blinding factor algorithm is also done.

KEYWORDS:Cloud, Encryption, Sanitizer, RC4 Algorithm, Sensitive Information Hiding.

I. INTRODUCTION

Cloud computing is a new Method that was created after grid computing, utility computing and distributed computing. The main idea of cloud computing is application hosting and service outsourcing etc. With cloud computing technology, users do not need to spend much on the maintanance and cost of the system. In addition, strong computing and storage space makes users are depending on the cloud to take care of difficult tasks.Cloud computing provides security,information storage , and computing power. Using of cloud computing, all information can be Accessed services anytime and anywhere.Cloud computing allows to store large amout of data and many services to the user. When storing large amount of data issues generated. The main issues is data security.

When we store a large number of information to the cloud platform, the cloud system becomes the cloud storage system.Cloud storage systems gives storage ability at low price, and gives a method for sharing data between people but According to the survey conducted on 2007 only 60 percent of the total companies uses this cloud computing platform due to privacy concerns.
The information put away in the cloud is frequently shared by utilizing the applications, for example, Google Drive, Dropbox and iCloud. Information sharing as one of the most utilized strategy in distributed storage, permits various clients to impart their information to other people. A strategy for taking care of this issue is to encode the whole document before sending it to the cloud, at last transfer this scrambled record. This can do the touchy data stowing away since just the information proprietor can decode this document. however, it will make the whole record not ready to be utilized by others. For instance, scrambling the EHRs of irresistible sickness patients can ensure the security of patient what's more, emergency clinic, however these encoded EHRs can't be adequately used by scientists any more. Disseminating the unscrambling key to the specialists is by all accounts a potential answer for the above issue. Be that as it may, it is infeasible to embrace this technique in genuine situations because of the accompanying reasons. Right off the bat,disseminating unscrambling key needs secure channels, which is difficult to be fulfilled in certain occurrences.

Integrity means message that has not been modified until it is reach to the authorized person or receiver. It means data is not modified, no delete and no insert at all. Availability means resources are available at all the time to the authorized user. It is guaranteed to provide the confidentiality between the client and data. In remote data integrity auditing schemes, the data owner firstly needs to generate signatures for data blocks before uploading them to the cloud.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 6, June 2020||

This method can realize the sensitive information hiding since only the data owner can decrypt this file. However, it will make the whole shared file unable to be used by others. For example, encrypting the EHRs of infectious disease patients can protect the privacy of patient and hospital, but these encrypted EHRs cannot be effectively utilized by researchers any more. Distributing the decryption key to the researchers seems to be a possible solution to the above problem. However, it is infeasible to adopt this method in real scenarios dueto the following reasons. Firstly, distributing decryption key needs secure channels, which is hard to be satisfied in some instances.

A cloud is mainly used for storage purpose, it is able to store a very large amount of data. But the companies nowadays don't use this cloud for storing because of the privacy concerns that the cloud faces. So this approach introduces a security while uploading data in to the cloud. Here the sensitive data such as Name, Age, Email ID, Weight etc from the original file and these data is blinded by using RC4 algorithm. These blinded data is send to sanitizer to for unifying the code inorder to rectify the messy data. Then the sanitizer will upload the data to the cloud. When a external user needs to download this file. It will send request to the cloud and the data will be retrieved only by the approval of sanitizer and the Doctor.

The rest of this paper is organized as follows: the related works are presented in Section II. In section III, the proposed solution is presented. In Section III (B), the implementation of the proposed model within a cloud storage application is developed. Experimental results are presented in Section IV. Section V concludes the research.

II. RELATED WORK

This section presents the related works regarding the encryption based approaches and other blinding techniques used.

A considerable amount of literature has been published on using encryption-based data hiding. There are many methods used for the security of data in cloud storage such as AES, DES, Attribute based encryption etc but the problem in encrypting the entire file is that it will be unable to use by the researchers. This method of solving this problem is to encrypt the whole shared file before sending it to the cloud, finally upload this encrypted file to the cloud. This method can realize the sensitive information hiding since only the data owner can decrypt this file. However, it will make the whole shared file unable to be used by others. Distributing the decryption key to the researchers seems to be a possible solution to the above problem. However, it is infeasible to adopt this method in real scenarios due to the following reasons. Firstly, distributing decryption key needs secure channels, which is hard to be satisfied in some instances. it is impractical to hide sensitive information by encrypting the whole shared file.

Manish Kothe,HarshalKarandikar,NikhilWani,SumitTamkhane[1] proposed Attribute-Based Encryption Method. Here an Attribute based encryption is a one-to-numerous encryption that permits clients to scramble and unscramble information dependent on client properties. Characteristic based encryption framework with unscrambling that to a great extent disposes of the decoding overhead for clients. There are two kinds of ABE, they are Cipher-Text Policy Attribute Based Encryption (CP-ABE) and Key Policy Attribute Based Encryption (KP-ABE). In the Cipher text policy, the data is encrypted as ciphertext and stored to the cloud. The limitations of this system is that resource limited devices the decryption is very expensive due to pairing operation.

Liang Liu and Jun Ye [2] proposed A Homomorphic Universal Re-encryptor for Identity-based Encryption Which uses identity-based proxy re-encryption(IBPRE) scheme. The re encryptor is the method of transforming cipher text under one public key in to new cipher text in another public key. The functional Homomorphic encryption is used to protect the master secret s. But the introduction of fully Homomorphic encryption decreases the efficiency of the system. An epic personality based intermediary re-encryption(IBPRE) plot which to the greatest degree diminishes the outstanding tasks at hand in the client side by conveying the re-encryption key (RK) age work to the intermediary server. It assumes a significant job in present day secure correspondence and data trade by means of different sorts of system framework.

Caihui Lan , Haifeng Li , Shoulin Yin, and Lin Teng [3] proposed a A New Security Cloud Storage Data Encryption Scheme Based on Identity Proxy Re-encryption which uses identity proxy re-encryption here Computational Overhead is present. When a data is shared with many users the scalability is weak. Plain text encryption scheme to cipher text security encryption scheme is used. The system consists of control all the other keys. Here secret shares are generated then deduplication is done and the hash valus are checked which is used to check the integrity and this hash values are uploaded in to the cloud.

Cong Wang, Sherman S.-M. Chow, Qian Wang,Kui Ren and Wenjing Lou[4] proposed Privacy-Preserving Public Auditing for Secure Cloud Storage, Here a homomorphic direct authenticator with arbitrary veiling procedure is proposed. In this paper a securely present a suitable TPA, the checking on technique should procure no new vulnerabilities towards customer data security, and familiarize no extra online load with customer. In this paper. plot is

the first to help open examining in the Cloud Computing.A open evaluating plan comprises of four calculations(KeyGen, SigGen ,Gen Proof ,Verify Proof). KeyGen is a key age calculation that is controlled by the client to arrangement the plan. Sig Gen is utilized by the client to create confirmation metadata, while Verify Proof is controlled by the TPA to review the evidence from the cloud server.

Mohammad Ubaidullah Bokhari and Qahtan Makki Shallal[5]proposed Hybrid Encryption Technique to Secure Data during Transmission in Cloud Computation. Here it utilizes half breed encryption and unscrambling process dependent on AES-128 and RSA calculation. This is mainly used for the integrity and Confidentiality. The limitation of this method is that Decryption and encryption time consumption is more. The plaintext is converted to cipher by using AES algorithm, the key used for this AES is got from RSA algorithm. And in order to check the integrity an HMAC algorithm is also used. Thus, paper have proposed a model to use a hybrid encryption and decryption process based on AES-128 and RSA algorithm.

Avinash Shukla, Sanjay Silakari and Uday Chourasia[6] proposed A Secure Data Storage over Cloud using ABE where Enhanced lockbox algorithm with more secure algorithm AES is performed. It has Less computation time and cost. The idea driving the exploration is taken a safe and dependable calculation, approach which can discover the answer for security just as de-duplication repetition enhancement over information. Distributed computing is a rising zone of research, where the vast majority of the IT framework is moving to make their administration and conveyance increasingly productive. Distributed computing make it progressively versatile, increasingly dependable, secure and open with a lot of choice to play out its best. In this paper our work approach leads behind the information proprietorship and security giving over the client meeting. The current method dependent on key age for the client meeting and further the augmentation is performed on Enhanced lockbox meeting idea for record sharing and the executives.

Swati V.Thakre, Prof. K.K.Chhajed and Prof.V.B.Bhagat [7] proposed Key Based Encryption Scheme for Secure Data Sharing on Cloud here a Key Based Encryption conspire is utilized and one of the to appropriate a solitary key to a client when offering bunches of records to the client. Conveyed stockpiling makes it functional for customers to remotely store their data and welcome the on demand choice cloud applications without the any weight of close by hardware and programming the board, which gloats a group great conditions like unfathomable amassing limit, wherever receptiveness, etc. Since Cloud figuring condition is developed on open structures and interfaces,it can possibly join different inward and outer cloud benefits together to give high interoperability.

Sowmya P , Revathi P , Dr. Thirumala Rao [8] proposed policy based data deduplication in cloud storage were The Deduplication techniques like file level block level and byte level data Deduplication mechanism is used. The Deduplication technology has gained more popularity in cloud storage system with the increase in growth rate of digital data. Deduplication techniques minimizes the redundant data by using data Deduplication techniques like file level, block level, byte level and identifies duplicate copy by calculating the hash values. Deduplication techniques minimizes the redundant data by using data Deduplication techniques like file level, block level, byte level and identifies duplicate copy by calculating the hash values. Now a day's maintaining the data reliability and the confidentiality to the stored data becomes a curial part in the cloud storage. Dekey Ramp Secret key and Dupless are used methods here. The deduplication is used in order to reduce the system traffic. A master key is generated and it is used to control all the other keys. Here secret shares are generated then deduplication is done and the hash valus are checked which is used to check the integrity and this hash values are uploaded in to the cloud.

Abiodun Esther Omolara [9] proposed Security and Verification of Data in Multi-Cloud Storage with Provable Data Possession were Cooperative Provable Data Possession Scheme is used and the advantage of this system is that This technique gives a strong proof of data integrity.Support Block less verification. The Encryption was said to provide security to data in the cloud. A cloud storage system stores large number of data in its storage server. Data that are stored to the cloud have to be secured in order to avoid the data from exploiting. There are several techniques and methods that are used to secure the data before uploading it to the cloud such as cryptographic encryption techniques with Sensitive Information Hiding for Secure Cloud Storage here sensitive data is only hided it does not use any encryption techniques, The hiding is done by using blinding and sanitizer. One of the advantage is that no encryption is used here. And it is more secure than any other above techniques.

Wenting Shen, Jing Qin, Jia Yu, Rong Hao, and Jiankun Hu, Senior Member[10] proposed Enabling Identity-Based Integrity Auditing and Data Sharingwith Sensitive Information Hiding for Secure Cloud Storage here sensitive data is only hided it does not use any encryption techniques, The hiding is done by using blinding and sanitizer. One of the

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 6, June 2020||

advantage is that no encryption is used here. And it is more secure than any other above techniques.Social affair customers register their new private keys SKID;RN by using the character key IDKID and the fragmentary key TKID;RN. The customer refusal is recognized by a key update methodology. The amount of customer denials RN expect a noteworthy activity in the key update. RN is a value set by the social affair boss, and besides known by pack customers and the cloud. Right when the structure is instated, set RN = 0. Exactly when customers are denied, set RN = RN + 1. The social affair chief delivers another mostly key contrasting with this new estimation of RN, and sends it to the total of the non-disavowed pack customers. By then the non-denied pack customers update their private keys using the new inadequate key. In this manner, the repudiated customer can't get the current private key related to the most state-of-the-art RN.

III. METHODOLOGY

The client right off the bat blinds the information squares relating to the individual touchy data of the record. At that point the client sends this blinded document to the sanitizer. In the wake of getting the message from the client, the sanitizer disinfects these blinded information squares and the information squares comparing to the association's delicate data. At last, the sanitizer sends this disinfected document to the cloud. The framework model includes four sorts of various elements: the cloud, the client, the sanitizer, the Private Key Generator (PKG). The PKG is a confided in substances, it is utilized in creating open parameters and private key.

The figure 3.1 shows the overall structure of the system. The RC4 algorithm are used to blind the data which occurs from extracting the sensitive information from the original file. At first PKG generates the public and private keys of the user. The user gives the the identity keys to the PKG and it generates the keys then the sensitive data is taken as using by user and RC4 algorithm is used to blind the data. These blinded data is given to the sanitizer. This sanitizer unify the the data and elimates the messy code. Then the sanitizer blindes the hospital data along with the original text file.

The files are uploaded to the cloud. When the clinical specialist needs the document, he sends a solicitation to the sanitizer. And afterward the sanitizer downloads the blinded record from the document data framework and sends it to the clinical specialist. At last, the clinical specialist recoups the first document from this blinded document. For the most part, these information squares are supplanted with special cases. Moreover, the sanitizer can change these information obstructs into legitimate ones for the cleaned document.

At the point when essential, the client can recuperate the unique document from the blinded one by utilizing this blinding element.The client sends the blinded document to the sanitizer. When the above messages from client are legitimate, the sanitizer initially cleans the blinded information hinders into a uniform arrangement and furthermore disinfects the information squares relating to the association's touchy data to ensure the protection of association,and afterward changes their relating marks into legitimate ones for disinfected record utilizing change esteem. At last, the sanitizer transfers the disinfected record. At that point the computational overhead of client is assessed by checking the general time of the client side.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 6, June 2020||

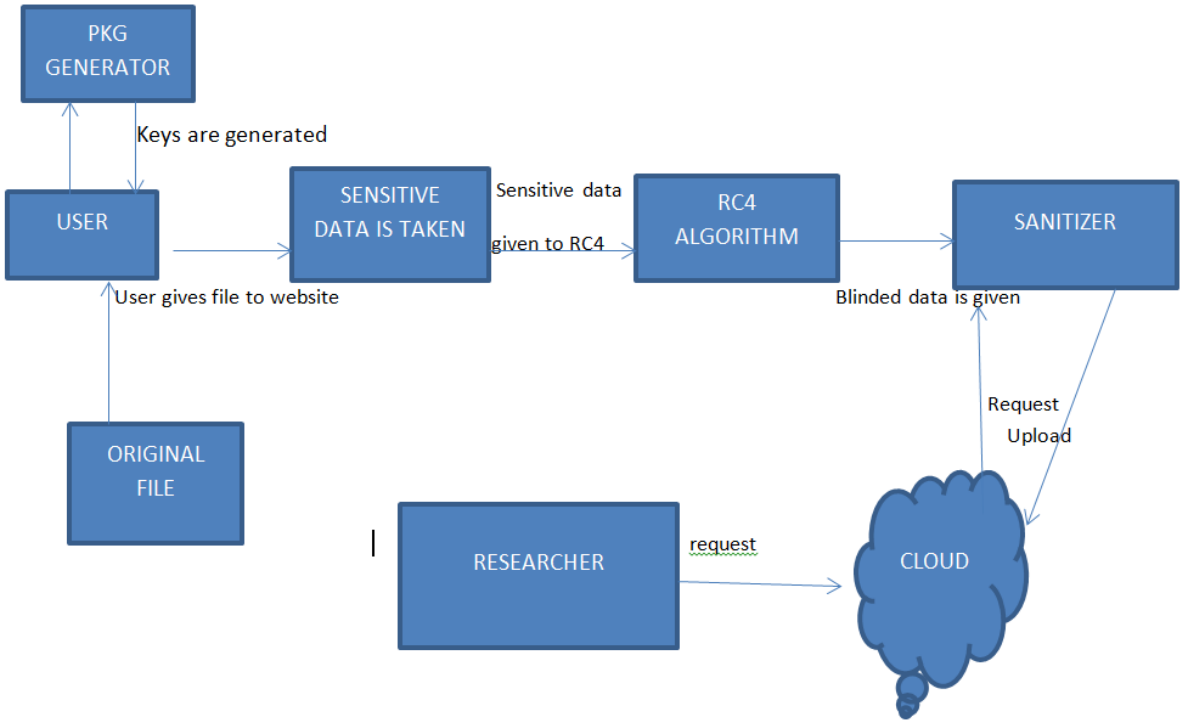

Fig 3.1 : A basic model

A. IMPLEMENTATION

The basic Sensitive Information Hiding system was developed using Python 2.0. The cloud Environment is used. A website is created by using Django platform . Doctor, Cloud , Sanitizer and other user are the main factors. A doctor can uploads his patients file in to the cloud. But while inserting the file, a PKG generator generates the patients private and public keys by taking the identities of the patients. This is created because of the diffie helman key exchange. The system takes the sensitive items which is provided by the user itself. Then the sensitive data is blinded by using the RC4 algorithm . The sensitive data is hided because encrypting the entire file will result in high computational overhead and also encrypting the whole file will result the researchers who need this data won't be able to get access of it. A decryption key will not be feasible enough. Along with the RC4 algorithm AES algorithm is also performed in parallel, The same data which is sensitive given by the user is taken by the AES algorithm and after performing both these algorithms the hided data along with the rest of the file is given to the sanitizer. The RC4 algorithm is having two stages such as key scheduling algorithm and also pseudo random generator algorithm .

The sanitizer is like an head of the hospital such as it is the one which also uploads the file in to the cloud and also helps the researcher to download the file. The sanitizer also uses this RC4 and AES algorithm to also hide the hospitals identity along with this. This is done to give high security while downloading the file. It is this sanitizer which uploads this data. Now if the Researcher needs a data . It will login to the website to the cloud which it have access and it request the cloud for the file. At this time a request is send to both the sanitizer and also the Doctor. Now if the Sanitizer login to the system to accept the request then the researcher will be able to see the file but the data which is the sensitive data will still be blinded And researcher will not be able to view the sensitive information. And when the Doctor login to the system and if the request is accepted then the researcher can see the data. MYSQL is used create the tables and also to save data to the website.

A related comparison between both the algorithm is also computed and it is marked as RC4 is much faster than AES algorithm. The computational Overhead of both the algorithm is conducted and it is noted that the time com plexity of the system which uses RC4 algorithm is much lesser than AES algorithm system.
Figure shows the flowchart showing how files are uploaded and how the researcher can download the files, then santizer and pkg generators work.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 6, June 2020||

IV. EXPERIMENTAL RESULTS

In this section, we compare the functionality comparison of our scheme and the scheme which uses the AES for blinding and the computation overhead comparisonbetween our scheme and the other one.

1. Functionality Comparison

We give the functionality comparison of our scheme with several related schemes [8, 9, 10]. Our scheme is the only scheme that can satisfy all of the followingproperties: data sharing and sensitive information hiding. Schemes [8, 9,10] all cannot support the sensitive information hiding.

2. Computational Overhead Comparison

The computational overhead in the user side is estimated. A comparative study between RC4 algorithm and AES algorithm is got as the result.The timeoverhead is calculated in millisecond. The computational time overhead of RC4 and AES from sending a message to the cloud, Downloading the file from the cloud and decrypting the file is included a variation of 2 sizes are considered here. And from the result it is able to see that the overhead of the RC4 algorithm is less than that of AES algorithm. In the figure the computational complexity of both Algorithms is given.

Schemes	Data Sharing and Sensitive Information Hiding	Sensitive Information Hiding(Time)	Total Computational Complexity(in ms)
Proposed Scheme with RC4 as Blinding Algorithm	YES	0.0041	0.055000ms
Proposed Scheme with AES as Blinding Algorithm	YES	0.0123	0.156000ms

Table4.1 :Comparison of both Algorithms

3. Performance Analysis

The computation complexity of different entities in our scheme respectivelydepends on the total number n of data blocks, the number d1 of data blocks corresponding to the personal sensitive information and the number d2 of data blocks corresponding to the organizations sensitive information. private key generation and private key verification spend nearly the same time, which are nearly 0.031ms. The time of sensitive information sanitization respectively is 0.0041ms. So we can conclude that in these processes, the sensitive information sanitization spends the shortest time.

V. CONCLUSION

In this paper, we proposed an identity-based data scheme for secure cloudstorage, which supports data sharing with sensitive information hiding. In our scheme, the file stored in the cloud can be shared and used by others on the condition that the sensitive information of the file is protected. The data is blinded using RC4 algorithm and given to the sanitizer and then the sanitizer encryptes the data again. And the blinded file is uploaded in to the cloud. The data can be downloadedfrom the cloud by another users request.The user will request the cloud. Then anotification will be given tothe sanitizer then the sanitizer will decrypt the blinded data of the organisation, the sensitive data will not be visible to the sanitizer. If only the user request the Doctor only the sensitive data will be displayed to the user. A comparison between the computational overheads of both RC4 algorithm and AES is considered. And the results it was able to find out that the Overhead of RC4 algorithm is less.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 6, June 2020||

REFERENCES

[1] Wenting Shen, Jing Qin, Jia Yu, Rong Hao, and Jiankun Hu, Senior Member Enabling Identity-Based Integrity Auditing and Data Sharing with SensitiveInformation Hiding for Secure Cloud Storage IEEE Transaction 2018.

[2] Abiodun Esther Omolara Security and Verification of Data in Multi-Cloud Storage with Provable Data Possession 2015,International Journal of ComputerApplications.

[3]Caihui Lan , Haifeng Li , Shoulin Yin, and Lin Teng. A New Security Cloud Storage Data Encryption Scheme Based on Identity Proxy Re-encryption 2016, International Journal of Network Security

[4] Swati V.Thakre, Prof. K.K.Chhajed and Prof.V.B.Bhagat Key Based EncryptionScheme for Secure Data Sharing on Cloud , International Research Journalof Engineering and Technology.

[5] Avinash Shukla, Sanjay Silakari and Uday Chourasia A Secure Data Storageover Cloud using ABE Approach 2017,International Journal of ComputeApplications.

[6] Mohammad Ubaidullah Bokhari and Qahtan Makki Shallal Hybrid EncryptionTechnique to Secure Data during Transmission in Cloud Computation2017,International Journal of Computer Applications

[7] Cong Wang, Sherman S.-M. Chow, Qian Wang,Kui Ren and WenjingLouPrivacy-Preserving Public Auditing for Secure Cloud Storage 2013,IEEETransactions on Computers.

[8] H. Shacham and B. Waters, Compact proofs of retrievability, J. Cryptology,vol. 26, no. 3, pp. 442483, Jul. 2013.

[9] Liang Liu and Jun Ye HoneyGen: A Homomorphic Universal Re-encryptorfor Identity-based Encryption, 2016,International Journal of Network Security

[10] ManishKothe,HarshalKarandikar,NikhilWani,SumitTamkhane Attribute-Based Encryption with Verifiable Outsourced Decryption 2016,InternationalResearch Journal of Engineering and Technology

[11] C. Erway, C. Papamanthou, and R. Tamassia, Dynamic provable data possession,in ACM Conference on Computer and Communications Security, 2009,pp. 213222.

[12] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, Enabling public auditabilityand data dynamics for storage security in cloud computing, IEEE Transactions on Parallel and Distributed Systems, vol. 22.