ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A COMBINED CLOUD ARCHITECTURE FOR PREVENTION OF DUPLICATION AND ANONYMOUS USER AUTHENTICATION

S.Sutar*, A.V.Survase, N.K.Malvadkar, M.V.Patil, P.V.Raut
Computer Department, Dr.D.Y.Patil Collage of Engineering, India.

ABSTRACT

Cloud computing is a rising computing standard in which assets of the computing framework are given as a service over the Internet. As guaranteeing as it may be, this standard additionally delivers a lot of people new challenges for data security and access control when clients outsource sensitive data for offering on cloud servers, which are not inside the same trusted dominion as data possessors. In any case, in completing thus, these results unavoidably present a substantial processing overhead on the data possessor for key distribution and data administration when fine-grained data access control is in demand, and subsequently don't scale Itll. The issue of at the same time accomplishing fine-grainedness, scalability, and data confidentiality of access control really still remains uncertain. This paper addresses this open issue by, on one hand, characterizing and implementing access policies based on data qualities, and, then again, permitting the data owner to representative the majority of the calculation undertakings included in fine-grained data access control to un-trusted cloud servers without unveiling the underlying data substance. It accomplish this goal by exploiting and combining techniques of decentralized key policy Attribute Based Encryption (KP-ABE). Extensive investigation shows that the proposed approach is highly efficient and secure. Data deduplication is one of important data compression techniques for eliminating duplicate copies of repeating data, and has been widely used in cloud storage to reduce the amount of storage space and save bandwidth. To protect the confidentiality of sensitive data while supporting deduplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing It also presents several new deduplication constructions supporting authorized duplicate check in a hybrid cloud architecture.

KEYWORDS: Cloud Storage, Access control, Key Distribution Center, Data Deduplication.

INTRODUCTION

Cloud computing is increasing computing standard in which resources of the computing framework are given as a service over the Internet. Cloud computing provides “virtualized” assets to users as services across the whole Internet, while hiding details of policy and implementation as Itll as platform. In cloud computing, users can farm out their calculation and storage to servers using Internet. The data stored most of in clouds is highly responsive, for example, medical records and social networks. Thus security and privacy are very essential issues in cloud computing.

Basically, the user should verify itself before initiating any transaction, and then it must be ensured that the cloud does not interfere with the data that is outsourced. To avoid the identification of the user from cloud or other user, the requirement of user confidentiality is must. The cloud can hold the user responsible for the data it outsources, and the cloud is itself responsible for the services it provides. The validity of the user who stores the data is also verified.

Cloud computing provides seemingly unlimited “virtualized” resources to users as services across the whole Internet, while hiding platform and implementation details. Today’s cloud service providers offer both highly available storage and massively parallel computing resources at relatively low costs. As cloud computing becomes prevalent, an increasing amount of data is being stored in the cloud and shared by users with specified *privileges*, which define the access rights of the stored data. One critical challenge of cloud storage services is the management of the ever-increasing volume of data.

Access control in clouds is very important because it gives permission to only authorized users to have access to suitable service. A huge amount of information is being stored in the cloud, and much of this is responsive information. Access control is also in advance important in social networking where users store their personal information, and

share them with selected faction of users, this data are being accumulate in clouds. It is very important that only the authorized users are given access to that information.

To make data management scalable in cloud computing, deduplication has been a Itll-known technique and has attracted more and more attention recently. Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data in storage. The technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. Instead of keeping multiple data copies with the same content, deduplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy. Deduplication can take place at either the file level or the block level. For file level deduplication, it eliminates duplicate copies of the same file. Deduplication can also take place at the block level, which eliminates duplicate blocks of data that occur in non-identical files.

Data deduplication is one of significant data compression techniques for removing duplicate copies of replicated data, and has been broadly used to reduce the amount of storage space and save bandwidth in cloud storage. The protection of the confidentiality of responsive data while supporting deduplication is done by using the convergent encryption technique that has been proposed to secure the data before outsourcing. Different from traditional deduplication systems, the differential privileges of users are further considered in duplicate check besides the data itself. The technique is used to get better storage consumption and can also be applied to network data transfers to reduce the number of bytes that must be sent. Instead of keeping multiple data copies with the same content, deduplication removes unnecessary data by keeping only one physical copy and referring other unnecessary data to that copy. The place taken deduplication is at either the block level or the file level. Deduplication can take place at the block level, in the occurrence of non-identical files where it removes duplicate blocks of data. It removes duplicate copies of the same file, for file level deduplication. Due to this they required minimum space and contained more advantages over the system.

LITERATURE SURVEY

In this system, It first define the notations used in this paper, review some secure primitives used in our secure deduplication.

Privacy Preserving Access Control with Authentication for Securing Data in Clouds"S. Ruj, M. Stojmenovic, and A. Nayak Proc.IEEE/ACM Int'l Symp. Cluster, Cloud and Grid Computing, pp. 556-563, 2012." they propose a new privacy preserving authenticated access control scheme for securing data in clouds. In the proposed scheme, the cloud verifies the authenticity of the user without knowing the user's identity before storing information. Our scheme also has the added feature of access control in which only valid users are able to decrypt the stored information. The scheme prevents replay attacks and supports creation, modification, and reading data stored in the cloud. Moreover, our authentication and access control scheme is decentralized and robust, unlike other access control schemes designed for clouds which are centralized. The communication, computation, and storage overheads are comparable to centralized approaches.

Toward Secure and Dependable Storage Services in Cloud Computing"C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou IEEE Trans. Services Computing, vol. 5, no. 2, pp. 220-232, Apr.- June 2012." Cloud storage enables users to remotely store their data and enjoy the on-demand high quality cloud applications without the burden of local hardware and software management. Though the benefits are clear, such a service is also relinquishing users' physical possession of their outsourced data, which inevitably poses new security risks toward the correctness of the data in cloud. In order to address this new problem and further achieve a secure and dependable cloud storage service, It propose in this paper a flexible distributed storage integrity auditing mechanism, utilizing the homomorphic token and distributed erasure-coded data. The proposed design allows users to audit the cloud storage with very lightItight communication and computation cost. The auditing result not only ensures strong cloud storage correctness guarantee, but also simultaneously achieves fast data error localization, i.e., the identification of misbehaving server. Considering the cloud data are dynamic in nature, the proposed design further supports secure and efficient dynamic operations on outsourced data, including block modification, deletion, and append. Analysis shows the proposed scheme is highly efficient and resilient against Byzantine failure, malicious data modification attack, and even server colluding attacks.

A Secure Cloud Backup System with Assured Deletion and Version Control "P. Anderson and L. Zhang. Fast and secure laptop backups with encrypted de-duplication. In *Proc. of USENIX LISA*, 2010".,It present the design and implementation of Fade Version, a system that provides secure and cost effective backup services on the cloud. Fade Version is designed for providing assured deletion for remote cloud backup applications, while allowing version control of data backups. It use a layered encryption approach to integrate both version control and assured deletion into one design. Through system prototyping and extensive experiments, It justify the performance overhead of Fade Version in terms of time performance, storage space, and monetary cost. It note that the main performance overhead of Fade Version is the additional storage of cryptographic keys in data backups. In future work, It explore possible approaches of minimizing the number of keys to be stored and managed. In this section, It first define the notations used in this paper, review some secure primitives used in our secure deduplication. Cloud storage is an emerging service model that enables individuals and enterprises to outsource the storage of data backups to remote cloud providers at a low cost. HoItver,cloud clients must enforce security guarantees of their out sourced data backups. It present *Fade Version*, a secure cloud backup system that serves as a security layer on top of today's cloud storage services. Fade Version follows the standard version-controlled backup design, which eliminates the storage of redundant data across different versions of backups. On top of this, Fade Version applies cryptographic protection to data backups. Specifically, it enables fine-grained assured deletion, that is, cloud clients can assuredly delete particular backup versions or files on the cloud and make them permanently inaccessible to anyone, while other versions that share the common data of the deleted versions or files will remain unaffected. It implement a proof-of-concept prototype of Fade Version and conduct empirical evaluation atop Amazon S3. It show that Fade Version only adds minimal performance overhead over a traditional cloud backup service.

Message-Locked Encryption for Lock-Dependent Messages M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In *EUROCRYPT*, pages 296–312, 2013[14],Prior dentitions and schemes for message-locked encryption (MLE) admit onlyan adversary who is oblivious to the scheme's public parameters during the initial interaction. It explore two avenues for extending security guarantees of MLEtowards a more poItrful adversarial model, where the distribution of plain text scan be correlated with the scheme's parameters (lock-dependent messages). In our first construction It augment the dentition of MLE to allow fully random cipher texts by supporting equality-testing functionality. One challenging aspect of the construction is ensuring cipher text consistency in the presence of random oracles without in acting the length of the cipher text. It achieve this goal via combination of a cut-and-choose technique and NIZKs. The resulting scheme is secure against a fully adaptive adversary. Our second construction assumes a predetermined bound on the complexity of distributions specified by the ad-versary. It the original framework of deterministic MLE while satisfying a stronger security notion. It formulate the following several directions for further research. First, It ask whether a fully adaptive randomized MLE2 can be constructed and proven secure in the standard model. Second, a randomized scheme for deduplication creates a potential leakage channel that allows one user to test whether her plain-text has already been uploaded to the system (similar to the attack described by Harnik et all. where the deduplication event was observable via traffic anal-ysis). Designing a scheme resistant to this attack, for example, by supporting server-side rerandomization of cipher texts, constitutes an interesting research question. Note that deterministic MLEs are immune to this problem. Finally, our first scheme requires a pair wise application of the equality-testing algorithm to identify all duplicate cipher texts, and uses computationally expensive NIZK as a building block. It leave reducing the overhead of the scheme as an open problem.

GQ and Schnorr Identication Schemes: M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. *J. Cryptology*, 22(1):1–61, 2009, The Guillou-Quisquater (GQ) and Schnorr identication schemes are amongst the most ecient and best-known Fiat-Shamir follow-ons, but the question of whether they can be proven secure against impersonation under active attack has remained open. This paper provides such a proof for GQ based on the assumed security of RSA under one more inversion, an extension of the usual one-wayness assumption that was introduced in. It also provides such a proof for the Schnorr scheme based on a corresponding discrete-log related assumption. These are the first security proofs for these schemes under assumptions related to the underlying one-way functions. Both results extend to establish security against impersonation under concurrent attack.

EXISTING SYSTEM

Generally alive work based on access control in cloud is centralized in nature.ABE use by all schemes. It makes the use of symmetric key there is no need of authentication. It provides privacy preserving valid access control in cloud. HoItver, the author uses a centralized approach where a single key distribution center (KDC) distributes secret keys and attributes to all users.

Deduplication of data is one of important data solidity techniques for eliminating replica copies of repeating data, and it also beneficial in the cloud to reduce the amount of storage space and save bandwidth. Deduplication of data systems, the classified cloud is involved as a proxy to allow data users to securely perform duplicate verify with disparity privileges.

Deduplication systems cannot support *differential authorization duplicate check*, which is important in many applications. In such an authorized deduplication system, each user is issued a set of privileges during system initialization Each file uploaded to the cloud is also bounded by a set of privileges to specify which kind of users is allotted to perform the duplicate check and access the files. Before submitting his duplicate check request for some file, the user needs to take this file and his own privileges as inputs. The user is able to find a duplicate for this file if and only if there is a copy of this file and a matched privilege stored in cloud. In order to save cost and efficiently management, the data will be moved to the storage server provider (SCSP) in the public cloud with specified privileges and the deduplication technique will be applied to store only one copy of the same file. Because of privacy consideration, some files will be encrypted and allotted the duplicate check by employees with specified privileges to realize the access control. Traditional deduplication systems based on convergent encryption, although providing confidentiality to some extent, do not support the duplicate check with differential privileges. In other words, no differential privileges have been considered in the deduplication based on convergent encryption technique. It seems to be contradicted if It want to realize both deduplication and differential authorization duplicate check at the same time.

Disadvantages of Existing system

- The system in uses asymmetric key approach and it does not support for authentication.
- Cloud environment contains the large number of user so it is difficult to maintain.
- Traditional encryption, while providing data privacy, is incompatible with deduplication of data.
- The same data copies of dissimilar users will lead to different Ciphertexts, making deduplication impossible.

PROPOSED SYSTEM

To give better data security, this system makes the first attempt to formally address the problem of authorized deduplication of data.[1]It proposed the system that verifies the validity of the series without knowing the user's identity before storing data. In this schema, it also include feature of access control in which only valid users are able to decrypt the stored information. It also prevents replay attacks and supports formation adaptation, and evaluation data stored in the cloud and also address user reversal. It proposed a fully dispersed ABE where users could have one or more attributes from each authority and need not require a trusted server. To get over this problem, the decryption task to a swap server, so that the user can compete with smallest resources.

Although data deduplication brings a lot of benefits, security and privacy concerns arise as users' sensitive data are susceptible to both inside and outside attacks. Traditional encryption, while providing data confidentiality, is incompatible with data deduplication. Specifically, traditional encryption requires different users to encrypt their data with their own keys. Thus, identical data copies of different users will lead to different cipher texts, making deduplication impossible. Convergent encryption [8] has been proposed to enforce data confidentiality while making deduplication feasible. It encrypts/ decrypts a data copy with a *convergent key*, which is obtained by computing the cryptographic hash value of the content of the data copy. After key generation and data encryption, users retain the keys and send the cipher text to the cloud. Since the encryption operation is deterministic and is derived from the data content, identical data copies will generate the same convergent key and hence the same cipher text. To prevent unauthorized access, a secure proof of ownership protocol [11] is also needed to provide the proof that the user indeed owns the same file when a duplicate is found. After the proof, subsequent users with the same file will be provided a pointer from the server without needing to upload the same file. A user can download the encrypted file with the

pointer from the server, which can only be decrypted by the corresponding data owners with their convergent keys. Thus, convergent encryption allows the cloud to perform deduplication on the Ciphertexts and the proof of ownership prevents the unauthorized user to access the file.

The KP-ABE is a public key cryptography primal for more than one correspondence. In KP-ABE, [2] information is associated with attributes for each of which a public key part is characterized. The encryption authority associates the set of attributes to the message by scrambling it with the comparing public key parts. Every client is assigned an access structure which is normally characterized as an access tree over information attributes, i.e., inside hubs of the access tree are limit doors and leaf hubs are connected with attributes. Client secret key is characterized to reflect the access structure so the client has the ability to decode a cipher-text if and just if the information attributes fulfill his access structure. The convergent encryption technique has been proposed to encrypt the data before outsourcing. To have a better data security, the problem of authorized data deduplication is introduced in the proposed system.Different privileges of users are considered in replica check as Itll the data itself. This Schema represent some new deduplication sustaining authorized duplicate check in combined cloud . The scheme is secure in requisites of the definitions specified in the proposed security model.[5] It apply a model of this proposed authorized replica check .In this system ,it proposed authorized duplicate check scheme incurs minimal overhead compared to normal operations.

Advantages of Proposed system

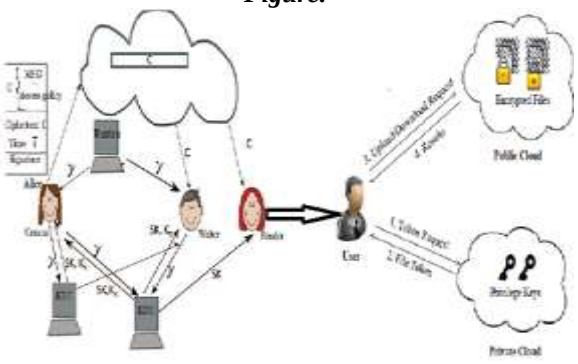
- This system proposed a new dispersed access control scheme for secure data storage in clouds that supports unspecified authentication.
- The cloud verifies the validity of the users without knowing the user's identity before storing data.
- The system also has the feature of access control in which only legal users are able to decrypt the stored information.
- The system prevents replay attacks and supports development, variation, and evaluation data stored in the cloud.
- The uniqueness of the user is protected from the cloud during validation.

Modules

The proposed system consists of the following modules:

1. System Initialization
2. User Registration
3. KDC setup
4. Attribute generation
5. Sign
6. Verify

Figure:



System Architecture

Owner Module

Owner module is to upload their files using some access policy. First they get the public key for particular upload file after getting this public key owner request the secret key for particular upload file. Using that secret key owner upload their file.

User Module

This module is used to help the client to search the file using the file id and file name .If the file id and name is incorrect means It do not get the file, otherwise server ask the public key and get the encryption file. If u want the decryption file means user have the secret key.

Distributed key policy attribute based encryption

KP-ABE is a public key cryptography primitive for one-to-many correspondences. In KP-ABE, information is associated with attributes for each of which a public key part is characterized. The encrypted associates the set of attributes to the message by scrambling it with the comparing public key parts. Every client is assigned an access structure which is normally characterized as an access tree over information attributes, i.e., inside hubs of the access tree are limit doors and leaf hubs are connected with attributes. Client secret key is characterized to reflect the access structure so the client has the ability to decode a cipher-text if and just if the information attributes fulfill his access structure. The proposed scheme consists of four algorithms which is defined as follows.

Set up

This algorithm takes as input security parameters and attribute universe of cardinality N. It then defines a bilinear group of prime number. It returns a public key and the master key which is kept secret by the authority party.

Encryption

It takes a message, public key and set of attributes. It outputs a cipher text.

Key Generation

It takes as input an access tree, master key and public key. It outputs user secret key.

Decryption

It takes as input cipher text, user secret key and public key. It first computes a key for each leaf node. Then it aggregates the results using polynomial interpolation technique and returns the message.

File Assured Deletion

The policy of a file may be denied under the request by the customer, when terminating the time of the agreement or totally move the files starting with one cloud then onto the next cloud nature's domain. The point when any of the above criteria exists the policy will be repudiated and the key director will totally evacuates the public key of the associated file. So no one can recover the control key of a repudiated file in future. For this reason It can say the file is certainly erased. To recover the file, the user must ask for the key supervisor to produce the public key. For that the user must be verified. The key policy attribute based encryption standard is utilized for file access which is verified by means of an attribute connected with the file. With file access control the file downloaded from the cloud will be in the arrangement of read just or write underpinned. Every client has connected with approaches for each one file. So the right client will access the right file. For making file access the key policy attribute based encryption.

Secure Deduplication System

To support authorized deduplication, the tag of a file F will be determined by the file F and the privilege. To show the difference with traditional notation of tag, It call it file token instead. To support authorized access, a secret key kp will be bounded with a privilege p to generate a file token. Let $\phi' F;p$ = TagGen(F, kp) denote the token of F that is only alloItd to access by user with privilege p. In another word, the token $\phi' F;p$ could only be computed by the users with privilege p. As a result, if a file has been uploaded by a user with a duplicate token $\phi' F;p$, then a duplicate check sent from another user will be successful if and only if he also has the file F and privilege p. Such a token generation function could be easily implemented as $H(F, kp)$, where $H(\_)$ denotes a cryptographic hash function.

Security of Duplicate Check

It consider several types of privacy It need protect, that is, enforceability of duplicate-check token: There are two types of adversaries, that is, external adversary and internal adversary. As shown below, the external adversary can be vie It as an internal adversary without any privilege. If a user has privilege p, it requires that the adversary cannot forge and output a valid duplicate token with any other privilege p' on any file F, where p does not match p'. Furthermore, it also requires that if the adversary does not make a request of token with its own privilege from private cloud server, it cannot forge and output a valid duplicate token with p on any F that has been queried.

Send Key

Once the key request was received, the sender can send the key or he can decline it. With this key and request id which was generated at the time of sending key request the receiver can decrypt the message.

CONCLUSION

It presented a Hybrid access control technique with unknown authentication, which provides user revocation and prevents replay attacks. The cloud does not know the individuality of the user who stores information, but only verifies the user's important data. Key distribution is done in a hybrid way. In future, it hides the attributes and access policy of a user.

One limitation is that the cloud knows the access policy for each verification stored in the cloud. Here furthermore it given many new deduplication constructions supporting approved duplicate sign up hybrid cloud design, during which the duplicate-check tokens of files as generated by the personal cloud server with personal keys. Refuge analysis demonstrates that our scheme is vulnerable in terms of business executive and outsider attacks laid out in the planned security model.

REFERENCES

[1] Sushmita Ruj, Milos Stojmenovic, and Amiya Nayak, "Decentralized Access Control with Anonymous Authentication of Data Stored in Clouds", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 2, FEBRUARY 2014

[2] H.K. Maji, M. Prabhakaran, and M. Rosulek, "Attribute-Based Signatures: Achieving Attribute-Privacy and Collusion-Resistance," IACR Cryptology ePrint Archive, 2008.

[3] M. Green, S. Hohenberger, and B. Waters, "Outsourcing the Decryption of ABECiphertexts," Proc. USENIX Security Symp., 2011.

[4] H. Li, Y. Dai, L. Tian, and H. Yang, "Identity-based authentication for cloud computing," in *CloudCom*, ser. Lecture Notes in Computer Science, vol.5931. Springer, pp. 157–166, 2009.

[5] S. Ruj, M. Stojmenovic and A. Nayak, "Privacy Preserving Access Control with Authentication for Securing Data in Clouds", *IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 556–563, 2012.

[6] C. Wang, Q. Wang, K. Ren, N. Cao and W. Lou, "Toward Secure and Dependable Storage Services in Cloud Computing", *IEEE T. Services Computing*, vol.5, no. 2, pp. 220–232, 2012.

[7] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in *IEEE INFOCOM.* , pp.441–445, 2010.