International Journal of Innovative Research in Engineering & Management (IJIREM)
ISSN: 2350-0557, Volume-5, Issue-1, January-2018
DOI: 10.21276/ijirem.2018.5.1.9

Security and Privacy Based Data Sharing in Cloud Computing

Dr.Prasanna Kumar.R
Professor, Department Of Computer Science And Engineering
S.A. Engineering College, Chennai, Tamil Nadu,India

Porselvan G
UG Student, Department Of Computer Science And Engineering
S.A. Engineering College, Chennai, Tamil Nadu,India

Prem Kumar S
UG Student, Department Of Computer Science And Engineering
S.A. Engineering College, Chennai, Tamil Nadu,India

Robinlash F
UG Student, Department Of Computer Science And Engineering
S.A. Engineering College, Chennai, Tamil Nadu,India

ABSTRACT

Data sharing in cloud storage is playing major role in Information Communication, since it can provide users with efficient and effective storage services. In order to provide security to the shared secret data, the cryptographic techniques are usually applied. However, the data cryptographic key is protected by the two factors. Only if both the factors works, the secrecy of the cryptographic key is held. 2) The cryptographic key can be revoked efficiently by integrating the proxy re-encryption and key separation techniques. 3) The data is protected in a fine-grained way by adopting the attribute based encryption technique. Furthermore, the security analysis and performance evaluation show that our proposal is secure and efficient, respectively.

Keywords: Cloud computing, privacy, security, attribute based key, encryption, and decryption

1. INTRODUCTION

Driven by the attractive on-demand features and advantages, the development and deployment of cloud-based applications have gained tremendous impetus in the industry and research community in recent years. Cloud storage is one of the most successful cloud-based applications, since it matches the huge data sharing demand quite well. Sharing huge data with several data sharers is a cost-consuming task, and the cost on the data owner side is usually proportional to the number of data sharers. While this cost could be reduced to the size of shared data with the help of cloud storage. The only thing the data sharer needs to do is to upload the data to the cloud and grant the access right to the data sharer. After that, data sharers can obtain the data from the cloud instead of the data owner. Despite the benefits of data sharing in cloud storage, it also introduces many chances to the adversary to access the shared data without authorization. To protect the confidentiality of the shared data, the cryptographic schemes are usually applied. The security of cryptographic schemes stem from the security of underlying cryptographic key. Currently, the cryptographic key is simply stored in the computer in most of existingcryptographic schemes. While it has been reported that the stored keys can be revealed by some viruses .To deal with the key exposure problem, many techniques have been proposed, such as key-insulated public key technique, and parallel key insulated public key technique. To the best of our knowledge, the cryptographic key exposure and revocation problems in cloud storage are unrevealed till the work by Liu et al. named LLS+15 afterwards). In, they proposed a novel two-factor data protection mechanism. The cryptographic key is divided into two parts. One is kept in user's computer and the other is stored in a security device (e.g., smart card), which is similar to the e-banking. Only if one of these two parts are kept secret from the adversary, the confidentiality of the cryptographic key is held. Hence, the "two-factor" is named. Furthermore, once the user's security device was either lost or stolen, it could be revoked by using the proxy re-encryption technique. While LLS+15 aims to solve the security problem of the data storage but not the data sharing scenario in cloud computing. Especially, one ciphertext in LLS+15 is essentially an identity-based ciphertext that can be decrypted by only one user but not a group of users as in data sharing scenario. Recently, the data sharing is rising a heated concern. While privacy is still the key concern and an equally striking challenge that reduce the growth of data sharing in cloud

2. CLOUD SECURITY TECHNIQUES

2.1 Delay-Optimized File Retrieval under LT-Based Cloud Storage

Luby Transform (LT) code is one of the popular fountain codes for storage systems due to its efficient recovery. In this paper, it is that multiple stage retrieval of fragments is effective to reduce the file-retrieval delay. In this first develop a delay model for various multiple stage retrieval schemes applicable to the considered system. This paper, it is focused on the file-retrieval delay, defined as the duration between the time for the portal receiving an LT-coded file request and the time when the last LT-coded packet is sent out by the portal. The file-retrieval delay is a good indicator of user experience. We formulate a delay-optimal file-retrieval problem, which aims to minimize the retrieval delay by strategically scheduling packet retrieval requests. Therefore, we aim to reduce the file-retrieval delay by strategically scheduling the LT-coded packet requests. In our proposed multi-stage request scheme, the designing objective is to minimize the average file-retrieval[1] delay, for a given number of stages. The issue with this protocol is that the problem of delay optimal file-retrieval under a distributed cloud storage system is rectified. Using this model, we derived an optimal two-stage request scheme for a given decoding probability. Both simulation and numerical results confirm that this optimal scheme can reduce the average delay dramatically. The analysis offers a way for storage system operators to design an optimized storage retrieval scheme for LT-based distributed cloud storage systems. Using this model, we derived an optimal two-stage request scheme for a given decoding probability. Both simulation and numerical results confirm that this optimal scheme can reduce the average delay dramatically.

2.2 Quick Sync: Improving Synchronization Efficiency for Mobile Cloud Storage Services

Mobile cloud storage services have gained great success in recent few years. In this paper, identify, analyse, and address the synchronization inefficiency problem of modern mobile cloud storage services. The results demonstrate that existing commercial sync services fail to make full use of available bandwidth, and generate a large amount of unnecessary sync traffic in certain circumstances even though the incremental sync is implemented. Based on the findings, it is proposed Quick Sync[2], a system with three novel techniques to improve the sync efficiency for mobile cloud storage services, and build the system on two commercial sync services. We further evaluate the capability of the Batched Synced in improving the bandwidth utilization efficiency. Finally the overall improvement of the sync efficiency using real-world workloads. If we compare the performances of the original Sea file and Drop box clients with those when the two service frameworks are improved with Quick Sync To address the inefficiency issues, it is proposed a Quick Sync, a system with three novel techniques. Quick Sync to support the sync operation with Drop box and Sea file. Our extensive evaluations demonstrate that Quick Sync can effectively save the sync time and reduce the significant traffic overhead for representative sync workloads

2.3 Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage

Dynamic hash table[3], which is a new two-dimensional data structure located at a third parity auditor to record the data property information for dynamic auditing. Differing from the existing works, the proposed scheme migrates the authorized information from the CSP to the TPA, and thereby significantly reduces the computational cost and communication overhead. To support privacy preservation by combining the homomorphism authenticator based on the public key with the random masking generated by the TPA, and achieve batch auditing by employing the aggregate BLS signature technique. The proposed scheme can effectively achieve secure auditing for cloud storage, and outperforms the previous schemes in computation complexity, storage costs and communication overhead. In addition, for privacy preservation, it introduces a random masking provided by the TPA into the process of generating proof to blind the data information. It further exploits the aggregate BLS signature technique from bilinear maps to perform multiple auditing tasks simultaneously, of which the principle is to aggregate all the signatures by different users on various data blocks into a single short one and verify it for only one time to reduce the communication cost in the verification process. Thus, it may be a new trend to design a more effective scheme, including different audit strategies for various types of cloud data.

2.4 KSF-OABE: Outsourced Attribute-Based Encryption with Keyword Search Function for Cloud Storage

Attribute-based encryption technology[4] has been used to design fine-grained access control system, which provides one good method to solve the security issues in cloud setting. Outsourced ABE with fine-grained access control system can largely reduce the computation cost for users who want to access encrypted data stored in cloud by outsourcing the heavy computation to cloud service provider. As the amount of encrypted files stored in cloud is becoming very huge, which will hinder efficient query processing. To deal with above problem, a new cryptographic primitive called attribute-based encryption scheme with outsourcing key-issuing and outsourcing decryption, which can implement keyword search function (KSFOABE). The time-consuming pairing operation can be outsourced to the cloud service provider, while the slight operations can be done by users. Thus, the computation cost at both users and trusted authority sides is minimized. The proposed scheme supports the function of keywords search which can greatly improve communication efficiency and further protect the security and privacy of users.

2.5 Minimum-Cost Cloud Storage Service Across Multiple Cloud Providers

Many cloud service providers provide data storage services with data centres distributed worldwide. These data centers[5] provide different get/put latencies and unit prices for resource utilization and reservation. Then we propose three enhancement methods to reduce the payment cost and service latency: 1) coefficient-based data reallocation; 2) multicast-based data transferring and 3) request redirection-based congestion control. According to the operations of a customer's clients, the customer data center generates read/write requests to a storage datacenter storing the requested data. For a customer, *DAR* aims to find a schedule that allocates each data item to a number of selected datacenters, allocates request serving ratios to these datacenters and determines reservation in order to guarantee the SLO and minimize the payment cost of the customer This work aims to minimize the payment cost of customers while guarantee their SLOs by using the worldwide distributed data centers belonging to different CSPs with different resource unit prices. In the first model this cost minimization problem using integer programming. Due to its NP-hardness, so they introduced the *DAR* system as a heuristic solution to this problem, which includes a dominant-cost based data allocation algorithm among storage data centers and an optimal resource reservation algorithm to reduce the cost of each storage data center.

2.6 An Economical and SLO-Guaranteed Cloud Storage Service Across Multiple Cloud Service Providers

A multi-cloud Economical and SLO-guaranteed Storage Service[6], which determines data allocation and resource reservation schedules with payment cost minimization and SLO guarantee. ES3 incorporates a coordinated data allocation and resource reservation method, which allocates each data item to a datacenter and determines the resource reservation amount on datacenters by leveraging all the pricing policies, a genetic algorithm based data allocation adjustment method, which reduce data Get/Put rate variance in each datacenter to maximize the reservation benefit. The problem to find the optimal data allocation and resource reservation schedules for cost minimization and SLO guarantee using an integer programming is done by Payment Minimization Objective method. They propose a multi-cloud Economical and SLO-guaranteed cloud Storage Service for a cloud broker over multiple CSPs that provides SLO guarantee and cost minimization even under the Get rate variation. ES3 is more advantageous than previous methods in that it fully utilizes different pricing policies and considers request rate variance in minimizing the payment cost. ES3 has a data allocation and reservation method and a GA-based data allocation adjustment method to guarantee the SLO and minimize the payment cost.

2.7 ASSER: An Efficient, Reliable, and Cost-Effective Storage Scheme for Object-Based Cloud Storage Systems

An ASSembling chain of Erasure coding and Replication. ASSER[7] stores each object in two parts: a full copy and a certain amount of erasure-coded segments. We establish dedicated read/write protocols for ASSER leveraging the unique structural advantages. On the basis of elementary protocols, we implement sequential and PRAM consistency to make ASSER feasible for various services with different performance/consistency requirements. Evaluation results demonstrate that under the same fault tolerance and consistency level, ASSER outperforms N-way replication and pure erasure coding in I/O throughput under diverse system and workload configurations with superior performance stability. ASSER delivers stably efficient I/O performance at much lower storage cost than the other comparatives. MPL is an extended mechanism of prevalently-adopted Parity Logging technique. The benefits brought by MPL mechanism are two folded. First, parity logging facilitates efficient handling towards update requests. The segment-chain in ASSER takes charge of receiving and handling update requests, thus reducing the amount of disk space needed to be overwritten. Second, introducing multi version into traditional parity logging enables ASSER to naturally support multiple consistency levels. Each object can have more than one recoverable versions in ASSER, and whether a version is valid to return is determined by the consistency level that ASSER is configured. ASSER, a hybrid storage scheme that aims at providing balanced trade-off between I/O performance and space efficiency with low storage cost. They proposed a mechanism called multiversional parity logging to facilitate efficient read/write handling in ASSER. We evaluated the performance of ASSER and the robustness of its implementation. According to their experimental results, with only half of extra space overhead, ASSER outperformed CRAQ in write-heavier workload and stayed evenly matched in read-heavier workload. Finally, they verified the feasibility of ASSER in practical environment through real-world traces driven experiment.

2.8 Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage

It shows how to securely, efficiently, and flexibly share data with others in cloud storage. We describe new public-key cryptosystems[8] that produce constant-size cipher texts such that efficient delegation of decryption rights for any set of cipher texts are possible. The novelty is that one can aggregate any set of secret keys and make them as compact as a single key, but encompassing the power of all the keys being aggregated. This compact aggregate key can be conveniently sent to others or be stored in a smart card with very limited secure storage. We provide formal security analysis of our schemes in the standard model. We also describe other application of our schemes. In particular, our schemes give the first public-key patient-controlled encryption for flexible hierarchy, which was yet to be known. The design is based on the collusion-resistant broadcast encryption scheme proposed by Boneh et al. Although their scheme supports constant-size secret keys, every key only has the power for decrypting cipher texts associated to a particular index. We, thus, need to devise a new Extract algorithm and corresponding Decrypt algorithm.

2.9 Anonymous and Traceable Group Data Sharing in Cloud Computing

With cloud computing, how to achieve secure and efficient data sharing in cloud environments is an issue to be solved. In addition, how to achieve both anonymity and traceability is also a challenge in the cloud for data sharing. In this paper enabling data sharing and storage for the same group in the cloud with high security and efficiency in an anonymous manner[9]. By leveraging the key agreement and the group signature, a novel traceable group data sharing scheme is proposed to support anonymous multiple users in public clouds is been focused. On the one hand, group members can communicate anonymously with respect to the group signature, and the real identities of members can be traced if necessary. On the other hand, a common conference key is derived based on the key agreement to enable group members to share and store their data securely.

The architecture of our cloud computing scheme is considered by combining the system model contains three entities: cloud, group manager and group members. Cloud provides users with seemingly unlimited storage services. In addition to providing efficient and convenient storage services for users, the cloud can also provide data sharing services. The cloud will not deliberately delete or modify the uploaded data of users, but it will be curious to understand the contents of the stored data and the user's identity. The cloud is a semi trusted party in our scheme. By presenting a secure and fault-tolerant key agreement for group data sharing in a cloud storage scheme. Based on the SBIBD and group signature technique, the proposed approach can generate a common conference key efficiently, which can be used to protect the security of the outsourced data and support secure group data sharing in the cloud

2.10 Block Design-based Key Agreement for Group Data Sharing inCloud Computing

Data sharing in cloud computing enables multiple participants to freely share the group data, By taking advantage of the symmetric balanced incomplete block design [10], we present a novel block design-based key agreement protocol that supports multiple participants, which can flexibly extend the number of participants in a cloud environment according to the structure of the block design. Based on the proposed group data sharing model, we present general formulas for generating the common conference key K for multiple participants. To support a group data sharing scheme for multiple participants applying an SBIBD, we design an algorithm to construct the (v; k +1; 1)-design. Moreover, the constructed (v; k + 1; 1)-design requires some transformations to establish the group data sharing model such that v participants can perform the key agreement protocol. With the help of the conference key agreement protocol, the security and efficiency of group data sharing in cloud computing is been greatly improved .The block design-based key agreement protocol that supports group data sharing in cloud computing. Due to the definition and the mathematical descriptions of the structure of a (v; k + 1; 1)- design, multiple participants can be involved in the protocol and general formulas of the common conference key for participant are been derived

2.11 Data Security for Cloud Environment with Semi-Trusted Third Party

Data security for cloud environment with semi-trusted third party (DaSCE)[11], a data security system that provides key management, access control, and file assured deletion. The DaSCE utilizes Shamir's (k, n) threshold scheme to manage the keys, where k out of n shares are required to generate the key. We use multiple key managers, each hosting one share of key. Multiple key managers avoid single point of failure for the cryptographic keys. The DaSCE makes use of both symmetric and asymmetric keys. The confidentiality and integrity services for data are provided through symmetric keys that are secured by using asymmetric keys. Asymmetric key pairs are generated by third party KMs. We modeled and analyzed FADE. The analysis highlighted some issues in key management of FADE. DaSCE improved key management and authentication processes. The working of the DaSCE protocol was formally analyzed using HLPN, SMT-Lib, and Z3 solver. The performance of the DaSCE was evaluated based on the time consumption during file upload and download. The results revealed that the DaSCE protocol can be practically used for clouds for security of outsourced data. The fact that the DaSCE does not require any protocol and implementation level changes at the cloud makes it highly practical methodology for cloud.

2.12 A New Privacy-Aware Public Auditing Scheme for Cloud Data Sharing with Group Users

New privacy-aware public auditing mechanism for shared cloud data by constructing a homomorphic verifiable group signature[12]. Moreover, our scheme ensures that group users can trace data changes through designated binary tree; and can recover the latest correct data block when the current data block is damaged. To ensure the integrity of data stored in cloud servers, a number of mechanisms based on various techniques have been proposed. The system model contains four entities cloud, TPA, trusted private key generator (PKG), and group users. To achieve integrity checking of the shared data in the cloud, NPP is expected to the following design objectives Public auditing Authorized auditing Identity privacy Traceability Support data traceability and recoverability Support group dynamics The multi-level privacy pre-serving public auditing scheme for cloud data sharing with multiple managers. During the process of auditing, the TPA cannot obtain the identities of the signers, which ensures the identity privacy of the group users. Moreover, unlike the existing schemes, the proposed NPP requires at least group managers to work together to trace the identity of the misbehaving user. Therefore, it eliminates the abuse of single-authority power and ensures non-frame ability

2.13 A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud

It is a secure data sharing scheme for dynamic members. First, we propose a secure way for key distribution without any secure communication channels, and the users can securely obtain their private keys from group manager[13]. Second, our scheme can achieve fine-grained access control, any user in the group can use the source in the cloud and revoked users cannot access the cloud again after they are revoked. Third, we can protect the scheme from collusion attack, which means that revoked users cannot get the original data file even if they conspire with the untrusted cloud. The main design goals of the proposed scheme including key distribution, data confidentiality, access control and efficiency as follows: Key distribution, Access control, Data confidentiality, Efficiency. First of all, the group member chooses a unique data file identity IDdata and a random number k 2 Zq; then computes When a user i with identity IDi is revoked, the group manager performs the operations This operation is performed by the group member and the cloud, as illustrated in Fig. 6, the group member encrypts IDdata with his key Ai and sends ID group; IDi; EncAi (IDdata)as a request to the cloud. A secure anti-collusion data sharing scheme for dynamic groups in the cloud. In our scheme, the users can securely obtain their private keys from group manager Certificate Authorities and secure communication channels. Also, our scheme is able to support dynamic groups efficiently, when a new user joins in the group or a user is revoked from the group, the private keys of the other users do not need to be recomputed and updated.

2.14 Dropping Activation Outputs With Localized First-Layer Deep Network for Enhancing User Privacy and Data Security

Dropping Activation Outputs With Localized First-Layer [14]Deep learning methods are used in anomaly detection, prediction, and supporting decision making for many applications. The current architecture of deep networks suffers the privacy issue that, users need to give out their data to the model (typically hosted in a server or a cluster on Cloud) for training or prediction. In addition to this, there is also a security risk of leaking these data during the data transmission from user to the model (especially when it is through the Internet). Targeting at these issues, in this paper, we proposed a new architecture for deep network in which users do not reveal their original data to the model. In our method, feed forward propagation and data encryption are combined into one process: we migrate the first layer of deep network to users' local devices and apply the activation functions locally, and then use the "dropping activation output" method to make the output non-invertible. The resulting approach is able to make model prediction without accessing users' sensitive raw data. The experiment conducted in this paper showed that our approach achieves the desirable privacy protection requirement and demonstrated several advantages over the traditional approach with encryption/decryption. We firstly introduce our new architecture of localized first layer deep network, then we introduce the proposed encrypt methods for both invertible activation (like sigmoid) and noninvertible activation (like rectifier). The main content will focus on a method for encrypting data for invertible activation function. Some of the methods used are: We proposed a new architecture for deep network, with the localized first layer of the network. how this architecture can support better privacy protection in model prediction is. Invertible activation function through Dropping activation outputs during feed-forward propagation are proved to be able to encrypt the original input data and preserving privacy. The whole encryption process can be improved through combining feed-forward propagation and data encryption into one process, which means no need for a specialized data encryption process on the local device, nor data decryption process on the server. During the encryption process, both invertible and noninvertible activation functions have been discussed and mathematically proved possible to do encryption. In error-back propagation, splitting the neural network into local device and server can provide data privacy during training. In other words, the server is able to provide model learning service by using error-back propagation, without accessing the original input data from the local device.

2.15 User-Level Runtime Security Auditing for the Cloud

User-Level Runtime Security [15],Cloud computing is a best solution for enabling ubiquitous, convenient, and on-demand accesses to a shared pool of configurable computing resources. But the adoption of cloud is still far behind because of the lack of transparency and accountability, which has traditionally been ensured through security auditing techniques. Auditing in cloud poses many unique challenges in data collection and processing and in verification. To this end, existing runtime auditing techniques do not offer a practical response time to verify a wide-range of user-level security properties for a large cloud. In this paper, we propose a runtime security auditing framework for the cloud with special focus on the user-level including common access control and authentication mechanisms e.g., RBAC, ABAC, SSO, and we implement and evaluate the framework based on Open Stack, a widely deployed cloud management system. RBAC Model. We focus on verifying multi-domain role-based access control (RBAC), which is adopted in real world cloud platforms ABAC Model. ABAC is considered as a strong candidate to replace RBAC in Sandhu which identifies several limitations of RBAC and thus emphasizes the importance of ABAC specially for large infrastructures (e.g., cloud). In fact, major cloud platforms have started supporting ABAC. SSO Mechanism. SSO, which is a popular cloud authentication extension and supported by major cloud platforms only requires a single user action to permit a user to access all authorized computers and systems. In this work, we detail two SSO protocols: OpenID and SAML supported by Open Stack and many other cloud platforms.

we proposed a runtime security auditing framework for the cloud with special focus on the user-level including different access control and authentication mechanisms. Our experimental results showed that our incremental approach in runtime verification reduces the response time to a practical level. This response time is satisfactory when the management operations are manually done by the administrators. The current approach would be insufficient to provide the same response time in the case of batch execution for management operations, As future work, to address this use case, we consider maintaining a scheduler including an event queue with different threads for different tasks in order to verify properties concurrently and therefore reduce the response time in this case.

2.16 Privacy-Preserving Data Processing withFlexible Access Control

Privacy-Preserving Data Processing [16],Cloud overcomes the bottlenecks of resource-constrained user devices and greatly releases their storage and computing burdens. Due to the lack of full trust in cloud service providers, the cloud users generally prefer to outsource their sensitive data in an encrypted form, which, however, seriously complicates data processing, analysis, as well as access control. Homomorphic encryption (HE) as a single key system cannot flexibly control data sharing and access after encrypted data processing. With the cooperation of a data service provider (DSP) and a computation party (CP), our scheme, based on Paillier’s partial homomorphic encryption (PHE) In addition, our scheme, based on the homomorphism of attribute-based encryption (ABE), is also designed to support flexible access control over processing results of encrypted data.
Secure Data Processing Based on SMC:
Secure multi-party computation enables computations over multi-user outsourced data without revealing any input.
Secure Data Processing Based on Homomorphic Encryption FHE schemes are designed to realize arbitrary computations over encrypted data. Due to high computation overhead, some extended schemes were proposed to improve FHE efficiency. Secure Data Access Control Cloud storage enables cloud users to upload their data to the cloud for storage and further sharing. The semantic security of HRES has been proved in our previous work Hence, we skip its security proof and focus on the security analysis of our proposed schemes.Performance Evaluation In this section, we analyze the computational complexity and the communication overhead of our proposed seven computing operation schemes. Further, we implemented them and tested their performances through simulations. The efficient and secure scheme are used to achieve privacy-preserving data processing with ABE based flexible access control. It can support seven basic operations and achieve fine-grained access control without the need of fully trusted cloud servers. Security analysis, performance evaluation and performance comparison with existing work further demonstrated that our scheme is efficient and effective with regard to big data processing operations.

2.17 Efficient Proofs of Retrievability with Public Verifiability for Dynamic Cloud Storage

Efficient Proofs of Retrievability with PublicVerifiability [17],Cloud service providers offer various facilities to their clients They can outsource their bulk data to the cloud server. The cloud server maintains these data in lieu of monetary benefits a malicious cloud server might delete some of these data to save some space and offer this extra amount of storage to another client. Therefore, the client might not retrieve her file (or some portions of it) as often as needed.

Notation: We take λ to be the security parameter. An algorithm denoted by A(1λ) is a probabilistic polynomial-time algorithm when its running time is polynomial in λ and its output y is a random variable which depends on the internal coin tosses of A.

Erasure Code: An (˜ m, ˜n,d)Σ-erasure code [is an error-correcting code that comprises an encoding algorithm Enc: Σ˜ n → Σ˜ m (encodes a message consisting of ˜ n symbols into a longer codeword consisting of ˜ m symbols) and a decoding algorithm.

Merkle Hash Tree: A Merkle hash tree is a binary tree where each leaf-node stores a data item. The label of each leaf-node is the data item stored in the node itself .A collision-resistant hash function hCR is used to label the intermediate nodes of the tree. Each of the outputs of hCR on different inputs is a binary string of length O(λ).

Digital Signature: Scheme Diffie and Hellman introduce the public-key cryptography and the notion of digital signatures in their seminalpaper"NewDirections in Cryptography" Rivest et al. propose the first digital signature scheme based on the RSA assumption Boneh et al. introduce the first signature scheme where the signatures are short (e.g., such a signature of size 160 bits provides the security comparable to that of a 1024-bit RSA signature).

Discrete Logarithm Assumption: The discrete logarithm problem over a multiplicative group Gq = hgi of prime order q .Dynamic Proofs of Retrievability: A dynamic PoR scheme consists of the following protocols between two stateful parties: a client (data owner) and a server.

Homomorphic Hash Function: A homomorphic hash function h : Fm → Gq (for a finite field F and a multiplicative group Gq of prime order q) is defined as a collision-resistant hash function satisfying the following two properties: 1) for vectors u,v∈ Fm and scalars α,β ∈ F, it holds that h(αu + βv) = h(u)α ·h(v)β, and 2) it is computationally hard to find vectors u,v∈ Fm (u6= v) such that h(u) = h(v).

PoR scheme where the client can update her data file after the initial outsourcing of the file to the cloud server and retrieve all of her data at any point of time. Our scheme is publicly verifiable, that is, anyone having the knowledge of the public parameters of the scheme can perform an audit on the client's behalf, and it offers security guarantees of a dynamic PoR scheme.

2.18 Decentralized Server-aided Encryption for Secure Deduplication in Cloud Storage

Decentralized Server-aided Encryption [18],Cloud storage provides scalable and low cost resources featuring economies of scale based on multi-tenant architecture. As the amount of data outsourced grows explosively, data deduplication, a technique that eliminates data redundancy, becomes essential. However, deduplication leads to problems with data confidentiality Server-aided encryption schemes have been proposed to achieve the strongest confidentiality but with the cost of managing a key server (KS). The key idea of our proposed scheme is to construct an inter-KS deduplication algorithm, by which a cloud storage service provider can perform deduplication over cipher texts from different KSs within a tenant or across tenants.

Convergent Encryption: Many secure deduplication solutions try to achieve data confidentiality as the primary goal while addressing other issues such as ownership management authorization authenticity access control and reliability These solutions commonly utilize convergent encryption (CE) algorithms to enable deduplication over encrypted data.

Server-aided Encryption: In order to solve the problem of convergent encryption and resist brute-force attacks, the key generation method has to be strengthened so that the complexity of convergent key space is preserved. Recent studies try to achieve this goal by using an additional key server from which users obtain convergent keys independent of the message.

Bilinear pairings: Bilinear map. Let G and GT be two multiplicative cyclic groups of prime or derp. Let g be a generator of G. A bilinear map is an injective function e : G × G → GT

(GDH) group Let G be a multiplicative group of a sufficiently large prime order. We consider the following two problems and a GDH group in G.

- Computational Diffie-Hellman (CDH) problem
- Decisional Diffie-Hellman (DDH) problem.

GDH group

A blind signature is a form of digital signature in which a message is blinded before it is signed. The goal of blind signature protocols is to allow a user to obtain a signature from a signer and to verify that signature. Our proposed scheme was implemented so as to fully utilize the parallelism supported by modern CPU architecture. By conducting extensive experiments and micro benchmarks on the real servers, we analyzed the performance with respect to the computational efficiency for the file upload operation and the inter-KS deduplication algorithm, and demonstrated that the proposed scheme outperforms the previous schemes.

2.19 Providing User Security Guarantees in Public Infrastructure Clouds

Security Guarantees in Public Infrastructure Clouds [19], The infrastructure cloud (IaaS) service model offers improved resource flexibility and availability, where tenants – insulated from the minutiae of hardware maintenance – rent computing resources to deploy and operate complex systems. Large-scale services running on IaaS platforms demonstrate the viability of this model; nevertheless, many organizations operating on sensitive data avoid migrating operations to IaaS platforms due to security concerns. Trusted Cloud Compute Platform" (TCCP) to ensure VMs are running on a trusted hardware and software stack on a remote and initially untrusted host. To enable this, a trusted coordinator stores the list of attested hosts that run a "trusted virtual machine monitor" which can securely run the client's VM. Cooper and Martin described in in a secure platform architecture based on a secure root of trust for grid environments precursors of cloud computing. Trusted Computing is used as a method for dynamic trust establishment within the grid, allowing clients to verify that their data will be protected against malicious host attacks. We share the threat model with which is based on the Dolev-Yao adversarial model and further assumes that privileged access rights can used by a remote adversary ADV to leak confidential information Problem Statement The introduced ADV has far-reaching capabilities to compromise IaaS host integrity and confidentiality. We define a set of attacks available to ADV in the above threat model. These protocols are successively applied to deploy a cloud infrastructure providing additional user guarantees of cloud host integrity and storage security.

2.20 Achieving Efficient and Secure Data Acquisition for Cloud-Supported Internet of Things in Smart Grid

Secure Data Acquisition for Cloud-Supported Internet of Things [20], The IoT front-ends are responsible for data acquisition and status supervision, while the substantial amount of data is stored and managed in the cloud server. Achieving data security and system efficiency in the data acquisition and transmission process are of great significance and challenging, because the power grid-related data is sensitive and in huge amount. we present an efficient and secure data acquisition scheme based on cipher text policy attribute-based encryption. Let G0 and G1 be two multiplicative cyclic groups of prime order p and g be the generator of G0. The bilinear map e is, e : G0 ×G0 →G1, for all a,b∈ Zp. 1) Bilinearity: ∀u,v∈G1,e(ua,vb) =e(u,v)ab. 2) Nondegeneracy: e(g,g) =1. 3) Symmetric: e(ga,gb) =e(g,g)ab =e(gb,ga). Problem Let G be a multiplicative cyclic group of prime order p and g be its generator. Given a tuple <g,gx>, where g ∈R G and x∈ ZP are chosen as input uniformly at random, the discrete logarithm (DL) problem is to recover x.

Policy Attribute-Based Encryption Definition 2: Let P ={ P1,P2,...,Pn} be a set of participants, let U = 2{P1,P2,...,Pn} be the universal set. If ∃ AS ⊆U \{∅ }, then AS can be viewed as an access structure. If A ∈ AS,∀B ∈ U,A ⊆ B, and B ∈ AS, AS is considered as a monotonic AS. Then the sets in AS are defined as authorized sets, while the other sets are regarded as unauthorized sets. Secret sharing scheme is used for sharing a secret among a group of parties, each of whom only obtain a piece of the secret (namely a share of the secret). No single party can infer any information about the secret with its own share. The only way to reconstruct the secret is to combine a certain number of shares. We parallel the transmission and computation by partitioning the data and access tree into chunks, which reduces the response time of severs and the DR's waiting time. The ways to encrypt and decrypt will affect the security of the system A complete tree will be partitioned into several subtrees, and each DB will be encrypted with the CP-ABE. The difference is that each subtree only contains two levels: 1) one root node and 2) its child nodes. Our scheme is secure against the adversaries with polynomial time in the length of the access tree information. we propose a secure and efficient data acquisition scheme for Cloud-IoT in smart grid. In the proposed scheme, the large data is partitioned into several blocks, and the blocks are encrypted/decrypted and transmitted in sequence. In addition, we adopt the dual secret sharing scheme, which realizes the privacy-preserving.

3. CONCLUSION

Privacy and security are among the most important requirements in Big Data. Here we noticed the challenges in big data and also the issues that are faced for providing security due to its enormous size. We have seen the possible methods and solutions for implementing the security and privacy in the big data analytics. While these techniques provides a good starting point for securing the big data, further research is needed to turn them into practical solutions that can achieve privacy and security in the real world.

REFERENCES

[1]. Haifeng Lu, Chuan Heng Foh, Yong gang Wen, and Jianfei Cai, "Delay-Optimized File Retrieval under LT-Based Cloud Storage", IEEE transactions on cloud computing, vol. 5, no. 4, october-december 2017

[2]. Yong Cui , Zeqi Lai, Xin Wang, and Ningwei Dai," QuickSync: Improving Synchronization Efficiency for Mobile Cloud Storage Services", IEEE transactions on mobile computing, vol. 16, no. 12, december 2017

[3]. Hui Tian, Yuxiang Chen, Chin-Chen Chang,Hong Jiang, Yongfeng Huang, Yonghong Chen, and Jin Liu," Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage", IEEE transactions on services computing, vol. 10, no. 5, september/october 2017

[4]. Jiguo Li, Xiaonan Lin, Yichen Zhang, and Jinguang Han," KSF-OABE: Outsourced Attribute-Based Encryption with Keyword Search Function for Cloud Storage", IEEE transactions on services computing, vol. 10, no. 5, september/october 2017

[5]. Guoxin Liu and Haiying Shen, "Minimum-Cost Cloud Storage Service Across Multiple Cloud Providers", IEEE/ACM transactions on networking, vol. 25, no. 4, august 2017

[6]. Guoxin Liu, Haiying Shen, and Haoyu Wang, " An Economical and SLO-Guaranteed Cloud Storage Service Across Multiple Cloud Service Providers", IEEE transactions on parallel and distributed systems, vol. 28, no. 9, september 2017

[7]. Jianwei Yin, Yan Tang, Shuiguang Deng, Ying Li, Wei Lo, Kexiong Dong, Albert Y. Zomaya, and Calton Pu," ASSER: An Efficient, Reliable, and Cost-Effective Storage Scheme for Object-Based Cloud Storage Systems", IEEE transactions on computers, vol. 66, no. 8, august 2017

[8]. Cheng-Kang Chu, Sherman S.M. Chow, Wen-Guey Tzeng, Jianying Zhou, and Robert H. Deng "Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage" IEEE transactions on parallel and distributed systems, vol. 25, no. 2, february 2014

[9]. Jian Shen, Tianqi Zhou, Xiaofeng Chen, Jin Li, Willy Susilo"Anonymous and Traceable Group Data Sharing in Cloud Computing", IEEE Transactions on Information Forensics and Security

[10]. Jian Shen, Tianqi Zhou, Debiao He, Yuexin Zhang, Xingming Sun and Yang Xiang, Block Design-based Key Agreement for Group Data Sharing in Cloud Computing" IEEE Transactions on Dependable and Secure Computing .

[11]. Mazhar Ali, , Saif U. R. Malik, and Samee U. Khan, " Data Security for Cloud Environment with Semi-Trusted Third Party"IEEE transactions on cloud computing, vol. 5, no. 4, october-december 2017

[12]. Anmin Fu, Shui Yu, Yuqing Zhang, Huaqun Wang, and Chanying Huang," NPP: A New Privacy-Aware Public Auditing Scheme for Cloud Data Sharing with Group Users" , IEEE Transactions on Big Data.

[13]. Zhongma Zhu and Rui Jiang," A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud", IEEE transactions on parallel and distributed systems, vol. 27, no. 1, january 2016

[14]. Hao Dong, Chao Wu , Zhen Wei, and Yike Guo," Dropping Activation Outputs With Localized First-Layer

Deep Network for EnhancingUser Privacy and Data Security"IEEE transactions on information forensics and security, vol. 13, no. 3, march 2018

[15]. Suryadipta Majumdar, Taous Madi, Yushun Wang,Yosr Jarraya, Makan Pourzandi, Lingyu Wang, and Mourad Debbabi, " User-Level Runtime Security Auditing for the Cloud"

[16]. Wenxiu Ding, Zheng Yan and Robert H. Deng, Fellow," Privacy-Preserving Data Processing withFlexible Access Control"information, IEEE Transactions on Dependable and Secure Computing,

[17].Binanda Sengupta, Student Member, IEEE and Sushmita Ruj," Efficient Proofs of Retrievability with PublicVerifiability for Dynamic Cloud Storage", IEEE Transactions on Cloud Computing.

[18]. Youngjoo Shin, Dongyoung Koo, Joobeom Yun and Junbeom Hur Member," Decentralized Server-aided Encryption for Secure Deduplication in Cloud Storage",IEEE Transactions on Services Computing.

[19]. Nicolae Paladi, Christian Gehrmann, and Antonis Michalas," Providing User Security Guarantees in Public Infrastructure Clouds" IEEE transactions on cloud computing, vol. 5, no. 3, july-september 2017

[20]. Zhitao Guan, Jing Li, Longfei Wu, Yue Zhang, Jun Wu, and Xiaojiang Du," Achieving Efficient and Secure Data Acquisition forCloud-Supported Internet of Things in Smart Grid" IEEE internet of things journal, vol. 4, no. 6, december 2017