Indian Journal of Science and Technology

Privacy Preserving Approaches in Cloud: a Survey

T. Jothi Neela[1]* and N. Saravanan[2]

[1]PG Student, School of Computing, SASTRA University, Thanjavur, 613401, Tamilnadu, India; jothineela@gmail.com
[2]Assistant Prof., School of Computing, SASTRA University, Thanjavur, 613401, Tamilnadu, India; saranindia@gmail.com

Abstract

Cloud Computing is continuously evolving and showing consistent growth in the field of computing. But, the security issues and threats associated with it still stay as a cumbersome. The focal point of this paper is Privacy preserving in cloud computing. This paper analyses and discusses various methods like adopting cryptographic methods, writing access rights and policies, anonymising data, segregating or fragmenting and then reconstructing the data, etc. All these approaches would preserve the privacy of user and data and while performing public auditing on the cloud data. The surveyed approaches are showcased and compared with one another, stating their pros and cons. Finally, the results are centralized and the issues to be concentrated in the future are drawn out. Enabling complete user control over his data, anonymising or encrypting the sensitive data before outsourcing, notifying the data owner upon data access, altogether with the tied up security mechanisms would nullify the privacy issue. This would serve as a helping note in the progress of strengthening the privacy preserving approaches in Cloud Computing.

Keywords: Cloud Computing, Privacy Preserving, Access Control, Public Auditing.

1. Introduction

The construction of cloud and storing data in it has tremendous benefits. It facilitates the authenticated and authorized cloud users to access enormous resources that are outsourced and shared in the cloud. Whenever required, the user can request and gain the access (only, if the users' credentials are validated [13]) in an easy way and at low cost, irrespective of the user location. Also, cloud computing takes away the expenses spent on installing all hardware and software, by allowing users to rent the resources based on their needs. Despite all these benefits, cloud computing still faces many challenges which forbid the successful implementation of the cloud. These include both the traditional as well as cloud security challenges [12]. Specific to cloud computing, the issues are many, of which some are: identity management of cloud users, multi-tenancy support, securing the security of applications, preserving privacy of the users, attaining control over the life cycle of outsourced data, etc. Among which, the issues related to privacy preserving are alone looked at in this survey.

Preserving the privacy of user, his identity and data in the cloud is very mandatory. With the rise in growth of cloud computing, the concerns about privacy preserving are also getting increased [11, 14]. But reaching the peak in providing and assuring privacy-preserved data access in cloud is yet in progress and still needs much attention to attain the goal. The tied in issues of privacy which acts as the barrier are listed in Table 1. Addressing all these issues and designing a system which could not be compromised by the intruders or attackers would mark the success of Cloud Computing.

2. Privacy Preserving Methods

Several methods have been put forward to tackle this issue of privacy preserving. This work studies some of those approaches and provides a brief overview. It is important,

* *Corresponding author:*
T. Jothi Neela (jothineela@gmail.com)

Table 1. Issues in attaining privacy in cloud computing

Issue	Description
Insufficient user control	Owner of the data lacks control over their data in the cloud, especially when their data are accessed or processed in the cloud.
Information disclosure	Disclosure of sensitive data while data moves across the cloud. Sensitive information may be user's identity, usage data etc.
Unauthorized secondary storage	Possibility of accessing and retrieving the sensitive information and backing up.
Uncontrolled data proliferation	Data flow in the cloud is unpredictable and uncontrollable.
Dynamic Provision	Legally responsible entity in the cloud to assure privacy remains unclear, due to the dynamic nature of the cloud.

that the privacy has to be preserved anytime and anywhere. So, the work takes us in both tracks: preserving the privacy of the data as well as preserving the privacy while we prefer some third party auditing to assure the data correctness. Table 2 and 3 gives the brief overview of the work and the detailed discussion is as follows:

Table 2. Comparison of privacy preserving methods in cloud computing

Approach	Description	Usage of crytographic techniques	Future enhancements
Anonymity-based method.	Anonymises the sensitive data before storing in cloud.	NO	Automating the anonymisation.
Architecture for privacy-preserving database storage.	Prevents both internal and external attacks.	YES	Effective generation of right expressions.
Privacy-preserved Access Control.	Determines access rights for users and achieves access control.	YES	Reducing the cost of encryption/ re-encryption techniques.
Privacy-preserving Authorisation System.	Puts forth a policy based authorisation infrastructure.	NO	Virtualising the infrastructure services.
Privacy Preserving Data Outsourcing.	Guarantees privacy by means of data fragmentation.	NO	Construction of hypergraph rather than two-dimensional graph.
PccP approach.	Preserves both user identification and information.	NO	Enhancing the efficiency of user id generation.
Dynamic reconstruction of metadata.	Designs schemas for the database. Performs segregation and reconstruction of metadata.	YES	Deploying the work in private cloud.

Table 3. Illustration of privacy preserving public auditing schemes in cloud computing

Approach	Description	Advantage
Public Auditing for Data Storage Security[8]	Third party auditing.	Guarantees the correctness of data in cloud server. Secured batch auditing.
Public Auditing for Secure Cloud Storage[9]	Enhanced and secured third-party auditing.	Public auditing with zero-knowledge leakage.
Oruta[10]	TPA assuring data correctness.	Identity is preserved which is not achieved in[8] and [9].

2.1 Anonymity-based Method

Jiang Wang et al. put forward an Anonymity-based method to achieve and preserve privacy in cloud computing [1]. The anonymity algorithm processes the data and anonymises all or some information before releasing it in the cloud milieu. When required, the cloud service provider makes use of the background knowledge it has and incorporates the details with the anonymous data to mine the needed knowledge. This approach differs from the traditional cryptography technology for preserving user's privacy as it gets rid of key management and thus it stands simple and flexible. While anonymising is easier, the attributes that has to be made anonymous varies and it depends on the cloud service provider. This approach will be suitable only for limited number of services. Thus, the method has to be bettered by automating the anonymisation.

2.2 A Privacy-Preserving Architecture

Architecture for database storage [2] in cloud is proposed in this paper, which preserves the privacy of users' data. This approach prevents the risk of both external and internal attacks to the outsourced data. The main architectural elements are the user interface, user engine, rule engine and the cloud database. Through user interface, the request for accessing database is obtained, which is sent as an XML/RPC request to the user engine, rule engine and finally to the cloud database. By means of encrypting and assigning secured identities for each request and response at each stage, together with the maintenance of machine readable usage /access rights, privacy is preserved. While it is easier to carry out the encryption schemes, there exists a difficulty in providing machine readable access rights. This problem of effective right expressions generation is the future work that has to be carried out.

2.3 Privacy-Preserved Access Control

Miao Zhou et al. [3] considered the privacy of users in the cloud environment and proposed a flexible method of access control. Each cloud user is linked with certain attributes, which determines their access rights. The paper propounded a two-tier encryption model in which the base phase and surface phase builds up the two tiers of the model respectively. At the first phase, the data owner performs local attribute-based encryption on the data that has to be outsourced. The surface phase on the other hand is performed by the cloud servers, after the initialization done by the cloud data owner. This phase implements the Server re-encryption mechanism (SRM). The SRM dynamically re-encrypts the encrypted data in the cloud, when the owner of that data requests. The request for SRM arises either when a new user has to be created or an existing user has to be repealed. Though the re-encryption takes place in cloud server, the privacy of users data is not compromised as the access policies remains hidden to the cloud servers. Thus, in this paper privacy of data is preserved by providing full access control to the owner of the data and by disallowing the cloud provider to gain knowledge about the data.

2.4 A Privacy Preserving Authorisation System

David W. Chadwick et al. explained a policy based authorization infrastructure for the cloud with the intention of preserving privacy of user's data [4]. Users can define their access policies and cling it to their data. This assures the controlled access of data in the cloud. Policy Decision Points (PDPs) and Policy Enforcement Points (PEPs) are used for making authorization decisions and enforcing these decisions respectively. Master PDP is launched which figures out and solves the conflicts among various decisions of different PDPs. Obligation service is provided as a part of the authorisation infrastructure, by means of which the data owner is intimated about the authorized or unauthorized access of their data. This authorisation infrastructure trusts the cloud providers and considers only the threats that come from outsiders. As the cloud provider is trusted, encryption of outsourced data is not done. The enhancement of this approach could be done by focusing on security threats from cloud providers and also by partitioning the infrastructure into separate services, each running in a distinct virtual machine. This would step-up the performance of the system.

2.5 A Privacy Preserving Data Outsourcing

In [5], one another method is constructed for preserving the confidentiality of users' data. The privacy constraints are illustrated by means of a graph. The nodes and links represent the attributes and confidentiality between the corresponding nodes respectively. Sensitive attributes are the subset of the entire group of attributes. These attributes should not be leaked out to the external party. A relation is drawn over such attributes, which is then vertically fragmented. While the owner has one fragment, the other fragment is placed at the external server. By making use

of a common id, the relation can then be reconstructed. A graph coloring algorithm is used to perform fragmentation and placing the fragments at the appropriate location, as well. While fragmenting, it is necessary to check that the workload is kept minimized at the source and also the confidentiality constraints not been breached by the server fragment. The fragmentation is carried out based on certain metrics like Min-Attr, Min-Query and Min-Cond. These metrics combined with the appropriate fragmentation guarantees that the outsourced data will always be protected from third party attacks, thus ensuring keeping up of the privacy. Thus this work adopted only fragmentation to attain privacy effectively and efficiently, keeping the cryptographic techniques aside. The effectiveness can still be improved, by constructing a hyper graph rather than a two-dimensional graph.

2.6 PccP Model for Cloud

Preserving cloud computing Privacy (PccP) model is explained in [6], which is one another approach to attain privacy. The Consumer Layer forms the basement of the model, where the cloud users have to submit their request for accessing cloud services. The second layer is the Network Interface or Address Mapping layer. The function of this layer is to modify the original IP address related to the access request. Thus, it assures the privacy of users' IP address. Next layer is the Privacy Preserved Layer, which is the topmost layer of the model. This layer has an associated Unique User Cloud Identity Generator. Hence, this layer preserves the privacy of users' sensitive information by implementing the Privacy check mechanism. This mechanism enables the user to specify the access control and the amount of data transparency in the cloud. If a particular Personal Data Attribute (PDA) of a user has to be specified with the transparency level, then a Boolean function of the attribute is to be carried out, which is named as Transparency Purpose in Cloud (TPC). Thus, PccP forecloses both the access of user identification and data content.

2.7 Dynamic Metadata Reconstruction

Adeela Waqar et al. [7] focused on the possibility of metadata exploitation in the cloud. By gaining knowledge of the metadata, the attacker could compromise users' privacy. As a solution, a framework is proposed to preserve the data privacy. First, the metadata that has to be put in cloud's database are segregated. The segregated attributes are then grouped as exclusively private, partially private and non-private depending on the sensitivity of data. Following this data classification, the next phase called table splitting comes up, where the database tables are divided both horizontally and vertically. The splitting of the database table ensures the database normalization. Next is to perform metadata reconstruction as and when required by the cloud. This phase is called ephemeral referential consonance. This phase guarantees that data is not leaked from the cloud database both before and after splitting. These steps are illustrated by considering the possible attacks on metadata kept in Eucalyptus database files and ensuring the prevention of attacks by the proposed framework. Thus, the method proves to be efficient.

2.8 Public Auditing for Secured Data Storage

C. Wang et al. proposed a privacy-preserving method to carry out public auditing on the cloud information [8]. In case of cloud computing, it is not sufficient to adopt the traditional cryptographic measures to achieve security. The reason is due to data outsourcing and the ubiquitous nature of the data. So, in this paper they opt the concept of Third Party Auditing (TPA). Homomorphic authenticator and random masking ensures that TPA could not gain any knowledge during the process of auditing. Thus, TPA is trusted and capable of accessing the cloud storage to perform auditing. The audit report brings out the risks, if any is present in the data. The public auditing system is built using four algorithms and two phases. KeyGen and SigGen algorithms make up the first phase called Setup, in which initialization of secret parameters and generation of verification metadata are done. Following this, the Audit phase carries out the auditing process and ascertains the correctness of data in the cloud server. This is done in this second phase using GenProof and VerifyProof algorithms. The approach guarantees the correctness of data in cloud server, preservability of privacy and security for batch auditing (simultaneous auditing for multi-user setup).

C. Wang et al. in [9] enhanced their previous proposal by improving the security strength of data storage. A new protocol for privacy-preserving public auditing is designed for this purpose. Public auditing with zero-knowledge leakage is also achieved. Batch auditing is also enhanced with the improvement in main auditing scheme. As an extension to the previous work, the authors put forward the support for data dynamics and generalization of the auditing scheme in

this paper. An experiment is conducted on an instance of Amazon EC2 and the better performance of the proposed design is proved.

2.9 Oruta

Boyang Wang et al. [10] analyzed the work of Wang et al. and propounded another public auditing mechanism. Oruta provides data privacy, identity privacy. Also it ensures correctness and unforgeability while carrying out the public auditing. In [8] and [9], the identity privacy is not achieved. The approach considers three main entities: the cloud server, TPA and the users. Users are statically grouped as the original user (owner of the outsourced data) and group users. The original user can control their data and its flow in the cloud. All users request and depend on the TPA to carry out auditing for verifying the rightness of data. Homomorphic Authenticable Ring Structures (HARS) scheme, comprising three algorithms: KeyGen, RingSign and RingVerify are constructed here for achieving the privacy-preserving auditing. Still, the approach can be empowered by focusing on an efficient auditing approach to ascertain the integrity of shared data in dynamically grouped users' environment.

3. Conclusion

Cloud Computing is gaining popularity and advancement day-by-day. But still the security threat hinders the success of Cloud Computing. In this paper, some of the privacy threats are addressed and the techniques to overcome them are surveyed. While some approaches utilized traditional cryptographic methods to achieve privacy, some other approaches kept them away and focused on alternate methodologies in achieving privacy. Also, approaches to preserve privacy at the time of public auditing are also discussed. Thus, to conclude it is necessary that every cloud user must be guaranteed that his data is stored, processed, accessed and audited in a secured manner at any time. The user must be given complete access control over the published data. Also, powerful security mechanisms must always supplement every cloud application. Attaining all these would end up in achieving the long dreamt vision of secured Cloud Computing in the nearest future.

4. References

1. Wang J, Zhao Y et al. (2009). Providing Privacy Preserving in cloud computing, International Conference on Test and Measurement, vol 2, 213–216.
2. Greveler U, Justus b et al. (2011). A Privacy Preserving System for Cloud Computing, 11th IEEE International Conference on Computer and Information Technology, 648–653.
3. Zhou M, Mu Y et al. (2011). Privacy-Preserved Access Control for Cloud Computing, International Joint Conference of IEEE TrustCom-11/IEEE ICESS-11/FCST-11, 83–90.
4. Chadwick D W, and Fatema K (2012). A privacy preserving authorisation system for the cloud, Journal of Computer and System Sciences, vol 78(5), 1359–1373.
5. Sayi T J V R K M K, Krishna R K N S et al. (2012). Data Outsourcing in Cloud Environments: A Privacy Preserving Approach, 9th International Conference on Information Technology- New Generations, 361–366.
6. Rahaman S M, and Farhatullah M (2012). PccP: A Model for Preserving Cloud Computing Privacy, International Conference on Data Science & Engineering (ICDSE), 166–170.
7. Waqar A, Raza A et al. (2013). A framework for preservation of cloud users' data privacy using dynamic reconstruction of metadata, Journal of Network and Computer Applications, vol 36(1), 235–248.
8. Wang C, Wang Q et al. (2010). Privacy-Preserving Public Auditing for Storage Security in Cloud Computing, Proceedings IEEE INFOCOM'10.
9. Wang C, Chow S S M et al. (2013). Privacy-Preserving Public Auditing for Secure Cloud Storage, IEEE Transactions on Computers, vol 62(2), 362–375.
10. Wang B, Li B et al. (2012). Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud, IEEE Fifth International Conference on Cloud Computing, 295–302.
11. Gellman R (2009). WPF REPORT: Privacy in the clouds: Risks to privacy and confidentiality from cloud computing.
12. Rong C, Nguyen S T et al. (2013). Beyond lightning: A survey on security challenges in cloud computing , Computers & Electrical Engineering, vol 39(1), 47–54.
13. Takabi H (2010). Security and Privacy Challenges in Cloud Computing Environments, IEEE Security & Privacy, vol 8(6), 24–31.
14. Xiao Z, and Xiao Y. Security and Privacy in Cloud Computing, IEEE Communications Surveys & Tutorials, vol PP(99), 1–17.