# Privacy Preservation in Cloud Computing: An Experimental Analysis

**Smita Sharma, Sanjay Tyagi**

***Abstract*: Cloud Computing has become popular among organizations for managing data because of its low cost, robustness, scalability and high availability. Privacy issues are major concern while outsourcing the data on cloud. In this paper, the authors have provided a review of the existing techniques for preserving privacy in cloud computing environment.**

***Index Terms*: Cloud Computing, Privacy Preservation, Security.**

## I. INTRODUCTION

Cloud computing can be defined as the use of remote servers on the internet to store, process and manage data rather than on a local server or personal computer. Using cloud computing, the organizations or clients do not buy servers, but lease those on hourly basis from cloud service providers. The cloud service providers manage the servers and the clients need not to worry about the underlying infrastructure.

Cloud Computing has been defined by the National Institute of Standards and Technology (NIST) as [1]: "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

The model of cloud computing consists of deployment models (public, private, community and hybrid cloud) and service models (Infrastructure as a Service, Platform as a Service and Software as a Service).

The major models for deploying and accessing cloud computing environment are [2]:

- **Public Cloud:** The service providers establish the infrastructure that is available for general public on the basis of subscription. In this model, the data and applications of users are placed on datacenters on service providers' premises.
- **Private Cloud:** This cloud model is used by the organizations or institutions that want more control over the cloud consumers, infrastructure and computation resources those provided by public cloud. A private cloud can be deployed either at organization's premises or outside.
- **Community Cloud:** More than one organization that have common policies and security considerations can use the same cloud i.e. community cloud.
- **Hybrid Cloud:** This cloud model is a mix of two or more clouds, e.g. Combination of private and public cloudcan be used for accessing more services securely.

Cloud computing provides various services that can be treated as its Service Models. Basic cloud service models are [2]-[4]:

- **Infrastructure as a Service**
  It refers to providing on-demand services of computing resources in the form of software, servers and networking equipments. By using IaaS, purchasing and managing basic software and hardware components can be avoided.
- **Platform as a Service**
  Under this service model, the cloud service provider provides the computing platform as an on-demand service for developing, deploying and executing applications. The service provider is also responsible for managing fault tolerance and providing scalability whereas the role of the users is limited to the logic of application.
- **Software as a Service**
  It is the service model that provides applications and resources to execute them as and when requested by the consumers. The consumers need not to manage the infrastructure and applications. It reduces the overall cost of developing, operating and maintaining hardware as well as software.

Cloud Computing has various advantages like services can be used without human intervention whenever required, location independent resources are pooled over the network that can serve multiple users at a time and can be scaled as per the need and services cost as much as used by the consumer [5].

Despite of all the features, cloud models also have some security issues [6] that prevent the customers from outsourcing the data to cloud. A report generated after discussion with some security professionals by the Cloud Security

Revised Manuscript Received on May 25, 2019.

**Smita Sharma**, Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, India.

**Sanjay Tyagi**, Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, India.

Alliance stated that most of the participants considered storing data on cloud as risky [7]. The users' data is processed on the machines that is owned or operated by service providers and the users have no control on how the data is maintained. Also, cloud service providers have the opportunity to exploit and analyze users' personal data [8]. The cloud users may want to know about what Cloud Service Providers will do with their personal or confidential data. It increases the risk to privacy as the cloud consumers cannot afford to reveal their confidential data to any unauthorized user [2].

Cloud computing is a dynamic environment where data may move around within or across organizations, so it is required to protect the data and information [9].

There can be various issues in achieving privacy as [10]:

- Lack of user control
- Information leakage while moving across clouds
- Risk of unauthorized backup of sensitive data
- Uncontrolled data flow
- Dynamic nature of cloud

The role of cloud service providers is to ensure the privacy of consumers' data and the goal is to increase the trust in cloud services. The cloud consumer must be assured that the data can be accessed and audited securely any time. Also, the user can control the access of cloud data [10].

## II. LITERATURE REVIEW

Different authors have explored security and privacy issues in cloud computing environment and techniques to preserve the privacy of cloud consumers. A brief review of some of the privacy preservation approaches has been presented in the present paper.

Wang *et al.* [11] described the importance of protecting individuals' privacy in the cloud. The authors concluded that privacy is an important issue for cloud computing, as per the legal rules as well as to maintain user trust. The authors also discussed some of the privacy preserving techniques and proposed a new anonymity algorithm for the cloud computing services. The data was processed and sent to service providers. After that, it was analyzed by integrating the background knowledge retrieved from web or records to get the knowledge. In this algorithm, the service provider can directly use the data without any key. The authors concluded that the algorithm was more flexible than cryptography technology and safe to protect privacy.

An authenticated access control scheme was proposed by Ruj *et al.* [12] for securing cloud data and to preserve the privacy of users. Under this scheme, the authenticity of the user was verified before the data is stored on cloud and access control was also provided so that only authorized users could decrypt the data stored on cloud. Replay attacks were prohibited and the cloud user can create, modify & read data stored on the cloud. But the scheme had a limitation that the access policy of data is known to the cloud.

An algorithm to protect the data was proposed by Sayi *et al.* [13] that was applied at client side. The algorithm's results were then stored on the cloud so that unauthorized user cannot reconstruct and retrieve data from cloud. It preserved the privacy of data owner. The authors had prolonged the fragmentation approach for outsourcing the data while maintaining confidentiality and preserving privacy. A two-graph-coloring algorithm was proposed that determines which of the data can be outsourced on the cloud and which can be kept on local server. The solution meant to minimize the workload of data owner and the amount of data stored on local server such that data confidentiality and privacy can also be preserved. The approach used only fragmentation to maintain privacy effectively without using any cryptography technique.

Waqar *et al.* [14] concentrated on the chances of misuse of metadata in the cloud. If the attacker can understand the metadata, that could be unsafe for users' privacy. So the authors proposed a framework that redesigns the database schema and dynamically reconstruct the metadata to preserve the privacy. The cryptography and relational operations were used to modify the database schema and reconstruction to original schema was ensured. The authors showed that the proposed framework was suitable for private cloud.

Before migrating to cloud computing, it should guarantee data confidentiality, privacy preservation and resiliency. To ensure privacy protection, Jung *et al.* proposed AnonyControl [15] and AnonyControl-F [16] schemes. These schemes controlled the privilege to access the content stored on cloud. A private key was assigned to data owner by attribute authorities to allow data access. The authors showed that the proposed schemes were secure as well as efficient by conducting detailed analysis of their performance. AnonyControl-F was enhancement to AnonyControl, in which no information was disclosed to attribute authorities whereas AnonyControl was semi-anonymous that allowed little disclosure of information.

A theoretical analysis was presented by Pooja and Nagarathna [3], describing different types of security issues and threats that can affect the users' privacy. The authors concluded that privacy and security of data need to be considered to motivate more and more consumers for migrating to the cloud. The authors also analyzed existing approaches to preserve the privacy of data users.

Afterwards, Yang *et al.* [17] proposed a technique by hybridizing statistical analysis and cryptography that provide flexible access to medical data and preserve privacy while sharing data in cloud environment. The technique partitioned the data when stored on cloud and merged when access was requested. The

*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

queries were processed while maintaining privacy. The authors concluded that a better balance of privacy preservation and information utilization can be achieved if different technologies for preserving privacy were mixed.

Tang *et al.* [18] presented a detailed overview of security threats to outsourcing the data to a cloud and security techniques. The authors discussed existing protection techniques to attain secure, reliable and privacy-preserved cloud services. The challenges to security and future research directions were also discussed by the authors.

A semantic data splitting approach was proposed by Sanchez and Batet [8] that protected the privacy of cloud data. The approach evaluated semantic of data to detect that part of data which can cause information disclosure and then automatically split the data in such a way that ensured privacy preservation. The privacy requirements were defined by the data owners by giving the semantics of data that had to be stored securely. As the partitioning and storage was determined by the semantics, so data can be managed efficiently. To process the data securely before migrating it to the cloud, a local application was deployed on the user side. The application stored the metadata that was used to process the queries on data and to reconstruct the **results.**

Table I. Comparison of Privacy Preservation Schemes

| **Paper/ Comparison Factor** | **Encryption** | **Key Management** | **Access Control** | **Scheme** | **Signature** | **Authenti-cation** | **Implemented** |
|---|---|---|---|---|---|---|---|
| **Wang *et al.* [11]** | No | No | No | Anonymity Algorithm | No | No | On site |
| **Ruj *et al.* [12]** | Attribute Based Encryption | Decentralized Key Distribution | Fine grained | Authenticated Access Control Scheme | Attribute Based Signature | Yes | On site & in Cloud |
| **Sayi *et al.* [13]** | No | No | No | Fragmentation (Two-graph Coloring Algorithm) | No | No | On site |
| **Waqar *et al.* [14]** | Rivest-Shamir -Adleman (RSA) | Yes | No | Cryptography & Relational Operations | No | Yes | At Private Cloud Data |
| **Jung *et al.* [15][16]** | Ciphertext Policy-Attribute Based Encryption (CP-ABE) | Yes | Fine grained | Anonymous Attribute Based Privilege Control Schemes | No | No | On site & in Cloud |
| **Yang *et al.* [17]** | Symmetric & Asymmetric Key Cryptography | No | No | Hybrid of Statistical Analysis & Encryption | No | No | On site & in Cloud |
| **Sanchez & Batet [8]** | No | No | Yes | Semantic Data Splitting | No | No | On site |
| **Renuga & Jagatheeshw-ari [19]** | Yes | Yes | No | Data Sanitization using Optimal GSA | No | No | On site |

Renuga and Jagatheeshwari [19] proposed a method that uses sanitization algorithm for preserving privacy while transforming data. The gravitational search algorithm (GSA) was used to select the optimal key and that key was then used in the sanitization of data. The authors compared the proposed method with some existing approaches and concluded that the proposed method provided more security and improved the performance of privacy preservation in cloud computing.

Secure disintegration protocol (SDP) was proposed by Rawal *et al.* [20]. The proposed protocol protected the on-site privacy as well as privacy in the cloud. The authors concluded that the method not only provide security of data in cloud but also increase the throughput and performance of cloud environment.

Comparison of various approaches of preserving privacy has been presented in Table I based on scheme, encryption technique, key management, access control, authentication and signatures.

## III. CONCLUSION

Cloud computing offers many benefits for companies, public institutions and individuals who want to store and process their data in the cloud. However, there are security concerns that prevent many users from migrating to the cloud. In this paper, authors discussed privacy issues and reviewed some existing techniques used for privacy preservation. It is concluded that secure and efficient cloud services should be provided to facilitate the users with maximum benefits of cloud computing environment. To build users' trust in cloud services, it is important to ensure the privacy preservation.

## REFERENCES

1. P. Mell and T. Grance, "The *NIST* definition of cloud computing," 2011.
2. R. Buyya, C. Vecchiola, and S. T. Selvi, *Mastering cloud computing: foundations and applications programming*. Newnes, 2013.
3. H. Pooja and N. Nagarathna, "Privacy preserving issues and their solutions in cloud computing: A survey," *IJCSIT*, vol. 6, no. 2, pp. 1588–1592, 2015.
4. P. K. Tiwari and B. Mishra, "Cloud computing security issues, challenges and solution," *International journal of emerging technology and advanced engineering*, vol. 2, no. 8, pp. 306–310, 2012.
5. A. Jula, E. Sundararajan, and Z. Othman, "Cloud computing service composition: A systematic literature review," *Expert systems with applications*, vol. 41, no. 8, pp. 3809–3824, 2014.
6. Z. Xiao and Y. Xiao, "Security and privacy in cloud computing," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 843–859, 2013.
7. "Cloud usage: Risks and opportunities report," 2014.
8. D. Sánchez and M. Batet, "Privacy- preserving data outsourcing in the cloud via semantic data splitting," *Computer Communications*, vol. 110, pp. 187–201, 2017.
9. S. Pearson, "Taking account of privacy when designing cloud computing services," in *ICSE Workshop on Software Engineering Challenges of Cloud Computing, CLOUD'09* IEEE, 2009, pp. 44–52.
10. T. J. Neela and N. Saravanan, "Privacy preserving approaches in cloud: a survey," *Indian Journal of Science and Technology*, vol. 6, no. 5, pp. 4531–4535, 2013.
11. J. Wang, Y. Zhao, S. Jiang, and J. Le, "Providing privacy preserving in cloud computing," in *3rd Conference on Human System Interactions (HIS)*. IEEE, 2010, pp. 472–475.
12. S. Ruj, M. Stojmenovic, and A. Nayak, "Privacy preserving access control with authentication for securing data in clouds," in *12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*. IEEE, 2012, pp. 556–563.
13. T. Sayi, R. S. Krishna, R. Mukkamala, and P. K. Baruah, "Data outsourcing in cloud environments: A privacy preserving approach," in *Ninth International Conference on Information Technology: New Generations (ITNG)*. IEEE, 2012, pp. 361–366.
14. A. Waqar, A. Raza, H. Abbas, and M. K. Khan, "A framework for preservation of cloud users' data privacy using dynamic reconstruction of metadata," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 235–248, 2013.
15. T. Jung, X.-Y. Li, Z. Wan, and M. Wan, "Privacy preserving cloud data access with multi-authorities," in *Proceedings IEEE INFOCOM*. IEEE, 2013, pp. 2625–2633.
16. T. Jung, X.-Y. Li, Z. Wan, and M. Wan, "Control cloud data access privilege and anonymity with fully anonymous attribute based encryption," *IEEE transactions on information forensics and security*, vol. 10, no. 1, pp. 190–199, 2015.
17. J.-J. Yang, J.-Q. Li, and Y. Niu, "A hybrid solution for privacy preserving medical data sharing in the cloud environment," *Future Generation Computer Systems*, vol. 43, pp. 74–86, 2015.
18. J. Tang, Y. Cui, Q. Li, K. Ren, J. Liu, and R. Buyya, "Ensuring security and privacy preservation for cloud data services," *ACM Computing Surveys (CSUR)*, vol. 49, no. 1,p 13, 2016.
19. S. Renuga and S. Jagatheeshwari, "Efficient privacy-preserving data sanitization over cloud using optimal GSA algorithm," *The Computer Journal*, vol. 61, no. 10, pp. 1577–1588, 2018.
20. B. S. Rawal, V. Vijayakumar, G. Manogaran, R. Varatharajan, and N. Chilamkurti, "Secure disintegration protocol for privacy preserving cloud storage," *Wireless Personal Communications*, pp. 1–17, 2018.

## AUTHORS PROFILE

**Smita Sharma** is pursuing Ph.D. in Computer Science & Applications from Department of Computer Science & Applications, Kurukshetra University, Kurukshetra. She completed her M.Tech in Computer Science & Engineering from Department of Computer Science & Applications, Kurukshetra University, Kurukshetra. Her research area is Cloud Computing.

**Sanjay Tyagi** is an Assistant Professor at Department of Computer Science & Applications, Kurukshetra University, Kurukshetra. He holds an experience of 27 years. He received his Ph.D. from Kurukshetra University. His area of research includes Software Testing, Cloud Computing, MANETs and Information System. He has published about 70 papers in National and International Journals.