International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Survey on Privacy Preserving and Data Security Techniques

Pramila Kharat, Amar Buchade

[1,2] Department of Computer Engineering, Pune Institute of Computer Technology, Savitribai Phule Pune University, Pune 411043, India

Abstract: ***In day to day life cloud computing is most promising and widely used technology. As it provides infrastructure, platform and software as a service, it is used in almost all area of computing. We use cloud for data storage, data processing and computing purpose. In return we pay to cloud for using its resources. Here user is not sure whether his data is in safe hands as well as his identity is reveled with his data. In this Survey paper we enlist all the research carried out related to data security and users privacy preserving techniques in detail. Finally we proposed system which provides data ownership privacy using token generation for each user, each session and data security using RSA.***

Keywords: Cloud Computing, security**,** Privacy, Data Ownership**.**

1. Introduction

New revolving technologies in hardware, middleware, virtual machine and distributed systems required to complete the exponentially growing information technology industry need. Such technology should complete the need of all variety of customers from individual and organizations. All these facilities can be provided by cloud computing as it provides platform, infrastructure and software as a service using above mentioned technologies. This is a pay-as-you go model. As this technology requires resources, networking, hardware and virtualization in large scale there is combinations of all technologies drawbacks also which lead many loop whole in the systems. Among all the issues related to cloud computing we focus on the issues of user's privacy and data securities [9].

Researchers point out cloud security and user privacy issues. Some threats are enlisted over here.

- Sensitive private information can be disclosed while exchanging the data with cloud services.
- Once the data is given to cloud service user do not have control over the data and its security.
- Data is vulnerable to different network attacks such as DOS and DDOS [10].

In a secure multiparty computation author defined privacy and data security as follows:

1.1 Privacy

Means no other party should learn about the information that is more than its prescribed output [11].

1.2 Data Security

Means to protect data from any vulnerable or unwanted actions as well as only authorized users should access the data.

As government agencies are the big customer of the cloud computing services privacy and data security is the main is the main concern of such organizations. In this paper we carried out survey of the research which provides different solutions to above mentioned problems finally we proposed our novel technique for preserving privacy and data security.

1.3 Anonymous Cloud Framework

[39] To provide data ownership privacy in this frame work author uses tor circuit to create anonymous cloud for computing the data. Using a separate manager node it secure users identity and then provide data to the cloud for processing and the session is identified using tokens as well as data is secured using public key encryption techniques.
Limitation: Systems efficiency is dependent on number of nodes selected for tor circuit generation.

1.4 Secure Multiparty Computation

[11] secure multiparty means to consider the distributed computing where many devices wish to communicate simultaneously. In this privacy preserving technique author purposed a secure multiparty protocol construction and proved that it achieves privacy.
Limitations: its computation deals with secure functionality but the first priority should be given to complete the functionality in a polynomial time.

1.5 Digital Identity management

[40] In this technique of cloud computing author described the need of digital identity management services in the infrastructure of the cloud for privacy preserving. K-Anonymity is provided for linking of explicitly identifying information to its content map the information to at least k entities. And also illustrated that K- Anonymity can provide data integrity. Minimum generalization captures the property of the release process not to distort the data more then to achieve k- anonymity.

Limitations: additional techniques are required for this model to develop for specification and investigation with respect to different possible classes of data recipients.

1.6 Privacy Preserving Public auditing

[41] In this technique of public auditing cloud storage is of critical importance so that user can resort to third party auditor(TPA) to check the integrity of outsourced. Using privacy preserving protocol author enable external auditor to audit user's outsourced data in cloud computing without learning content. The system is efficient for multiple user data auditing also. Following are solutions provided for data security and privacy preserving in cloud techniques [10]:

- An encryption scheme to ensure data security in a highly interfering environment which maintains security standards against all threats.
- The service provider should be given limited access to the data and manage to see what exactly the data is.
- Strict access controls to prevent unauthorized and illegal access to the servers.
- An individual authority to separate user information from the data that is given to cloud before computing and then again link this user information to the user identity for billing purpose can be established.
- A strong authentication protocol to restrict unauthorized user from data access.
- Data should be encrypted with strong encryption public key techniques for network security and to prevent from other data attacks.

To fulfill these criteria we proposed a new system framework for cloud computing.

2. Proposed System

Here we described relation between the user and cloud communication, authorization and service computation. Here we cloud not only use for storage but we can use this system for accessing cloud services such as medical analysis or file conversions. As the slave nodes do not know like which user data is processing correctly by the slave node, also privacy preserving is achieved.

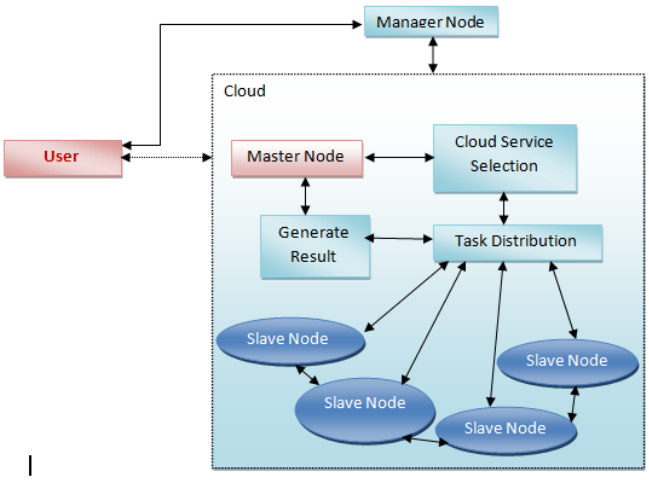

Figure 1: System Architecture

The proposed system is explained in four parts as follows:
Key Generation and key Distribution

- **Authentication Module**
- **Task Distribution**
- **Generate Result**

2.1 Key Generation and key Distribution

For security and authentication for each user we are generating the keys for each slave node and master node. All the public and private keys are generated using RSA algorithm. And keys are stored at manager.

2.2 Authentication module

Here user is authenticated by cloud server process by authentication process. And after the user can access the service provided by the cloud. Master node has the authority to send the data received from the user.

2.3 Task distribution

Task distribution is the most important module in proposed system. Receiving of data from master node and forward the data to slave nodes and load are done by this module. And finally result is generated from the slave nodes are combined then it is given to the user through manager node which also gives status of billing of cloud.

2.4 Generate Result

In this module the results generated from the slave nodes are combined and result is forwarded to the user through manager node which also takes care of billing of cloud.

For experiment we had taken diabetes detection data. This is a very sensitive data by analyzing it we can detect how many results are positive and how many are negative diabetes patients.

3. Related Work

Data security concerns are broadly identified as a significant obstruction to end users trust in cloud computing [7], [9], [10]. Associated provocations span at least three types of similar task: protected remote base attestation, protected data storage, and information-centric safety [17].

Trusted computing gives users high confirmation that they are interacting with a remote server containing of known, trusted hardware and software [18]. Safe storage regards the issue of protectively storing personal data in the cloud within computations that apply it (e.g., [19]). In opposite, information-centric approaches saturate data with self-safety attributes, such as by showing it in a form manageable to straight operation on cipher texts without decryption (e.g., [20]). Anonymous- Cloud's points of separating personal information from its hide information can be seen as an instance of the last of these approaches.

Common data anonymization is a vast research area spanning many years; however, the most broadly applied strategies for anonymization of data content are currently separate privacy

[21] and k-anonymity for privacy-preserving micro data lunched [22]. Such research important to our work by giving a mode for customers to anonymize personal data content before providing it to the cloud. We therefore assume that customer curious in privacy submit data that revels fewer secrets once it has been separated from hidden and semantic metadata, and that therefore benefits from our anonymization rules.

Previous task has also traversed separated document points from format and structure for more protective cloud storage and operation [23]. For example, HTML documents can be encoded in a format that decoupled their tree design from the textual points of items and properties. Since a group of personal data resides in the content, this permits differentiated operating of design-based queries in the cloud without disclosing the personal data.

To separated and hides origin metadata, Anonymous Cloud engaged onion routing based on Tor [13]. Tor has become the most successful public anonymity communication service in the Internet, with tens of millions of users worldwide [24]. In Tor, initiators choose a path through network and build a circuit in which each node or onion router in the path knows only its successor and predecessor, but no other nodes in the unit. Based on the selected route, the initiator first encrypts the data with one layer of encryption for each node in the route, from the last node to the first. This is likened to the levels of an onion, with each hop peeling one layer as the data is forwarded to its destination. The data can only be read in plaintext once it reaches the endpoint of the path and all layers have been peeled.

The Tor Cloud project [25] has applied a full-scale Tor system within a production-level cloud that operates on the Amazon EC2 cloud computing platform [2]. It facilitates a user-friendly way of utilize bridges to help users entry an uncensored Internet. Tor Cloud hides user pseudonyms (e.g., IP numbers) from untrusted third-party services, but does not adequate to anonymously access data from a third-party cloud [26], since clouds require a means of identifying users in sequence to control access to each user's personal information and bill them properly.

Our operation therefore enhanced cloud-based onion path with unknown credential system for identification [27]. Anonymous identification gives no knowledge proof of authentication, permitting information to be securely separated from origin for advance security. More precise anonymous credential systems (e.g., [28], [29], [30], [31], [32], [33]) helps additional security attributes, such as non-transferability, lazy revocation, and access hierarchies. These are not compulsory for our system, but could be alternated if such attributes are demanded for other reasons.

Our attack studies and examines do not consider the problem of end-to-end timing attacks Previous works have displays that these attacks are strongly affective against Tor and other onion path systems even when the attacker controls only a few ends [34], [16]. The Tarzan system secures against timing attacks through generation of artificial cover traffic that masks timing patterns in a sea of mimicry and noise [35]. Future work should consider the feasibility of supplementing Anonymous Cloud with similar protections.

Aside from applying security and rules that straight provides higher security, mechanisms that provide greater transparency for internal cloud operations— particularly distribution and management of security sensitive data—is critical for instilling greater confidence in end users [36], [37], [38]. Future work should therefore consider augmenting Anonymous Cloud with features that afford customers greater control over data distribution and scheduling details after Tor circuit construction, and without sacrificing anonymity.

3. Conclusion

In this survey we discussed various technique of privacy preserving and data security in cloud and thier limitations also. To overcome those limitations we proposed an approach to enhanced data security by separating private data content from metadata concerning its origin and semantics. An unknown identification system based on public-key cryptography provides billing of anonymous customers without connecting their private data to their authentications.

References

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View Of Cloud Computing," Communications of the ACM (CACM), vol. 53, no. 4, pp. 50–58, 2010.
[2] Amazon, "Amazon elastic compute cloud (Amazon EC2)," http://aws.amazon.com/ec2, 2012.
[3] Microsoft, "Windows Azure: Cloud Computing," http://www. windowsazure.com/, 2012.
[4] Apache, "Apache Hadoop," http://hadoop.apache.org/, 2012.
[5] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing And Emerging IT Platforms: Vision, Hype, And Reality For Delivering Computing As The 5th Utility," Future Generation Computer Systems, vol. 25, no. 6, pp. 599–616, 2009.
[6] A. Weiss, "Computing In The Clouds," networker – Cloud Computing: PC Functions Move Onto the Web, vol. 11, no. 4, pp. 16–25, 2007.
[7] Fujitsu Research Institute, "Personal Data In The Cloud: A Global Survey Of Consumer Attitudes," http://www.fujitsu.com/ global/news/publications/ dataprivacy.html, October 2010.
[8] G. Gross, "Cloud Computing May Draw Government Action," IDG News Service, September 2008.
[9] M. D. Ryan, "Cloud Computing Privacy Concerns On Our Doorstep," Communications of the ACM (CACM), vol. 54, no. 1, pp. 36–38, 2011.
[10] D. Chen and H. Zhao, "Data Security And Privacy Protection Issues In Cloud Computing," in Proceedings of the International Conference on Computer Science and Electronics Engineering (ICCSEE), 2012, pp. 647–651.
[11] Y. Lindell and B. Pinkas, "Secure Multiparty Computation For Privacy-Preserving Data Mining,"

Journal of Privacy and Confidentiality, vol. 1, no. 1, pp. 59–98, 2009.
[12] I. Roy, S. T. Setty, A. Kilzer, V. Shmatikov, and E. Witchel, "Airavat: Security And Privacy For Mapreduce," in Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation (NSDI), 2010, pp. 297–312.
[13] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second-Generation Onion Router," in Proceedings of the 13th USENIX Security Symposium, 2004, pp. 303–320.
[14] P. Syverson, "A Taxonomy Of Replay Attacks," in Proceedings of the 7th IEEE Computer Security Foundations Workshop (CSFW), 1994, pp. 187–191.
[15] A. Pfitzmann and M. Hansen, "A Terminology For Talking About Privacy By Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, And Identity Management," http://dud.inf.tu-dresden.de/Anon Terminology.shtml, August 2010, v0.34.
[16] N. Hopper, E. Y. Vasserman, and E. Chan-Tin, "How Much Anonymity Does Network Latency Leak?" ACM Transactions on Information and System Security (TISSEC), vol. 13, no. 2, 2010.
[17] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka, and J. Molina, "Controlling Data In The Cloud: Outsourcing Computation Without Outsourcing Control," in Proceedings of the ACM Workshop on Cloud Computing Security (CCSW), 2009, pp. 85–90.
[18] C. Mitchell, Ed., Trusted Computing. London, UK: The Institution of Engineering and Technology, 2005.
[19] R. Huang, X. Gui, S. Yu, and W. Zhuang, "Research On Privacy-Preserving Cloud Storage Framework Supporting Ciphertext Retrieval," in Proceedings of the International Conference on Network Computing and Information Security n(NCIS), 2011, pp. 93–97.
[20] Q. Liu, G. Wang, and J. Wu, "Secure and privacy preserving keyword searching for cloud storage services," Journal of Network and Computer Applications (JNCA), vol. 35, no. 3, pp. 927–933, 2012.
[21] C. Dwork, "Differential privacy: A Survey of results," in Proceedings of the 5th International Conference on Theory and Applications of Models of Computation (TAMC), 2008, pp. 1–19.
[22] P. Samarati, "Protecting Respondents' Identities In Microdata Release," IEEE Transactions on Knowledge and Data Engineering (TKDE), vol. 13, no. 6, pp. 1010–1027, 2001.
[23] J.-S. Xu, R.-C. Huang, W.-M. Huang, and G. Yang, "Secure Document Service For Cloud Computing," in Proceedings of the 1st International Conference on Cloud Computing (Cloud- Com), 2009, pp. 541–546.
[24] A. Greenberg, "The Tor Project's New Tool Aims To Map Out Internet Censorship," Forbes, April 2012.
[25] ExpressionTech and The Tor Project, "Tor Cloud project," https://cloud.torproject.org/, 2012.
[26] R. Laurikainen, "Secure And Anonymous Communication In The Cloud," Aalto University School of Science and Technology, Department of Computer Science and Engineering, Tech. Rep. TKK-CSE-B10, 2010.
[27] D. Chaum, "Security Without Identification: Transaction Systems To Make Big Brother Obsolete," Communications of the ACM (CACM), vol. 28, no. 10, pp. 1030–1044, 1985.
[28] J. Camenisch and E. V. Herreweghen, "Design And Implementation Of The Idemix Anonymous Credential System," in Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS), 2002, pp. 21–30.
[29] S. Zarandioon, D. Yao, and V. Ganapathy, "K2C: Cryptographic Cloud Storage With Lazy Revocation And Anonymous Access," in Proceedings of the 7th International ICST Conference on Security and Privacy in Communication Networks (SecureComm), 2011, pp. 491–510.
[30] D. Slamanig, "More privacy for cloud users: Privacy Preserving Resource Usage In The Cloud," in Selected Papers from the 4th Hot Topics in Privacy Enhancing Technologies (HotPETs), 2011, pp. 15–27.
[31] J. Camenisch and A. Lysyanskaya, "Signature Schemes And Anonymous Credentials From Bilinear Maps," in Proceedings of the 24th Annual International Cryptology Conference (CRYPTO), 2004, pp. 56–72.
[32] M. Jensen, S. Sch¨age, and J. Schwenk, "Towards An Anonymous Access Control And Accountability Scheme For Cloud Computing," in Proceedings of the IEEE 3rd International Conference on Cloud Computing (CLOUD), 2010, pp. 540– 541.
[33] M. Backes, J. Camenisch, and D. Sommer, "Anonymous yet accountable access control," in Proceedings of the ACM Workshop on Privacy in the Electronic Society (WPES), 2005, pp. 40–46.
[34] T. Abbott, K. Lai, M. Lieberman, and E. Price, "Browserbased Attacks on Tor," in Proceedings of the 7th International Conference on Privacy Enhancing Technologies (PET), 2007, pp. 184–199.
[35] M. J. Freedman and R. Morris, "Tarzan: A Peer-To-Peer Anonymizing Network Layer," in Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS), 2002, pp. 193–206.
[36] J. Abawajy, "Determining Service Trustworthiness In Intercloud Computing Environments," in Proceedings of the 10th International Symposium on Pervasive Systems, Algorithms, and Networks, 2009, pp. 784–788.
[37] R. K. L. Ko, P. Jagadpramana, M. Mowbray, S. Pearson, M. Kirchberg, Q. Liang, and B.-S. Lee, "TrustCloud: A Framework For Accountability And Trust In Cloud Computing," in Proceedings of the IEEE World Congress on Services. (SERVICES), 2011, pp. 584–588.
[38] T. Nguyen and W. Shi, "Improving Resource Efficiency in Data Centers Using Reputation-Based Resource Selection," in Proceedings of the International Conference on Green Computing (GREENCOMP), 2010, pp. 389–396.
[39] Safwan Mahmud Khan and Kevin W. Hamlen, "Anonymous Cloud :A Data Ownership Privacy Provider Framework in Cloud Computing".
[40] Elisa Bertino, Federica Paci and Rodolfo Ferrini, " Privacy Preserving Digital Identity Management for Cloud Computing". Bulletin of IEEE computer society technical Committee on Data Engineering .

[41] C. wang, Q Wang, K. Ren and W. Lou, " Privacy – Preserving Public Auditing for Data Storage Security in Cloud Computing" in the proceeding of IEEE on INFOCOM , 2010, pp 1-9.

Author Profile

Pramila Kharat: is a postgraduate student at Pune institute of computer technology, Pune. She has received B. E. from Marathwada Institute of Technology, Aurangabad in 2002. Her research area is cloud computing and security .

Amar Buchade: is Research Scholar at College of Engineering, Pune. He has received B.E. and M.E. in Computer Engineering from Walchand College of Engineering, Sangli in 2002 and 2005 respectively. His research area is Distributed System, Cloud computing and Security.