INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 3)
Available online at: www.ijariit.com

Data sharing and privacy preserving on personal healthcare information in cloud

A. Aaro
aaroantony96@gmail.com
Anna University BIT-Campus, Tiruchirappalli, Tamil Nadu

S. Usha
ushaanbu2014@gmail.com
Anna University BIT-Campus, Tiruchirappalli, Tamil Nadu

ABSTRACT

Cloud computing is an Internet-based computing model which provides several resources through the cloud service provider to cloud users on demand based. Cloud computing storage, users are able to access and shared resources offered by cloud services provider at a lower cost. Data protection, reliability, dependency, transparency, data portability are the most important challenges of the various cloud environment. In Healthcare system privacy is the ability to protect the patient's personal information against unauthorized access. While sharing the healthcare data in the cloud, the clients have handled security and scalable key management. In this proposed work of Merkle hash tree algorithm used to significantly reduce the amount of data that a TPA has to maintain to proof the integrity of the data. CP-ABE perform private verification, delegated verification, and public verification. A privacy-preserving OPOR auditing scheme for shared data in the cloud environment and utilize with signature, to compute verification information on shared data

Keywords— *Healthcare System, Data security, Merkle hash tree algorithm, CP-ABE, OPOR*

1. INTRODUCTION

Cloud Computing is one of the popular technology in IT that provides various services to the user via the Internet. Cloud system empowers the information sharing system which gives the variety of services to the user. Sharing of data is the higher priority task which plays an important role in any organization by which the productivity in the cloud environment is increased. The common cloud services are effectively available by the on-request network access service as well as it is flexible which is available at a lower cost. At the time of the sharing of information, the medical information or data sharing assumes a fundamental part in light of the fact that the patient data are effortlessly open with the least cost. Healthcare system has collected valuable information about the individuals in our societies that contain sensitive information, e.g. medical data. Researchers need to access and analyze such data using big data technologies in cloud computing, while organizations are required to enforce data protection compliance. Now day to day life the health record of the person is exchanged technology in applications of medications that are utilized for generating, managing as well as modifying the health data related to the patient in a very effective way. The health records of the person have different data related to the patient such as identification sheet, issues, medical records, progress notes, details of the consultation, lab reports, immunization records, consent forms, imaging, and x-ray reports, etc. Data records must be stored on the cloud as well as an access mechanism that is utilized for controlling the activities of the patient. In the personal health record, sharing of the data is fine-grained access control, security, data confidentiality, authorization, and authentication is a crucial challenge while sharing the personal health records in the third party storage. At the time of uploading of personal health care data in the cloud the owner of data losses the physical control also it can be hacked by hackers. Hence the providing security is a big issue while sharing personal health care data in a cloud environment. This can be solved by using an encryption mechanism at the time of data sharing that will increase the confidentiality of the data as well as information security in the third-party storage service. By making use of several encryption techniques user can store the data on a cloud without worrying about the security.

1.1 Challenges faced by the Healthcare organization

(a) Security and privacy: Increase the trust of electronic data needs to secure and protect information. Security breaches can have a negative effect on the organization's brand and credibility, usually resulting in reduced revenue.

(b) Quickness: Organizations want to deploy their new applications quickly to their user populations, in order to respond quickly to market events. Increasing the popularity of mobility has made the ability to have a common security approach across both Web and mobile apps to be a strong requirement.

(c) User expediency: Doctors and nurses don't have a long time to do the redundant and lengthy process; they wish to use their valuable time efficiently. This system provides easy and quick access to store patient information.

(d) Cost control: Because of the cost pressure health care system need an innovative idea to reduce IT expenses.

Cloud computing models and automation of key identity-related process can help to improve the efficiencies and reduce the cost.

(e) Integration with leading healthcare applications: Because of the organizational infrastructure the data move inside and outside the network perimeter, this scenario facing the several threads Insiders, and especially administrators, pose a significant risk due to the damage they can cause through malicious or inadvertent actions. In addition, external attacks are increasing in sophistication and frequency. With the help of an integrated security scheme above said challenges to be removed.

2. RELATED WORK

L. M. Kaufman et al (2009) has proposed this environment strives to be dynamic, reliable, and customizable with guaranteed quality of service. Seal Cloud Services is a consistent and secured managed services where seal deploys and manages the entire software stack and infrastructure. The end-to-end, the mean that content is encrypted at cloud-based storage and delivery channels. Design k-out-of-n secret sharing and broadcast revocation protocols to renew the shared secret key in a scalable fashion.

Min Chen et al (2016) is build up a novel healthcare system by utilizing the flexibility of cloudlet. The functions of cloudlet consist of privacy protection, data sharing, and intrusion detection. In the phase of Number Theory Research Unit (NTRU) method to encrypt user as body data collected by wearable devices. Those data will be an end to nearby cloudlet in an energy efficient fashion. Secondly, present a new trust model to help users to select trustable partners who want to share stored data in the cloudlet. Next, divide users medical data stored in the remote cloud of the hospital into three parts, and give them proper protection. To protect the medical data from malicious users, design an intrusion detection system (IDS) method depend on cloudlet mesh, which can effectively prevent the remote healthcare big data cloud from attacks.

Ning Cao et al (2014) is defined and solved the challenging issue of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE). The efficient similarity measure of coordinate matching to capture the relevance of data documents to the search query. The additional use of inner product similarity to quantitatively evaluate such similarity measure. First, suggest a basic idea for the MRSE based on secure inner product computation, and then give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. A thorough study of inspecting privacy and efficiency guarantees of proposed schemes is given.

K. Zhang et al (2014) has proposed a priority based health data aggregation (PHDA) scheme with privacy preservation for cloud assisted WBANs to improve the aggregation efficiency between different types of health data. Exactly, first, explore social spots to help forward health data and enable users to select the optimal relay according to their social ties. Affording to distinct data priorities, the adjustable forwarding methods can be selected to forward the user as health data to the cloud servers with reasonable communication overheads. The security analysis describes that the PHDA can achieve identity and data privacy preservation, and resists the forgery attacks.

R. Lu et al (2013) has developed a secure and privacy-preserving opportunistic computing framework, called SPOC, for an m-Healthcare emergency. With SPOC, smartphone resources involving computing power and energy can be opportunistically collected to process the computing-intensive personal health information (PHI) during an m-Healthcare emergency with minimal privacy disclosure. They introduce an efficient user-centric privacy access control in SPOC framework, which depends on an attribute-based access control and a new privacy-preserving scalar product computation (PPSPC) technique and permits a medical user to decide who can participate in the opportunistic computing to support in processing his vast PHI data.

Jiawei Yuan et al (2015) has proposed a public auditing scheme consists of four algorithms (Key Gen, Sig Gen, Gen Proof, and Verify Proof). First one is a key generation algorithm that is run by the user to set up the scheme. Sig Gen is used by the user to generate verification metadata, which may consist of MAC, signatures or other related information that will be used for auditing [2]. Generation Proof is run by the cloud server to generate a proof of data storage correctness, while Verify Proof is run by the TPA to audit the proof from the cloud server.

Zhu.Y et al (2012) have introduced a model, for provable data possession (PDP) that allows a client that has stored data at an untrusted server to verify that the server possesses the original data without retrieving it. The data owner keeps a constant amount of metadata to verify the proof. A PDP protocol checks that an outsourced storage site retains a data, which consists of a collection of n blocks. The data owner pre-processes the data, generating a piece of metadata that is stored locally, transmits the data to the server and may delete its local copy. The server stores the data and responds to challenges issued by the data owner.

Zheng. Q et al (2011) have proposed Proof-of-Retrievability to allow data owners to efficiently and securely verify that the storage servers stores their data correctly. Proof-of-retrievability (POR) schemes have been proposed wherein storage server must prove to a verifier that all of a client's data are stored correctly. Though existing POR schemes offer decent solutions addressing various practical issues, they either have a linear or quadratic communication complexity and only the data owner can verify the remotely stored data. It remains open to design a POR scheme that achieves both public verifiability and constant communication cost simultaneously.

Chen and Zhao (2011&2012) have discussed the consumer's concern regarding moving the data to the cloud. According to Chen and Zhao, one of the foremost reasons why large enterprises still would not move their data to the cloud is security issues. Authors have developed an outstanding analysis on data security and privacy protection issues related to the cloud.

3. PROPOSED SYSTEM

This project focuses on the issues related to data security in cloud computing. By way of information will be shared with a third party, cloud computing users want to avoid an untrusted cloud provider. Protecting private information, such as a patient's medical records from attackers or malicious insiders is of critical importance. To product, the medical information using Ciphertext-Policy attribute-based encryption tool perform private verification, delegated verification and public verification. A Merkle hash tree algorithm is dividing a file into fragments, and replicate the fragmented data over the cloud nodes. All the nodes store only a single fragment of data and ensure in the case of a successful attack, no meaningful information is revealed to the attacker.

3.1 Drawbacks of the existing system

- Identity privacy can be leaked at the place of TPA
- Duplicated data can be stored in the cloud service provider.
- Data encryption is large so small users with limited computational power (PDAs, mobile phones, etc.).
- PDP mechanism only provides public auditing in cloud storage
- Provide large time for retrieving the documents

Cipher one of most preferable identity-based cryptographic systems where attributes are taken as input and cryptographic operations are done on those attributes based on defined policies. In user's identity is composed of a set, S, of strings which serve as descriptive attributes of the user. For our case, the patient's attributes are contact information, disease, medicines, medical tests, etc. There is another set of attributes S1 that helps us to decrypt a message if his identity S has at least k attributes matched with the set S1, where k is a parameter set by the administrator of the system. The Attribute-Based Encryption system only needs to know the description of the user for determining his secret key.

(a) Setup (k): Algorithm takes security parameter and attributes value as input, k and outputs as a master key MK to generate secret keys in the Key generation algorithms and a set of public parameters PK

(b) Key-Gen (S, MK): The authority executes the Key-Gen algorithm for generating a new secret key SK. The algorithm takes as input a set of user's attributes, S and the master-key MK and outputs as a secret key SK corresponding to S.

(c) Encryption (M, S1, PK): This Encryption algorithm takes input a message M, with a set of attributes S1 (access structures), and the public parameters PK. It outputs a ciphertext, CT.

(d) Decryption (CT, S1, S, SK): The Decryption algorithm is run by a user with identity S and secret key SK to decrypt a ciphertext C that has been encrypted with S1. With the help of identity attributes S and secret key SK, this decryption algorithm is used to decrypt a ciphertext CT that has been encrypted by S. If the set |S union S1| is greater than or equal to k this algorithm shows the decrypted message M.

4. SYSTEM MODEL

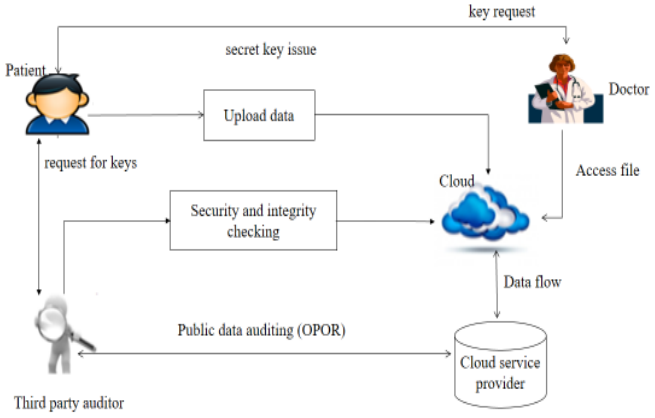


Fig. 1: System framework

The system model involves the parties are the patient, a group of users (doctors), cloud service provider and a public verifier (TPA). The original user and a number of group users are available in a group. The original user, to begin with, creates shared data in the cloud and shares it with group users. Each member of the group is allowed to access and modify shared data. Shared data and its verification metadata (i.e. signatures) are both stored in the cloud server. A third-party auditor (TPA) providing authority data auditing services or a data user outside the group intending to utilize shared data, in public verify the integrity of shared data stored in the cloud server. When a TPA needs to check the integrity of shared data, it first sends an auditing challenge to the cloud server. After receiving the auditing challenge, the cloud server responds to the public verifier with an auditing verification of the control of shared data. Then, this public verifier checks the correctness of the entire data by verifying the correctness of the auditing proof. Basically, the process of public auditing is a challenge and reply procedure between a public verifier and the cloud server.

4.1 Cloud Framework

In this module, cloud data storage service three different entities such as the cloud user, who has large amount of data files to be stored in the cloud; the cloud server, which is managed by the cloud service provider to provide data storage service and has significant storage space and computation resources; the third-party auditor, who has expertise and capabilities that the cloud users do not have trusted to assess the cloud storage service reliability on behalf of the user upon request.

4.2 Registration

In this module is used to store the cloud owner registration details. The Registration details include a user name, password, address, phone number, E-mail, etc. The verification code will be sent to the owner mail id for authentication purpose. When the cloud owner wants to enter this system he must enter this code.

4.3 Upload Data

This module is used to upload healthcare data into a cloud server at any time. When uploading the data is secret will be generated by using a key generation algorithm. Each uploaded data's will have a unique key and this key is forwarded to the user registered mail id.

4.4 Admin module

The Admin will manage the entire application and storage device. Admin can monitor the entire files and users stored on the server at any time. But admin cannot download user files without user permission. Admin can view files and content at any time but he cannot download the file without user acceptance. If Admin wants to download the file he will send the request information to the user through this system. After the response from the user only can download the data because of privacy.

4.5 Third Party Auditor Module

In this module of a third party, the auditor is to verify the users and files stored on the cloud. They can audit those things but they cannot view the content or download the files stored on the cloud.

This module classified into two categories

1) Data integrity : This module is used to perform the privacy-preserving public auditing in Cloud Computing, TPA may concurrently handle multiple auditing delegations upon different users' requests. The separate auditing of these tasks for TPA can tedious and very inefficient. Batch auditing is not only allowed TPA to perform with multiple auditing tasks simultaneously but also greatly reduces the computation cost on the TPA side and also checking data integrity.

2) Data Dynamics Module: Hence, supporting data dynamics for privacy-preserving public risk auditing is Outsourcing proof

of retrievability also paramount importance. Now it shows how our main scheme can be adapted to build upon the existing work to support data dynamics, including block-level operations of modification, deletion, and insertion. They can adopt this technique in our design to achieve privacy-preserving public risk auditing with support of data dynamics.

4.6 Download

It is used to download the health data but without the secret key, this system will not allow downloading the file.

4.7 Key Management

In this module to protect the encryption keys from loss, corruption, and unauthorized access. More processes can be used to control key management, including changing the keys regularly, and managing how keys are assigned and who gets them. Merkle hash tree algorithm used to perform operations on encrypted data without knowing the private key, the client is the only holder of the secret key and MHT contains: KeyGen, Sign and Verify. In KeyGen, all user in the group generates his/her public key and private key. In Sign, a user in the group is able to generate a signature on a block and its block identifier with his/her private key and all the group member's public keys.

4.8 Secure Data Sharing

Each user is assigned to a data owner from the Provider. Each user can freely get the ciphertexts from the server. To decrypt a ciphertext, each user may submit their secret keys issued by data owner together with its global public key to the server and ask it to generate decryption token for some ciphertext. Upon receiving the decryption token, the user can decrypt the ciphertext by using its global secret key. The users those who are having matching keys as in the access policy defined in the ciphertext can retrieve the entire data content.

5. SYSTEM IMPLEMENTATION

This stage of implementation in our project when the theoretical design is revolved out into a working system. The most critical stage of achieving a new system in successful and giving the user, confidence that the new system work will be effective. These stage of implementation is investigated the previous system and its constraints on developing, designing of methods to achieve conversion and evaluation of changeover methods.

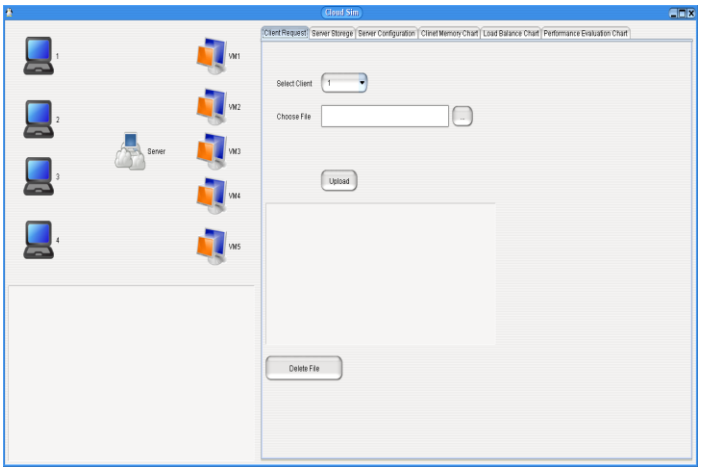
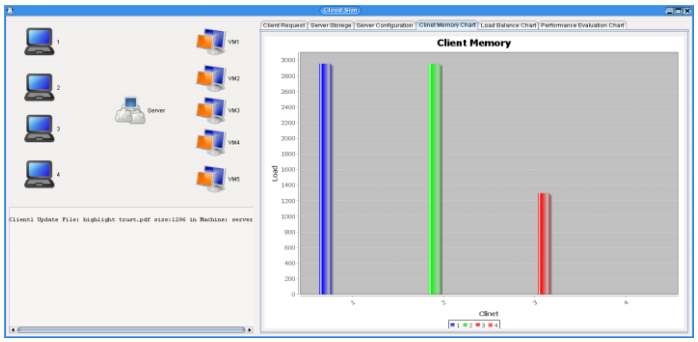


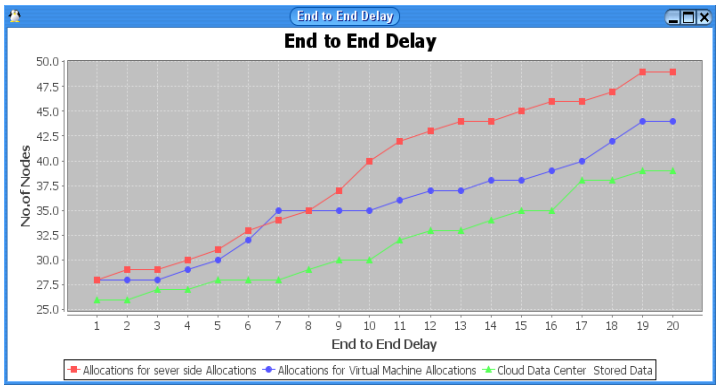
Fig. 2: Cloudsim toolkit

5.1 Performance Measurement

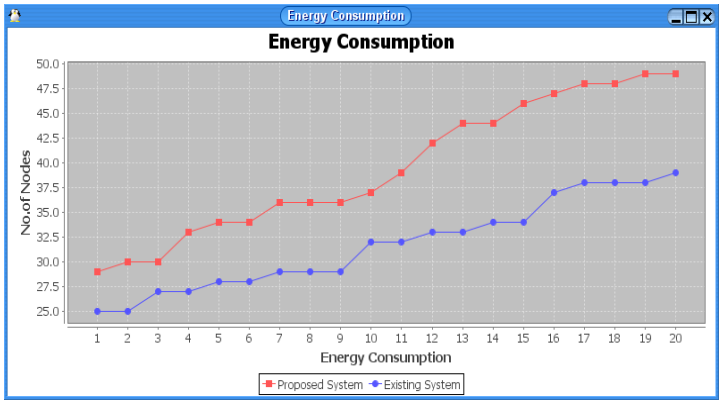
This paper has provided some simulation results that we have found using the clouds toolkit. Calculated total time utilization for generating key, encryption, and decryption. In figure 3 shows our simulation result based on memory allocation, load balancing time, end-to-end delay, data delivery ratio and finally calculated energy consumption.



(a) Memory allocation



(b) Data delivery ratio



(c) Energy consumption

Fig. 3: Performance analysis using Cloudsim toolkit

6. CONCLUSION

In this paper, demonstrated the challenges of storing data using cloud computing technology. Accordingly, we have proposed a Health Cloud architecture for secured data management of the patients. To implement the Ciphertext-Policy Attribute-Based Encryption **(**CP-ABE**)** algorithm within our proposed Security Manager module in our system and also using the Merkle hash tree algorithm is used to checking the data integrity. Through performance evaluation, we have measured the time needed for data delivery, energy consumption, and time utilization. We are claiming that if we implement the CP-ABE algorithm in the cloud then there will be less performance overhead for the security and confidentiality of the data. Because it does not need to check whether a user in the cloud is a doctor or other medical staff or patient. It would be more efficient if we can implement the whole "Health Cloud" architecture and find out the limitations. In our future research, we are planning to implement the whole system and go through the more

performance evaluation of key generation time, encryption time and decryption time.

7. REFERENCES

[1] L. M. Kaufman, "Data Security in the world of Cloud Computing," Security & Privacy, IEEE vol. 7, no.4, pp. 61–64, 2009.

[2] Min Chen, Yongfeng Qian, Kai Hwang, Shiwen Ma, and Long Hu, "Privacy Protection and Intrusion Avoidance for Cloudlet-based Medical Data Sharing", IEEE Transactions on Cloud Computing, 2016, pp. 1–1.

[3] Ning Cao, Cong Wang, Ming Li, Kui Ren, and Wenjing Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data", Parallel And Distributed Systems, IEEE Transactions on, vol. 25, no.1,2014, pp. 222-233, 2014

[4] K. Zhang, X. Liang, M. Baura, R. Lu, and X. S. Shen, "Phda: A priority-based health data aggregation with privacy preservation for cloud assisted banks," Information Sciences, vol. 284, pp. 130–141, 2014.

[5] R. Lu, X. Lin, and X. Shen, "Spoc: A secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency," Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 3, pp. 614–624, 2013.

[6] Jiawei Yuan and Shucheng Yu "Public Integrity Auditing for Dynamic Data Sharing With Multiuser Modification" IEEE Trans. Information Forensic and security. Syst., vol. 10, no.8, 2015.

[7] Zhu. Y, Hu.H, Ahn G.J, and Yu. M., "Cooperative provable data possession for integrity verification in multi-cloud storage," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 12, pp. 2231–2244, 2012.

[8] Zheng. Q and Xu. S, "Fair and dynamic proofs of retrievability," in CODASPY, 2011, pp. 237–248.

[9] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," J. Netw. Comput. Appl., vol. 34, no. 1, pp. 1–11, Jan.2011.

[10] F. Zhang and H. Chen, "Security-Preserving Live Migration of Virtual Machines in the Cloud," J. Netw. Syst. Manag.,pp.562–587,2011.

