| DOI: 10.15680/IJIRCCE.2020.0812028 |

A Data Privacy Protection Cloud-Based Systems for Protecting the Sensitive Health Data Stored on the Cloud

Ms. Vaishali Vijay Wagh, Prof. Kishor N. Shedge

PG Student, Dept. of Computer Engineering, SVIT, Nashik, Maharashtra, India

Assistant Professor, Dept. of Computer Engineering, SVIT, Nashik, Maharashtra, India

ABSTRACT: Cloud computing is a revolutionary mechanism that changing way to enterprise hardware and software design and procurements. Because of cloud simplicity everyone is moving data and application software to cloud data centres. The Cloud service provider (CSP) should ensure integrity, availability, privacy and confidentiality but CSP is not providing reliable data services to customer and to stored customer data, in order to preserve the customers' trust. Nevertheless, in certain applications, such as medical health records for example, the medical facility is responsible for preserving the privacy of the patients' data. Although the facility can offload the overhead of storing large amounts of data by using cloud storage, relying solely on the security measures taken by the CP might not be sufficient. Any security breach at the CP's premises does not protect the medical facility from being held accountable. This work aims to solve this problem by presenting a secure approach for storing data on the cloud while keeping the customer in control of the security and privacy of their data.

KEYWORDS: Cloud computing, security, privacy, encryption, hash functions. Cloud service provider (CSP), cloud data storage, security issues, policies & protocols.

I. INTRODUCTION

Cloud computing is one of the emerging technologies that has an increasing impact on both private and public sectors. It represents an on-demand service model for delivering resources ranging from storage and data access, via computation to software provisioning. Typical categories of cloud computing include Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). In the IaaS category, virtual machines are provided, such as Amazon EC2. PaaS offers development tools such as Microsoft Azure, while SaaS providing software OnDemand usually over the Internet, eliminating the need for installing and maintenance of the software on the client's computer, such as Google Docs. The most important benefits of cloud computing include scalability, cost reduction, data availability, reliability and resilience.

The Cloud Service Provider (CSPs) has promise to ensures the data security over stored data of cloud clients by using methods like firewalls and virtualization. These mechanisms would not provide the complete data protection because of its vulnerabilities' over the network and CSPs have full command on cloud applications, hardware and client's data. Encrypting sensitive data before hosting can deserve data privacy and confidentiality against CSP. A typical problem with encryption scheme is that it is impractical because of huge amount communication overheads over the cloud access patterns. Therefore, cloud needs secure methods to storage and management to preserve the data confidentiality and privacy.

Therefore, in this work, an approach that can complement any security measures taken by the CP is proposed. It can also be implemented even if the CP does not protect the data via encryption. The proposed approach does not involve redundancy or increased storage costs, but can easily be implemented in conjunction with redundancy techniques. The main objective of the proposed method is to offload the overhead of storing large amounts of data from the client to the CP, while still allowing the client to be in charge of the security and privacy of their data, thus providing an additional layer of privacy for critical applications such as health records.

II. RELATED WORK

Health data has precisely recorded people's illnesses, and medical records and the secure storage and sharing of medical data and patient privacy protection have increasingly become a priority to build intelligent hospitals. Traditional data access control technology builds and implements a safety access strategy with a completely reliable server, making it difficult to get adapted to the distributed network environment in modern times. Featured by decentralization and trust lessness, blockchain has given people a brand-new idea through distributed data storage, reliable point-to-point transmission, a consensus mechanism, and encryption algorithms. As an encryption mechanism that uses attribute as a public key, the mechanism, in essence, links users with ciphertexts through the attribute. Its flexibility of encryption and access control form has greatly ensured the security of cloud data storage. In the meanwhile, it has also achieved fine-grained access and become the key technique for secure cloud storage access control.

It consists of an approach to upload data to the cloud, an approach to download data from the cloud, and a management process implemented by the client in order to upload to, and retrieve data from, multiple cloud providers.

Dubovitskaya et al. propose a framework to create an infrastructure for medical data management that makes it possible for healthcare professionals to disclose patients' information; however, the privacy of the patient is preserved. For that purpose, they use generalization, represented by binary trees, and pseudonyms generated by multiple key searchable encryption. Their architecture consists of databases on both the client side and the server side, a cryptographic module on the client side, an anonymization module on both sides, and a standalone Certification Authority. The cryptographic module performs a multi-key searchable encryption, encrypts EHRs, and generates signatures to authenticate the data. The anonymization module allows healthcare professionals to upload patient's medical information for research purposes while preserving the anonymity of the research database. Certification Authority issues public key certificates and smartcards. Smartcards store private keys and have PINs known only to their owners. To share and access patients' data stored at the server side database (i.e., Data Repository), the patient generates a key and shares it with the caregiver using a card reader device located at the caregiver's office. The EHR is encrypted using that shared key and is signed using the caregiver's secret key. Although it is not easy to guess the PIN, the card is protected only with that PIN. If the PIN is found, the patient's private key will be disclosed which limits the security of this framework

III. PROPOSED APPROACH

In this section, the details of the proposed work are presented to develop a secure framework for sharing Health Data over Clouds. The framework enables patients at a healthcare provider, to retrieve their health data, in entirety or partially, from a remote healthcare provider. The framework also allows targeted healthcare providers to authenticate the patients and requesting healthcare providers before processing a sharing request. Requested EHRs are encrypted using a shared symmetric key between communicating healthcare providers.

A. *Storing Data with multiple cloud providers*

- The proposed approach to store data on the cloud consists of the following steps:

1) The client who wants to store data on the cloud subscribes with n CPs.

2) A file F to be stored on the cloud is subdivided into n parts, or subfiles: F1, F2, ... Fn.

3) Each part is encrypted with an encryption key KF. The encrypted parts are denoted by F1*, F2*, ..., Fn*, where Fi*={Fi}KF. It should be noted that, in this paper, symmetric encryption is assumed; i.e., the encryption and decryption keys are the same.

4) A random (or pseudorandom) permutation vector P(F) is generated.

5) The encrypted parts are stored on the n clouds according to P(F); for example, depending on P(F), F1* could be stored on CP3, F2* on CPn, etc.

B. Managing Data with Multiple CPs

In order to be able to retrieve their files, the client needs to maintain some information related to the distribution of the various parts and to the encryption of the file. Thus, the client maintains two tables.

1) The first table contains a hash of the file name H(F_name), and the key KF.

2) The second table contains a hash of the file name H(F_name), a hash of the file content itself (unencrypted) H(F), a hash of the encrypted file content H(F*), and the permutation vector P(F), encrypted with KFc.

C. *Retrieving Data Stored on the Cloud with Multiple CPs*

In order to obtain a file from the cloud using the proposed approach, the following steps are followed

1) The client enters the file name.
2) The data stored under this filename with the various CPs is retrieved. However, it is encrypted, and in a scrambled order.
3) Then, the system computes the hash value of the name, finds the corresponding entry in Table 1, and obtains the key.
4) The key is used to decrypt the permutation vector P(F*)
5) Afterwards, the encrypted parts that were downloaded from the different cloud providers are assembled in the correct order by reversing the permutation process (hence the notation P-1(F*).
6) Then, they are decrypted with KF.
7) Finally, the hash values of the encrypted and unencrypted versions are then computed and compared to their corresponding values stored in order to check for the integrity of the file downloaded from the cloud. It should be noted that the hash of the encrypted file H(F*) can be checked before decryption thus reducing the overhead of useless decryption in case the file was corrupted during transfer.

IV. PROPOSED ARCHITECTURE

The aim is to provide a secure EHR sharing framework that defines a set of cryptographic techniques, which ensures the consent of patients to healthcare providers requesting their EHR information from another healthcare provider. Also, the framework provides an authentication protocol for healthcare providers to authenticate patients and requesting healthcare providers in order to ensure the security and privacy of patients' EHR information. The infrastructure, which runs the framework, consists of an EHR Sharing Cloud (private cloud) and a dedicated EHR Sharing Server at each participating healthcare providers' site

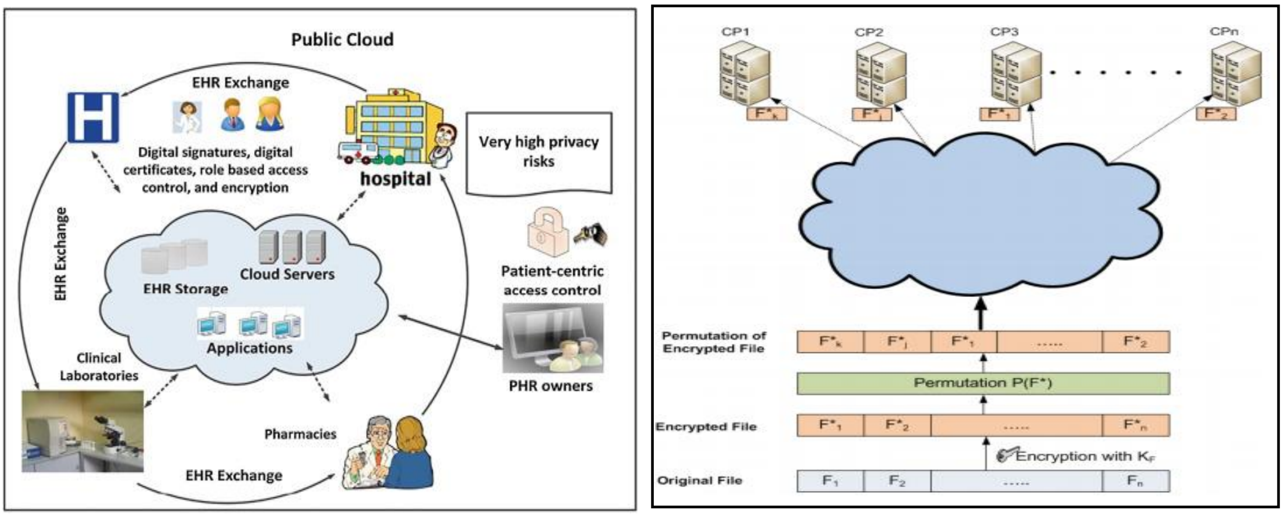

Fig- An example of a public cloud in context of Health & Storing a file on the cloud with the proposed approach.

The EHR Sharing Cloud is a private cloud that connects and serves different healthcare providers. It provides a stable, scalable infrastructure and dynamic secure environment to offer services, store data, and host and run desired applications. The framework requires an EHR Cloud Application to run on the Cloud infrastructure. The EHR Cloud Application provides all required communication between different healthcare providers while maintaining the required security measures in transferring EHRs over the cloud.

The EHR Cloud Application initially issues each participating healthcare provider a certificate, which contains its public key and appropriate parameters. The PKI is used for communication authentication between healthcare providers and the EHR Sharing Cloud. It can also be used to add a second level of confidentiality in transmission by encrypting the request using the EHR Sharing Cloud's public key if needed. The EHR Cloud Application also supports the use of a commitment scheme to allow participating healthcare providers, with no prior trust, to authenticate each other before sharing patients' EHR(s).

When a healthcare provider receives a request, it first authenticates the EHR Sharing Cloud and then authenticates the patient. The EHR Client Application initiates the commitment phase to authenticate the requesting healthcare provider and then the reveal phase takes place. After the success of the reveal phase, the symmetric key received in the request is used to encrypt the requested EHR(s) as a reply to the request. The reply is then sent to the EHR Sharing Cloud, which stores it and forwards it to the requesting healthcare provider. The requesting healthcare provider uses the symmetric key to decrypt the received reply.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented a simple and secure framework for protecting sensitive data stored on the cloud. It is based on splitting a given file into multiple parts, and storing each part, after encryption and permutation of the order of the parts, with a different cloud provider. The information to decrypt and reorder the file parts is stored in separate locations at the client premises. No extra cost is incurred since the total storage size is still the same, compared to the case of storing the file with a single cloud provider. The proposed approach allows the client to benefit from any security measures implemented by the cloud provider while still taking charge of the security and privacy of their data.

Future enhancements consist of adding redundancy techniques, e.g. by using secret sharing methods, to the proposed approach in order to be able to recover the whole data even in case of loss of one or more parts at the cloud providers' premises. Moreover, a proof of concept demonstrating the practical feasibility of the proposed approach will be prepared and described in future works

REFERENCES

1. G. Ramachandra, M. Iftikhar, F. A. Khan, "A Comprehensive Survey on Security in Cloud Computing", Procedia Computer Science, vol. 110, p.p. 465–472, 2017.
2. Ali Sakr1, Elias Yaacoub1,2, Hassan Noura1, Mohammed Al-Husseini2, Khalid Abualsaud3, Tamer Khattab4, Mohsen Guizani5, "A Secure Client-Side Framework for Protecting the Privacy of Health Data Stored on the Cloud" 2018 IEEE Middle East and North Africa Communications Conference, 978-1-5386-1254-5/18.
3. S. Islam, M. Ouedraogo, C. Kalloniatis, H. Mouratidis, and S. Gritzalis, "Assurance of Security and Privacy Requirements for Cloud Deployment Model," IEEE Transactions on Cloud Computing, vol. PP, no. 99, pp. 1–14, 2015.
4. J. H. Hine, W. Yao, J. Bacon, and K. Moody, "An Architecture for Distributed OASIS Services," in Proceedings of International Conference on Distributed Systems Platforms and Open Distributed Processing, Springer, vol. 1795, pp. 104 – 120, April 2000.
5. J. Jin, G.-J. Ahn, H. Hu, M. J. Covington, and X. Zhang, "Patient-centric Authorization Framework for Sharing Electronic Health Records," in Proceedings of 14th ACM symposium on Access control models and technologies. ACM, pp. 125–134, 2009.
6. R. Wu, G.-J. Ahn, and H. Hu, "Secure Sharing of Electronic Health Records in Clouds," in Proceedings of 8th International Conference Conference on Collaborative Computing: Networking, Applications and Worksharing. IEEE, pp. 711– 718, October 2012.
7. W. Itani, A. Kayssi and A. Chehab, "Privacy as a Service: Privacy Aware Data Storage and Processing in Cloud Computing Architectures", Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing (DASC 2009), p.p. 711-716, Chengdu, China, Dec. 2009.
8. A. Mallareddy, V. Bhargavi, K. D. Rani, "A Single to Multi-Cloud Security based on Secret Sharing Algorithm", International Journal of Research (IJR), vol. 1, no. 7, p.p. 910-915, Aug. 2014.
9. P. Pareek, "Cloud Computing Security from Single to Multi-clouds using Secret Sharing Algorithm", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 2, no. 12, p.p. 3261-3264, Dec. 2013.
10. M. K. Alam and K. S. Banu, "An Approach Secret Sharing Algorithm in Cloud Computing Security over Single to Multi Clouds", International Journal of Scientific and Research Publications, vol. 3, no. 4, Apr. 2013.
11. U. S. Department of Health and Human Services, Health Information Privacy; URL [Accessed Dec. 7, 2017] https://www.hhs.gov/hipaa/forprofessionals/privacy/index.html.