ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 5, May 2015

E-Health Care Solutions Using Anonymization

Chaitra.S [1], Narasimha Murthy M S [2]

M.Tech Scholar, CS&E, Acharya Institute of Technology, Bengaluru, India [1]

Assistant Professor, CS&E, Acharya Institute of Technology, Bengaluru, India [2]

Abstract: Cloud computing is a type of computing, instead of having local servers or personal devices to handle applications it trusts on sharing computing resources. Cloud services can help the healthcare industry to access and manage health records effectively in order to provide better patient care. A properly implemented cloud storage system allows hospitals to process tasks effectively and quickly, without causing a drop in performance. This paper mainly focuses on security concerned privacy data enhancement in the cloud environment.. One of the privacy preserving techniques called Anonymization is used, which removes identifying attributes from the database and making the data identification difficult to anybody except the data owners. In this project L-Diversity Technique is used for the Anonymization of the Patient data. the system proposes to integrate key management from pseudorandom number generator for unlink ability, a secure indexing method for privacy preserving keyword search which hides both search and access patterns based on redundancy.

Keywords: Anonymization, Cloud Computing, Data Storage Protection, Privacy Preserving, Security, Shared Data,Data Identification,L-Diversity Technique.

I. INTRODUCTION

The data is stored into cloud by cloud customers to enjoy the high quality networks, servers, services and applications from a shared pool of configurable computing resources. Advantages of cloud computing ubiquitous network access, transference of risk, location independent resource pooling. Sensitive data example personal health records may have to be encrypted by data owners before outsourcing to the commercial public cloud to protect data privacy and combat unsolicited accesses in the cloud and beyond.

Usersshare information in the cloud. Some purpose public or private organizations publish their database on the cloud for research purpose. This database may contain sensitive information about many people. The Hospital tracks its patients with help of this database. The privacy of this data must be preserved while disclosing it to third party or while placing it in long time storage. i.e. any sensitive information should not be disclosed [02].

In recent forefront of new technologies the patient records are being put in electronic format enabling patients to access their records via the Internet and also cloud computing environment. The remote patient is monitoring with more feasible at anytime, anywhere and any place also. The combination of these technologies will improve the quality of health care by making it more personalized and reducing costs and medical errors. While there are benefits to technologies, associated privacy and security issues need to be analyzed to make these systems socially acceptable.

In the cloud Storage Data sharing is important functionality. How efficiently, securely and flexibly share data with others in cloud storage is shown in [01]. According to government website around 8 million patient’s information was leaked in the past two years.

The costs of healthcare services rise and healthcare professionals are becoming hard to find the information. The healthcare organizations are concerned with health information technology (HIT) systems. HIT allows health organizations to provide services in a more efficiently and cost-effectively manner. Computer systems are used extensively in medical and healthcare systems, over the past three decades. for the storage, documentation, presentation of patients information storage devices and storage [06].

Most of the medical records are being stored in the form of electronic records in centralized databases. Cloud is divided into three types based on internally (private Cloud), outsourced (public Cloud) or a combination of the two (hybrid Cloud). E-Health Cloud is concerned with a Service-Based Applications and Gateway. Cloud based HIT solutions used for patients Details. Patient record contains sensitive information for example drug usage, patient disease etc. [08].

One of the privacy preserving techniques that manipulate the information, making the data identification difficult to anybody except the owners is anonymization. It is different from that of data encryption. Anonymization of data removes identifying attributes like names or social securityNumbers from the database [02].

The salient features are offered by the system including efficient key management, especially for retrieval at emergencies, efficient key management and auditability, for misuse of health data. Specifically, the system proposes to integrate key management from pseudorandom number generator for unlink ability, a secure indexing method for privacy preserving keyword search which hides both search and access patterns based on redundancy. The proposed method seeks to develop privacy technology and privacy-protecting infrastructures to facilitate the development of a health information system so that individuals can actively protect their personal information [13].

Privacy is the right of an individual to determine for themselves when, how and to what extent information about them is shared or transfer to others. Recently, health

DOI 10.17148/IJARCCE.2015.4591

information of hundreds of thousands of patients has been made liable to danger due to security lapses at hospital and government agencies [03]. Privacy-preserving health data storage is studied by Sun et al , where patients encrypt their own health data and store it on a third-party server [base paper].

Advantages of cloud computing ubiquitous network access, transference of risk, location independent resource pooling. To protect data privacy and combat unsolicited accesses in the cloud and beyond, sensitive data example personal health records may have to be encrypted by data owners before outsourcing to the commercial public cloud [04].

Major cloud infrastructure providers, such as Amazon, Google, and Microsoft, more and more third-party cloud data service providers are emerging which dedicate to offer more accessible and user friendly storage services to cloud customers. Secondly, to protect data privacy and combat unsolicited accesses in the cloud and beyond, sensitive data, e.g., personal health records,emails, financial transactions, tax documents, etc., may have to be encrypted by data owners before outsourcing to the commercial public cloud this, however, obsoletes the traditional data utilization service based on plaintext keyword search. The trivial solution of downloading all the data and decrypting locally is clearly impractical, due to the huge amount of bandwidth cost in cloud systems [09].

The burden of local data storage and maintenance is created by the Security problems in cloud computing. By utilizing, public auditability [11] the users can resort to an external audit party to the integrity of outsourced data when needed. The TPA's audit process should not bring any new vulnerability towards user data privacy and should not increase the burden of the user.To overcome these disadvantages and to introduce a secure cloud storage with aauditing [5].

Some early works on privacy protection for e-health data concentrate on the framework design, including the demonstration of the significance of privacy for e-health systems. Secret sharing is a mechanism for sharing secret information among multiple entities so that the cryptographic power is distributed which at the same time avoid single point of failure. For (k, n) threshold secret sharing, a piece of information I is divided into n pieces, such that knowledge of any k or more of these pieces can recover information I, while knowledge of (k − 1) or fewer pieces keeps information I completely undetermined.

K.Yang and X.Jia presented a TSAS: Third Party Storage Auditing Services. Traditionally, owners can verify the data integrity predicated on two-party storage auditing protocols. In cloud storage system, however, it is unfavorable to let either side of cloud service providers or owners conduct such auditing, because none of them could be ensured to provide impartial auditing result. In TSAS, both the Data Fragment Technique and Homomorphic Verifiable Tags are applied to improve the performance [07].

Rana.M.Pir presented a Data Integrity Verification in Cloud Storage without using Trusted TPA. TTPA is a reliable independent component which is trusted by both the cloud users and server many researchers recommend the support of trusted third party (TTP). But issues such as TTP becoming bottleneck, data leakage, introduction of new vulnerabilities, scalability, accountability, performance overhead, dynamic data support, extra hardware cost incurred etc[10].

II. LITERATURE SURVEY

Some early works on privacy protection for e-health data concentrate on the framework design, including the demonstration of the significance of privacy for e-health systems. Secret sharing is a mechanism for sharing secret information among multiple entities so that the cryptographic power is distributed which at the same time avoid single point of failure. For (k, n) threshold secret sharing, a piece of information I is divided into n pieces, such that knowledge of any k or more of these pieces can recover information I, while knowledge of (k − 1) or fewer pieces keeps information I completely undetermined.

Yue Tong, Jinyuan Sun, Sherman S. M. Chow, and Pan Li[13], proposed to build privacy into mobile health systems with the help of the private cloud. We provided a solution for privacy-preserving data storage by integrating a PRF-based key management for unlinkability, a search and access pattern hiding scheme based on redundancy, and a secure indexing method for privacy-preserving keyword search. We also investigated techniques that provide access control (in both normal and emergency cases) and auditability of the authorized parties to prevent misbehavior, by combining ABE-controlled threshold signing with role-based encryption.
Privacy preservation of sensitive information is a key factor. Anonmization offers more privacy options rather to other privacy preservation techniques (Randomization, Encryption, and Sanitization). However, Anonmization itself contains several techniques that require concluding best one[14]. According to the presented analysis there is close competition among K- Anonymity, L-Diversity, T-Closeness, P-Sensitive and M-invariance. Analytical comparative analysis is conducted to select optimal Anonymization methods.

Privacy-preserving health data storage is studied by Sun et al.,[11] where patients encrypt their own health data and store it on a third-party server. The backup mechanisms for emergency access rely on someone or something the patient trusts whose availability cannot be guaranteed at all times. Moreover, the storage privacy proposed in this paper is a weaker form of privacy because it does not hide search and access patterns.
It may be argued that medical information systems are subject to the same type of threats and compromises that plague general information systems, and that it does not require special attention from a research viewpoint[12]. D. Boneh and M. Franklin[15]. proposed a fully functional identity-based encryption scheme (IBE). The scheme has chosen cipher text security in the random oracle model assuming a variant of the computational Diffie Hellman problem.

DOI 10.17148/IJARCCE.2015.4591

ABE (Attribute-Based Encryption) has shown its promising future in fine-grained access control for outsourced sensitive data. Typically, data are encrypted by the owner under a set of attributes. The parties accessing the data are assigned access structures by the owner and can decrypt the data only if the access structures match the data attributes.

SSE allows data owners to store encrypted documents on remote server, which is modeled as honest-but-curious party, and simultaneously provides away to search over the encrypted documents[8]. More importantly, neither the operation of outsourcing nor keyword searching would result in any information leakage to any party other than the data owner, thus achieving a sound guarantee of privacy.

III. METHODOLOGY

An extrinsic auditor to audit the user's outsourced data is supported without learning knowledge on the data content. Cloud computing components are classified as:

- Cloud User (CU)
- Cloud Service Provider (CSP) & Cloud Server (CS)
- Third party Auditor (TPA)

Now let's get to know the component working for cloud computing in detail.

A. Cloud User (CU)

Cloud user has large amount of data files that is stored in the cloud. A user depends on the CS for cloud data storage and maintenance. They may also dynamically collaborate with the CS to access and update their stored data for various application purposes.

B. Cloud Service Provider (CSP) & Cloud Server (CS)

Services made available to users on demand via the Internet from a cloud computing provider's servers as opposed to being provided from a company's own on-premises servers. Cloud services are designed to provide simple, flexible access to applications, resources and services, and are fully handled by a cloud services provider.

Cloud servers work in the same way as physical servers but the functions they provide can be very different. Instead of renting or purchasing physical servers, clients are renting for virtual server space when electing for cloud hosting.

The key benefits of cloud servers:

- Flexibility and scalability
- Cost-effectiveness
- Ease of set up
- Reliability

C. Third party Auditor (TPA)

The third party auditor (TPA), who has expertise and capabilities that cloud users do not have and is trusted to assess the cloud storage service security when requested on behalf of the user. Users believe on the CS for cloud data maintenance and storage. They may additionally dynamically interact with the CS to access and update their stored data for various application purposes. The users may resort to TPA for ascertaining the storage security of their outsourced data, protect their data from TPA. However, due to their own benefits the CS might neglect to keep or deliberately delete rarely accessed data files which belong to ordinary cloud users when the cloud data storage based services are being provided. We assume the TPA, who is in the business of auditing, is independent and reliable, and thus has no incentive to collude with either the CS or the users during the auditing process.

Advantages:

- Avoiding local storage of data.
- Reduction in the costs of maintenance, storage and personnel.
- The chance of losing data is reduced by hardware failures.
- Not cheating the owner.

Here user needs to register himself to store the data in the cloud. When he successfully logs in, he can upload the data in the cloud. Then dataset is generated which is anonymized and encrypted in the cloud. In this module is used to send the auditor request to the TPA.if valid user can see the decrypted information. The detailed explanation is as shown in below modules with figure 1.

DOI 10.17148/IJARCCE.2015.4591

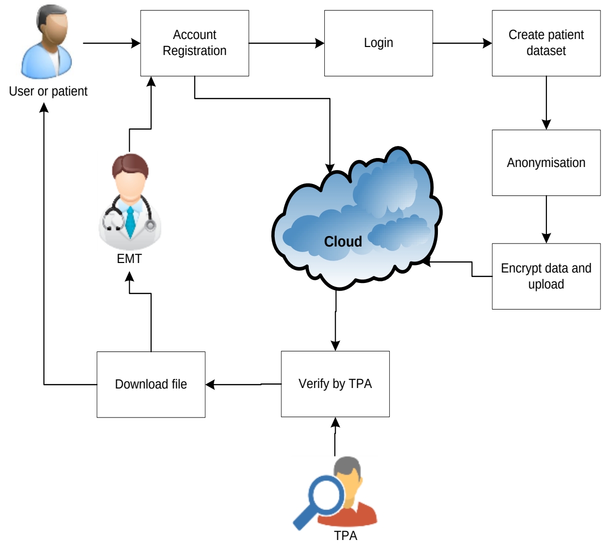

Figure 1: Block Diagram for Proposed System.

IV. Implementation

A. User Registration

In this user registration module contains the fields like Name, Username, Password, Confirm Password, Email-id and Mobile Number. In this user registration module all the fields must be filled otherwise an error message will display. User must register with the cloud then perform the remaining operation without registration can't perform the other operations so initially user register then go for the login. In the user registration form, user must enter the valid information otherwise user got the error message. If the user registered with the proper information or details then users can perform the remaining operations otherwise not.

B. Login

In this user login module, user must enter proper user name and password. User performs the further operations if any error in the user name and password error message should display to enter proper details.

C. File Upload

Once the user login successfully, now the user can create medical dataset with privacy fields and upload the created medical dataset into cloud. Once the user uploaded successfully display the message you have successfully uploaded the file.

From the user uploaded medical dataset we need to generate dataset in such a way that only the fields which are containing privacy fields to be hidden. Now apply the anonymize technique to the datasets. For the anonymized datasets are encrypted by using DES algorithm. This encrypted datasets are stored into the cloud.

D. File Download

The user can perform download the dataset file if the user selected option has a download file. Once the user successfully downloads the dataset file from cloud then the dataset file is decrypted using DES algorithm. This decrypted dataset file is stored into the local system.

E. Verification by the TPA

In this module is used to send the auditor request to the TPA. Generation proof is done by the cloud server to generate a proof of data storage correctness. Verification is done by the TPA to audit the proof form the cloud server.

F. User Authentication and Authorization Module

This module contains all the information about the authenticated user. User without his/her username and password can't enter into the login if he/she is only the authenticated user then he/she can enter to his/her login and he/she can see the all the information related to the project which he/she is developing. This module uses Form Based Authentication & Authorization to make security.

V. RESULTS

The proposed method provides secured health data access. In this work, anonymizationis done using l-diversity techniques, which add more security for private data present in the dataset. Then the whole dataset is encrypted using DES technique. Now the private data is double encrypted. The TPA is used for purpose of the audition in the server. Using RSA does not give better results than DES. So DES is considered as good encryption technique when compared to RSA. Even though AES is more efficient encryption technique, it consumes more time for computation. Therefore, DES is added with anonymization technique in order to achieve efficient security and also to reduce the computation time. Only related user can decrypt the data and view the anonymized data. In this work, private data get double encrypted from anonymization and DES. From this work, performance towards security is more when compared with other techniques. The anonymized data is not accessible to all. Only few users closely related the data can access this private data.

DOI 10.17148/IJARCCE.2015.4591

VI. CONCLUSION

The use of cloud computing is an important development in the world. Many IT businesses have started using cloud architecture because of its pay-as-you-go concept. Security is one most important factor that everyone thinks before uploading data's in the cloud. In this paper, we have proposed a combination of anonymization and encryption to enable securing of intermediate datasets. A privacy preserving searching algorithm is also proposed to identify which intermediate datasets are to be encrypted rather than encrypting all the intermediate datasets so as to enable cost effective security. Any dataset that is to be downloaded by the user, he should get verified by the TPA.

REFERENCES

1. C.K.Chu, S. M. Chow, W.G.Tzeng, J. Zhou andR. H. Deng, "Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage",Volume 25, Issue 2,pp. 1-11, 2014.
2. A. Sakhare and S. Ganar, "Anonymization: A Method to Protect Sensitive Data in Cloud", International Journal of Scientific & Engineering Research, Volume 25, Issue 2, pp. 1-11, 2013.
3. A.Omotosho,"A Criticism of the Current Security, Privacy and Accountability Issues in Electronic Health Records", International Journal of Applied Information Systems, Volume 7, Issue 8, pp. 11-18, 2014.
4. V. R.Patil and A.C.Lomte,"Challenges toward Achieving Privacy and Secure Searchable Outsource cloud data Storage Services",International Journal of Advanced Research in Computer Science and Software Engineering,Volume 3, Issue 11, pp. 400-403, 2013.
5. A.S.Anupriya, Ananthi and S.Karthik, "TPA Based Cloud Storage Security Techniques",International Journal of Advanced Research in Computer Engineering & Technology ,Volume 1, Issue 8, pp. 1-4, 2012.
6. K.P. Kulkarni and A.M.Dixit,"Privacy Preserving System Using Attribute Based Encryption for e-Health Cloud", International Journal of Science and Research, Volume 3 Issue 12, pp. 518-523, 2014.
7. Y. Zhu, G. J. Ahn, H. Hu, S. S. Yau,H. G. An, and C. J. Hu, "Dynamic Audit Services forOutsourced Storages in Clouds", IEEE Transactions on Services Computing, Volume 6, Issue 2, pp. 227-237, 2013.
8. R. Curtmola, J. Garay, S. Kamara and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions",
9. N. Cao," Secure and Reliable Data Outsourcing in Cloud Computing", 2012.
10. Y. Zhu, H. Wang, Z. Hu, G. J. Ahn, H. Hu and S. S. Yau, "Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds", Association for computing resources,2011.
11. J.Sun,X.Zhu,C.Zhang,andY.Fang,"HCPP:Cryptographybasedsecure EHR system for patient privacy and emergency healthcare," in Proc. IEEE Int. Conf. Distrib. Comput. Syst., Jun. 2011, pp. 373–382.
12. G. Ateniese, R. Curtmola, B. de Medeiros, and D. Davis, "Medical information privacy assurance: Cryptographic and system aspects," presented at the 3rd Conf. Security Commun. Netw., Amalfi, Italy, Sep. 2002.
13. Yue Tong, Jinyuan Sun, Sherman S. M. Chow, and Pan Li, "Cloud-Assisted Mobile-Access of Health Data With Privacy and Auditability", IEEE journal of biomedical and health informatics, vol. 18, no. 2, march 2014 419.
14. Abdullah Abdulrhman AlShwaier, Dr, Ahmed Zayed Emam, "Data Privacy On E-Health Care System", International Association of Scientific Innovation and Research.
15. D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing. Extended abstract in CRYPTO 2001," SIAM J. Comput., vol. 32, no. 3, pp. 586–615, 2003

DOI 10.17148/IJARCCE.2015.4591