ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

Machine Learning Algorithms for Secured Cloud Medical Data Storage Using Identity Based Encryption

[1] A.Hemlathadhevi, [2]Dr.Rajeshwari Mukesh
[1]Research Scholar, Department Of Computer Science and Engg, St. Peter's University, Tamil Nadu, India
[2] Professor, School of Computer Science and Engg, Hindustan University, Tamil Nadu, India

Abstract: Cloud computing will be the next major storage centre of the future for medical data. There are many powerful data centres that are usually heterogeneous and geographically distributed. Providing access to cloud computing data, however, is a big challenge. This article proposes a new method for securely storing and converting medical data into the cloud computing platform. Medical data are classified as sensitive or insensitive data by machine learning techniques and encrypted by encrypting security encryption method (IBE) for secure data access. Because the cloud's medical data is shared by multiple users, data protection is very important. In this way, the proposed system use of protection filtering mechanism user verification. Finally, compare the proposed method with other cloud-based data encryption methods with simulation and comprehensive analysis. The experimental results show that the solution proposed for cloud-based data transfer is very safe and feasible.

Keywords: Cloud Storage; Medical Data; Machine Learning Technique; Identity Based Encryption; Privacy Preserving Mechanism; Homomorphic Encryption.

I. INTRODUCTION

Cloud computing improves its quality in many medical applications. The exchange of medical data through the cloud platform has become very popular today. Transmission of online medical data facilitates a medical expert process that promotes storage, sharing, and access to medical data [1]. Cloud computing means that huge storage applications are handled by doctors from a remote location. In this way, the cloud platform received some attention from IT marketing. It is a replacement program where hospitals and other healthcare organizations to exchange patient records and also store medical records. Cloud computing technology in the medical record includes the advantages and disadvantages of data security as security and reliability are a problem.

Storing medical data on a cloud computing platform allows users to conveniently store their data. The user first encrypts personal information and sends them to the cloud storage. The cloud platform rearranges the encrypted data without knowing the original data. Encrypted cloud data is generally maintained with a third party [2]. Therefore, the integrity, confidentiality, and security of medical data become more and more uncertain than the private storage system. Therefore, it is important to make sure that unauthorized third parties can access or change their encrypted cloud data. Cloud platforms use very probable encryption techniques anywhere the third party is responsible for the cloud storage along with the secret key is generated by the third party. Yet, another problem is that an unauthorized user can search for search data, which results in data leaks for outsiders.

The document provides a secure data access scheme that provides greater security and integrity when storing and sharing confidential health data in the cloud. This proposed system contains four different methods in the Cloud Platform, which improves the security, privacy, and integrity of users' personal data. First, the medical records of the patient are classified as sensitive and insensitive data, which is possible from the vector support machine (SVM) from the machine learning algorithm. The second method is encryption (IBE), where access permission is used not only by the user but also by cloud data. The data controller will grant this permission to access the third-party data normally provided in the cloud. The third is homomorphic encryption, which is optional if the data owner changes the shared data on the cloud-based platform. Enhances the security of the cloud platform where the encoded data is modified without the original data being known and the authorized user retrieves and decrypts the data. Finally, a ring signature is created, a data protection mechanism that preserves the confidentiality of sensitive data through the user's confirmation process.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

The rest of the article is structured as follows. Section 2 describes related work with the proposed techniques. Shared sharing of cloud medical data is described in Chapter 3. Section 4 describes the results and discussion with the corresponding diagrams and finally, the completion of this report is presented in Chapter 5.

I. RELATED WORK

Qiu, Meikang, Keke Gay et al [3]. Allow permission to restrict access to data. Second, we offer a user-centered approach to proactively prevent the unexpected operation of users on the cloudy side. Finally, the proposed system is a certain resistance to a higher level because it cannot cope with dynamic threats, including emerging and future threats. We have seen that our proposed system has a quality performance that meets the expected target.

Li, Yibin, Keke Gai et al [4]. Proposed by the authoring system titled Model "SA-EDS", which primarily supports our algorithms, including algorithm for alternative distribution data (Ad2) algorithm for safe efficient allocation of data (SED2) and efficiently configured data for EDCon) algorithm. The experimental estimates are evaluated by the indicators of safety and efficiency, and the experimental results show that this approach effectively protects the main threat of cloud and requires an acceptable computation time.

Seenu Iropia and Vijayalakshmi et al [5]. The data controller has created sensitive data and data security, and access control has become a major problem here. Paper solves the problem of policies through access to data according to the data quality and allows the data owner to make the calculation access control without disclosing the content of the data. Therefore, the encryption key based on attributes (KP-ABE) is implemented in order to solve the problems mentioned in the article.

F. Ozgur CATAK and M. Erdal BALABAN et al [6]. The author of the mechanism is training Cloud SVM together with a technique in Map-Reduce cloud environment. The distributed cloud complies with the SVM algorithm to properly classify the data set. It then combines the support vector for each node in the cloud. We repeat these two steps until you find the optimal grader for the entire database cloud. Since it is difficult to find the optimal data set of huge data sets in the algorithm SVM system data files by dividing multiple repetitions and the results of the iteration combined and going for a few more reps to the optimal data set.

Dharani. R and M. Narmatha et al [7]. Computing cloud communication is a major problem with data security and access control. Three key services such as key data protection, key freshness, and key authentication are key to dynamic data sharing for a dynamically changing group. To achieve these three functions, the author offers the Dynamic Group key database protocol based on the Central Key Generation Centre. This method divides the keys to members of the group and rejects the distribution of members of non-members. In addition, anonymous access is governed by the short signing system.

Shashank Bajpai and Padmija Srivastava et al [8]. Encoding remotely stored data is a big problem in data sharing in the cloud environment. All this for the right reason that homomorphic encryption is a better approach that improves application security where sensitive data is found. Encrypted data are processed as an encryption input and changes the correct changes to the encrypted data without decrypting the data and the content of the data does not know the user during the changes. This homomorphic method of encryption demonstrates that remote sharing of personal data is made without prejudice to the privacy of personal data.

M. Saranya and R. Vasuki et al [9]. The key generator is usually used to decode all the messages belonging to the user. But if two users share the data, then the data must be kept confidential to third-party users except for the two users who shared the data. Therefore, the author offers a new encryption system that solves this problem by issuing a free access key that provides communication between the two parties data centre and the key generator. Security performance can be controlled using Third Party Control (TPA) and security analyzers using distributed partition data using RBAC.

Hassan Takabi et al [10]. Cloud providers limit privacy by CSP. There are many approaches to prevent CSP when user privacy is a threat to unreliable CSP. These existing approaches are burdened by the communication and the complexity of key management. This is the paper handling system that provides two levels of cloud data protection. CSP is an enhanced access control mechanism that protects user data, and the third service provider protects data with multiple encryption layers.

M. Divya Meena, AR. Arunachalam and T. Nalini et al [11]. There is a need for an algorithm for securely sharing personal data for multiple data users. Each node in the network provides a number from 1 to N. This is possible

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

with the authentication technique and thus make the data more secure than before. According to the identifiers, each node issued an identification number with the help of the central authority. This secure personal data sharing algorithm compares with existing algorithms and demonstrates that the proposed algorithm outperforms existing algorithms.

II. SECURED DATA SHARING ON CLOUD

A. Materials and Methods

Patient data consisting of clinical, demographic and general information from the database. About 6,000 patients were treated with chest and lung disease. These databases contain details about the patient, such as name, age, gender, family history of the disease, the results of the laboratory tests, physical examination by the physician, the results of the screening, etc. The severity of the disease changes in acute, chronic and chronic condition. The database subgroup (some attributes) is shown in Table 1 [12]. This table shows the state of the lungs based on the health criteria. These data have been continuously monitored since 2007. Most patients wanted to know their medical condition and receive a recipe from a remote location. So there is a need to store data on the cloud platform. As the patient's data is sensitive and insensitive, security and privacy are a challenge. Identification encryption is used with homomorphic encryption to overcome the security problem and the ring signature scheme is used to overcome the security.

TABLE.1 DATABASE SAMPLE

Gender	
Male	Idiopathic Pulmonary Fibrosis, Pneumoconiosis
Female	Lymphoid myomatomatosis, associated with ILD with connective tissue disease
Age	
20-40 years	An ILD related tissue disease, lymphangiolithiasis, sarcoidosis
Above 50 years	Cryptogenic organizing pneumonia, idiopathic pulmonary fibrosis.
Family history	The most important symptoms of the gene
Physical Examination	Clubs, lung symptoms, lung tests, such as wheezing or cracks
Pulmonary Function Testing	Lung Volumes, Resting and Ambulatory Oxygen Saturation, Arterial Blood Gas
Biopsy	Lung volume, resting and ambulatory oxygen saturation, arterial blood gas
Laboratory Testing	The hypersensitivity pneumonia panel, routine blood test, and others.
Imaging	CT High Definition, Chest X-ray (X-ray)

B. Data Classification

The dataset consists of general and private data for each patient. Prior to the encryption and sharing process, sensitive data (illness, the severity of the illness, etc.) and sensitive data (name, address, etc.) and the separation of the confidentiality of sensitive data are important [13]. Therefore, the most effective maintenance vector (SVM) is used to classify the sensitive and insensitive data of patient medical records. Think about medical record elements (name, age, illness, weight, stage, status, level, etc.) as Elements. Let xi be the set and elements i = name, age, disease, weight, phase, state, level, etc. our system defines two classes that are sensitive and insensitive. An individual input of elements, based on sensitivity, assigning the elements. If weight 1, sensitive data, such as illness, weight, etc. It belongs if weight is -1, this item belongs to non-insensitive data such as name, address, age, etc.

Let y be the two classes, and classification is represented as

$$y_i = \begin{cases} 1 \; if \; x_i \; insensitive \; class \\ -1 \; if \; x_i \; in \; insensitive \; class \end{cases}$$

C. Secure Transaction

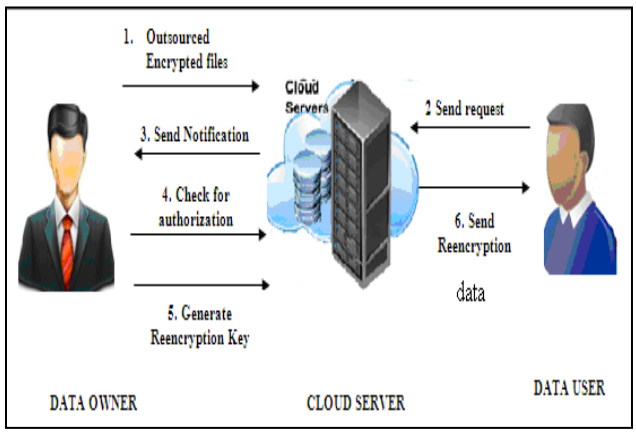
Classified data should be encoded preceding to storage inside the cloud. This is how the most important encryption technique called "Identity Coding" [14] is used. The IBE procedure is shown in Figure 1.

Initially, the data owner encrypts health records and send it to the cloud server. If the patient wants to know his medical records from the cloud server, the data owner will request a new encryption key (secret key). The data holder accepts the request and acknowledges the data user license that sends the request. If the user is an authorized then it creates the re-encryption key and sends it to the suitable user. After reaching the key encryption, the user can access the medical reports. The main benefit of identity-based encryption (IBE) is that you can give information of the special patients as a requesting. Data from other patients are not given without proper permission.

FIGURE.1 SECURED CLOUD DATA STORAGE USING IDENTITY BASED ENCRYPTION

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018



In our proposed system, There is only one G Group in our system, which consists of medical experts as data users and patients as data users. Let's see how identity encryption (IBE) works in our system. Like the standard IBEs, the proposed process consists of four steps: Create Names, Create Keys, Encrypt and Decrypt. The four steps in the proposed system are described as follows:

Setting: The public key generator (PKG) arbitrarily selects public T parameters such as k ← (L (T)) / (DS (T)), where k is a random pseudo square, is the set of elements, Jacobi 1 modulo T and T = ab, where the main secret key a and b.

Key Generation: A general identifier is created by the PKG and calculates the hash function for each data block. The secret key for every data block is calculated by key generation. The host of the selected global identifier ID is calculated as a function

$$S = H\,(ID) \qquad (1)$$

And the secret key for each block of the data is
Where

$$t = \begin{cases} \sqrt{S} \text{ if S is a square} \\ \sqrt{kS} \text{ if S is not a square} \end{cases}$$

Encryption: Creates a secret key for each block of data, and the data is encrypted by the generated secret key. For example, consider the secret key for block x and can be calculated as follows:

$$d_x = (f_x^2 + k^x S)/f_x \qquad (2)$$

With the encryption of the data block is completed by

$$C = \text{Encrypt}\,(T, ID_M, M) \qquad (3)$$

Where M is the original message, T is the public parameters and is the whole message ID

Encryption: Encrypted data stored in the cloud and data protection settings is applied to the cloud-based encrypted message based on the following session. These privacy settings are used to allow some users to read the file. This clearly states that the patient's medical records are not visible to the other patient, although they belong to the same group. Decrypt the encrypted message with public T parameters, C encrypted text, whole message and most of all with the secret key of the file (medical report of that patient).

$$Decrypt(T, C, ID_M) = M \qquad (4)$$

The patient's medical records are not durable, but changes have taken place depending on the patient's state of health. After the data has been encrypted and stored in the cloud, future changes are traditionally updated by decrypting, modifying, encrypting data, and storing in the cloud. However, for IBE, a homomorphic encryption is used besides safe cloud access. This method encodes itself without decoding and without knowing the private key of the data. The secret key was only the data owner knew.

D. Privacy-Preserving

The concept of secrecy refers to the preservation of the identity of the individual in the group. Consider whether the set of medical data stored in the cloud contains many patient data. Encrypted data is divided into blocks and each block is signed by one of the group member (user). Checking the licensee's license at IBE Phase, the ring signature is used to check the signature, regardless of whether it belongs to that group. Does not specify the particular user. By signing the signature and using known authorized users, the examiner cannot determine the identity of the signatory.

Signature to the ring consists of three basic steps: generating keys, creating ring signatures and ring checking [15]. In the key generation phase, all users in the group create their own private and public keys.

The u_i user is randomly selected by private key $x_i \overset{R}{\leftarrow} Z_p$ and calculate the public key using the private key such as $w\_i = w_i = g_2{}^{x_i}$.

During the ring signing phase, each user in the group generates a signature on a block and generates a block identifier for the public keys of all group members and the private key for that particular user. This block identifier allows both blocks to be distinguished from each other. The ring signature should be calculated as follows:

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

$$\sigma_s = \left(\frac{\beta}{\varphi(\prod_{i \neq s} w_i^{a_i})}\right)^{1/x_s} \in G_1 \quad (5)$$

$$\text{Where} R_s = H_1(\text{Id}) g1^m \in G_1 \quad (6)$$

Where H_1 hash function, it is blocked identifier, G_1 group.

The last step is to confirm that the ring is where the verifier verifies whether the given block is signed by members of the group. Calculate the number of d, private and public keys, identifiers, block m and σ_s ring signals, and then check equation (7).

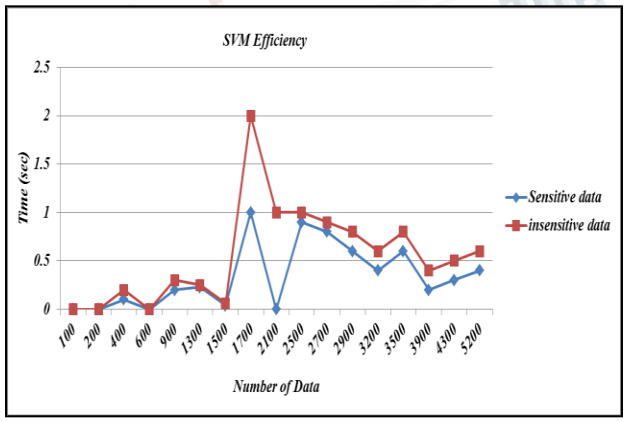
$$e(\beta, g_2) = \prod_{i=1}^{d} e(\sigma_i, w_i) \quad (7)$$

If equation (7) satisfies, members of the group sign a sentence or otherwise have access to unauthorized data. The main advantage of signing the privacy signature is that it does not require much room for the ringtone to be stored and dynamic operations are updated to the blocks without interfering with the secrecy process.

III. RESULT ANALYSIS

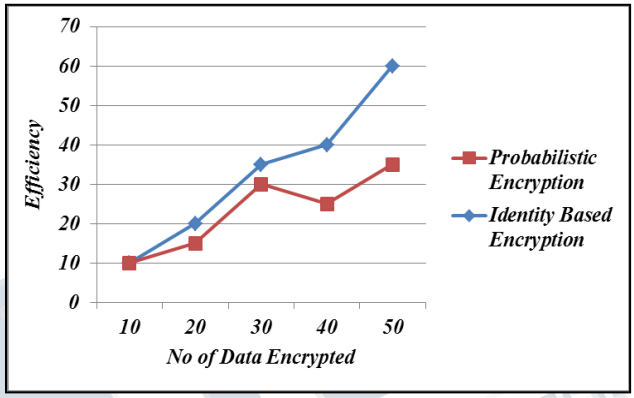
The important techniques used in the proposed system will be evaluated in this section. Virtual Machine Maintenance (SVM) is used to classify sensitive and insensitive data. Figure 2 shows that the SVM classifier provides consistent performance. Due to the experimental results, although the database provides less sensitive data than sensitive data, SVM keeps constant accuracy, as shown in Figure 2. This figure also shows that SVM provides better performance even if the number of data needs to be increased.

FIGURE.2 SVM CLASSIFIER EFFICIENCY



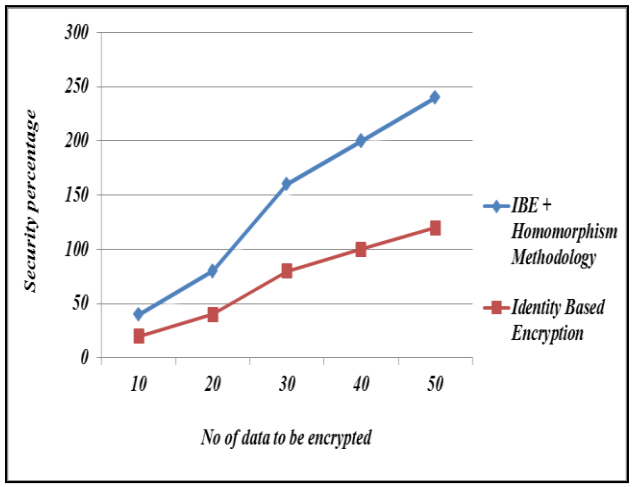
The Identity Based Encryption (IBE) is compared with the previous method of probabilistic encryption using the proposed method. For probabilistic encryption, plain text or partial text can be easily calculated from the encryption text. Therefore, IBE performs improved than existing Probabilistic encryption. Figure 3 shows the comparative graph for both probe encryption and personal encryption (IBE), indicating that the IBE shows better performance than the existing encryption method.

FIGURE.3 COMPARISON OF IBE WITH EXISTING PROBABILISTIC ENCRYPTION



The IBE has more coding performance than previous encryption methods. However, modification of data after encryption takes a lot of time for decoding, updating data, and re-encoding. Thus homomorphic encryption is provided with the IBE, which reduces the time needed to modify the data. The efficiency of IBE and IBE integrated with homomorphic encryption is shown in Figure 4.

FIGURE.4 EFFICIENCY OF IBE WITH THE HOMOMORPHIC ENCRYPTION



IV. CONCLUSION

The paper proposed a secure access to cloud health data using the most elegant, identity-based encryption that encrypts the data based on the user's identity. Medical data is classified as sensitive and insensitive data, thus facilitating security. Modifying data after cloud storage is done with homomorphic encryption without decryption. The confidentiality of medical data must be retained by the proposed ring signature, which is only part of the data collection team. Finally, the proposed system compares the existing probability of cloud encryption and indicates it is better than the existing method. The effectiveness of identity encryption is demonstrated with homomorphic encryption and the performance of the SVM classifier.

REFERENCES

[1]. Yu, Yong, Man Ho Au, Giuseppe Ateniese, Xinyi Huang, Willy Susilo, Yuanshun Dai, and Geyong Min. "Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage." IEEE Transactions on Information Forensics and Security 12, no. 4 (2017): 767-778.

[2]. Abhishek Kumar Gupta and Kulvinder Singh Mann, "Sharing of Medical Information on Cloud Platform-A Review", IOSR Journal of Computer Engineering (IOSR-JCE), Volume 16, Issue 2, and ISSN: 2278-0661, 2014.

[3]. Qiu, Meikang, Keke Gai, Bhavani Thuraisingham, Lixin Tao, and Hui Zhao. "Proactive user-centric secure data scheme using attribute-based semantic access controls for mobile clouds in the financial industry." Future Generation Computer Systems 80 (2018): 421-429.

[4]. Li, Yibin, Keke Gai, Longfei Qiu, Meikang Qiu, and Hui Zhao. "Intelligent cryptography approach for secure distributed big data storage in cloud computing." Information Sciences 387 (2017): 103-115.

[5]. Seenu Iropia and Vijayalakshmi, "Decentralized Access Control of Data Stored in Cloud Using Key-Policy Attribute-Based Encryption", International Journal of Inventions in Computer Science and Engineering, Volume 1, Issue 2, ISSN (Online): 2348 – 3539, ISSN (Print): 2348 – 3431, 2014.

[6]. F. Ozgur CATAK and M. Erdal BALABAN, "Cloud SVM: Training an SVM Classifier in Cloud Computing Systems", Springer Berlin Heidelberg, Volume-7719, ISSN: 0302-9743, pp.57-68, 2013.

[7]. Dharani. R and M. Narmatha, "Secured Data Sharing with Traceability in Cloud Environment", International Journal of Inventions in Computer Science and Engineering Volume 1, Issue 8, ISSN (Online): 2348 – 3539, ISSN (Print): 2348 – 3431, September 2014.

[8]. Shashank Bajpai and Padmija Srivastava, "A Fully Homomorphic Encryption Implementation on Cloud Computing", International Journal of Information & Computation Technology, Volume 4, Issue 8, pp. 811-816, ISSN 0974-2239, 2014.

[9]. M. Saranya and R. Vasuki, "Improving Data Security in KP-ABE with Third Party Auditing", International Journal of Inventions in Computer Science and Engineering, Volume 2, Issue 2, ISSN (Online): 2348 – 3539, ISSN (Print): 2348 – 3431, Feb 2015.

[10]. Hassan Takabi, "Privacy-Aware Access Control for Data Sharing in Cloud Computing Environments", proceeding of the second International Workshop on security in Cloud Computing, pp.27-34, ISBN: 4503-2805, 2014.

[11]. M.Divya Meena, AR. Arunachalam and T. Nalini, "Confidential Data Sharing With Anonymous Id Assignment Using Central Authority", Volume I Issue 2, International Journal of Inventions in Computer Science and Engineering, ISSN (Online): 2348 – 3539, ISSN (Print): 2348 – 3431, 2014.

[12]. Gulati, Mridu. "Diagnostic assessment of patients with interstitial lung disease." Primary Care Respiratory Journal 20, no. 2 (2011): 120.

[13]. Afif, M.H, and Hedar, A.-R., "Data Classification using Support Vector Machine Integrated with Scatter Search Method", Japan-Egypt Conference on Electronics, Communications and Computers (JEC-ECC), IEEE, pp. 168-172, ISBN: 4673-0485, 2012.

[14]. Varsha S. Agme and Archana C. Lomte, "Cloud Data Storage Security Enhancement Using Identity Based Encryption", International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 3, Issue 4, ISSN 2319 – 4847, April 2014.

[15]. Boyang Wang, Baochun Li, Hui Li, "Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud", IEEE Transactions on Cloud Computing, Volume 2, Issue 1, January-March 2014.