INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 3)
Available online at: www.ijariit.com

A survey paper on privacy of medical data storage on the cloud

Roomana Hasan
roomana.hasan@ymail.com
D. Y. Patil College of Engineering, Pune, Maharashtra

Dr. Kailash Shaw
kailash.shaw@gmail.com
D. Y. Patil College of Engineering, Pune, Maharashtra

ABSTRACT

The proliferation of smartphones and use of reliable cloud technologies have given rise to various cloud-based applications. Remote disease prediction based on real-time medical data is one such application that has become very popular among research communities. One of the challenges involved here is the privacy of medical data storage on the cloud. Encrypted data requires decryption before it can be used by machine learning algorithms for disease prediction. This paper presents an analysis of research methods proposed to provide privacy of medical data on the cloud and their limitations.

Keywords: *Privacy, Decryption, Encryption, Public key, Private key, Cloud, Medical data*

1. INTRODUCTION

Progression in individual wearable medicinal gadgets [1] and cloud innovation has made it conceivable to gather colossal measure of therapeutic data, process them and enhance the precision of disease prediction. This increasingly high interests in making individual wellbeing information accessible on the cloud offer ascend to cloud-based wellbeing consultation [2]-[4]. Lamentably, for any such cloud-based wellbeing data framework, protection of patient's data is the greatest concern [5] [6]. Privacy of medical data can be achieved by applying cryptography. Cryptography is not only the method but also the research of techniques that allow us to make secure communication possible even also in the presence of third parties called adversaries. Modern cryptography presents in almost every discipline like mathematics, computer science and electrical engineering [7]. Cryptography involves two methods called encryption and decryption. Encryption changes the plain text to cipher text using encryption algorithms such that no one other than the sender can make sense out of it using a key generated by the algorithm during the encryption process.

1.1 Cryptography concepts

The concepts of cryptography mostly used are summarized below:

1) **Plain Text**-The original message is called plain Text.
2) **Cipher Text**-The unreadable text after encryption is called ciphertext.
3) **Encryption**- Plain text converted into Cipher (non-readable) text. The process is called encryption. The encryption algorithm is used for this process.
4) **Decryption**-Conversion of ciphertext. In to plain text is called decryption.
5) **Key**-It is an Alpha-numeric text (mathematical formula) or numeric. Encryption and decryption take place with help of a key.
6) **Key size**-It is the key length in bits.

1.2 Cryptography types

There are two types of cryptography. They are:

1) **Symmetric Key Cryptography-** Where the same key is used for encryption and decryption.
2) **Asymmetric Key Cryptography-** Where two different keys are used, one for encryption and other for decryption.

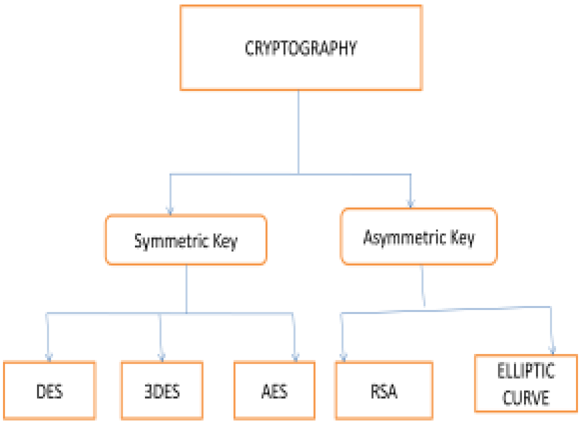

Fig. 1: Cryptography classification

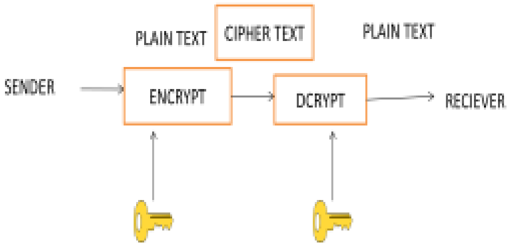

Fig. 2: Symmetric Cryptography

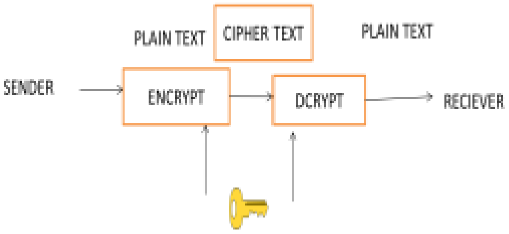

Fig. 3: Asymmetric Cryptography

There are three popular symmetric key algorithms:

a. **DES -**Data Standard Encryption is a block Cipher and used to decrypt and encrypt data blocks consist of 64 bit by a 64-bit key by means of substitution and permutation. A plaintext, DES groups the plaintext into 64-bit blocks. Decryption is the reverse of Encryption. It is designed by IBM in the year 1974. It is published by NIST (National Institute of Standards and Technology).

   Strength- the 2^56 possibilities of keys and it takes a decade to find out the correct key using brute-force attack.
   Weakness-It is vulnerable to linear cryptanalysis attack and moreover, it is weak key. It is exposed to bruteforce attack.

b. **Triple-DES-** It is an advanced version of DES. It designed in1978 and have higher security to encrypt and decrypt the data by applying the algorithm three times in succession with three different keys. 3DES is used in US government systems. It uses the plain text of 64 bits with 48 rounds and168-bits key length.

   Strength- It is an advanced version of DES. It designed to have higher security than DES.
   Weakness- It is exposed to differential and related-key attacks.

c. **AES-** Advanced Encryption Standard (1998) consists of three block cipher: AES-128, AES-192, and AES-256.Cipher decrypts or encrypts data in blocks of 128 bits using keys of 128-bits, 192-bits, 256-bits. There 14 rounds for 256- bit keys, There 12 rounds for 192- bit keys, There 10 rounds for 128- bit keys, where the round is processing steps which include transposition and substitution and mixing of the plaintext and convert into the final ciphertext output. The steps for performing operations by using AES algorithm has following types of operations. They are
   - Sub Bytes
   - Shift Rows
   - Mix Columns
   - XOR Round Key

   Strength- It is six times faster than triple DES. It is implemented in hardware and software and more robust against hacking.
   Weakness- It uses simple algebraic structure. Every block is encrypted in the same way.

Two popular asymmetric cryptographic algorithms are mentioned below:

a. **RSA-**It was designed in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman as a Public-key in Massachusetts Institute of Technology for cryptography. Random prime numbers p and q be taken which should be large, distinct and have similar bit length.

   Strength- RSA algorithm is more secure for users as it uses complex mathematics. RSA algorithm is not easy to hack as it uses prime numbers whose factorization is difficult to make, moreover it is a public key.
   Weakness-RSA algorithm is very slow when a huge amount of data is encrypted.

b. **Paillier Cryptography**- In 1999, after Pascal Paillier the cryptosystem invented is named as Paillier which is used for public key cryptography with the probabilistic asymmetric algorithm. There is computational difficulty present for computing n-th residue classes. The intractability of hypothesis with this cryptosystem which is based on the decisional composite residuosity assumption.

   Strength- It is an efficient and additively homomorphic cryptosystem.
   Weakness-Pallier key size is long and the speed of execution is low.

c. **ECC (Elliptic Curve Cryptography)**-It was created in 1985 by V. Miller as another form of Asymmetric cryptography. The equation of the elliptic curve is used in encryption. Plane curve over a finite field is an Elliptic curve.

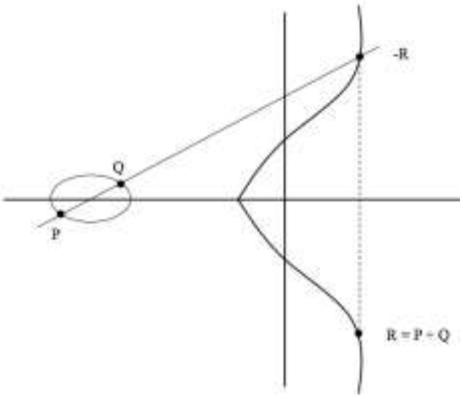

Fig. 4: Elliptic Curve Representation

Q and P, on an elliptic curve, find the integer n,

$$P = nQ \tag{1}$$

An elliptic curve consists of (x, y) is the set of real numbers of an elliptic curve and it satisfies the equation:

$$Y^2=X^3+AX+B \tag{2}$$

Curve shape deep depends on a and b. As a and b increases curve also increases. Hacker can know Q and P. Problem to find n. Where $Q = nP$ (public key) but n (private key).

Strength-Encryption and decryption are faster than any other public cryptography because it has smaller keys than any asymmetric keys. It requires less power and gives maximum security with 164 keys than any other cryptosystem keys.
Weakness-Text of encryption size is increased and ECC depends upon complex equation.

2. LITERATURE SURVEY

Our work is inspired by several research works published in this field that presented how healthcare system can be built on the cloud and how to send data to the cloud. As early as in the year 2004, Hung, Zhang, and Tai demonstrated a tele-home system which utilized wearable devices, wireless sensor technologies and multi-sensor data fusion methods to collect patient health information [1]. In [3] [4], monitoring and consultation system using smartphone device integrated with a medical device has been presented.

Use of personal handhelds and IoT devices for remote consultation brought the need for privacy and robust data security. To address privacy challenges associated with telehealth system, researchers proposed various schemes of data collection on the cloud. For example, authors in [5] [6] presents challenges associated with solving online privacy problem.

Other challenges associated with storing medical information on the cloud is related to making the information useful and pervasive with heterogeneous types of the system used by healthcare organizations for storing patient data. To address this problem, a standard for storing and retrieving medical data is developed called HL7. In [11], the author shows transmission of clinical information using HL7 CDA standard.

Since the use of cloud-based storage is excessively proposed by researchers, many research works conducted around the evaluation of various cloud-based solution available online and types of privacy and security issues associated with it. In [14], the author shows the limitation of most popular cloud storage Dropbox. In paper [15], researcher presents a hybrid solution for privacy-preserving medical data sharing in the cloud environment.

Hossain and Muhammad present technical IOT enabled scheme for medical analysis in [20] [21] where ECG & other healthcare information has been taken by sensors & mobile devices are ensured forward the cloud for logical access by a medical specialist. They validated their approach through practical analysis and replica by building IOT based ECG health monitoring service.

In [22], author analyses the problem of files sharing and updating in a multi-user environment. A new concept of duplicable dynamic storage and an efficient construction has been introduced.

In [23], Bethencourt, Sahai, and Waters introduce text-policy based encryption and decryption to ensure the privacy of the encryptor. By using this approach the receiver gets access policy in the form of a tree. M. Chase has proposed multi-authority attribute [24] based encryption methodology to encrypt the message. The recipient can decrypt message only if he has at least number of the given attributes from each authority.

The fundamental concept of the techniques was exploiting a large bit rate data hiding methods to insert patient data within real medical Information. In these techniques, the patient data was private and hidden during electronic transmission. This technique is robust to assault thus as, intrusion, compression, cryptanalysis, and cropping [25]. In [26], the author describes a systematic review of healthcare systems that use privacy protection in the cloud.

In the model [27], patient data is not necessarily surely shown in the form of the medical information. Alternatively, the patient information is firstly encoded by using the ASCII character encoding technique and this information encrypted into a non-detected scheme by using the Paillier encryption technique. The other patient information is more secure by building it in a section of the medical information that is external to the ROI (region of interest). Securely that the encoded and encrypted data is built in a position that will not impact the Medical Data quality. The region external to the ROI is position by using the Medical Data segmentation method.

Literature survey suggests that Paillier algorithm has certain drawbacks compared to Elliptic Curve Cryptography (ECC). Some of these are listed below:

- Key size used by the Paillier algorithm is very long
- Paillier encryption and decryption is slow
- ECC is more robust compared to Paillier

3. CLOUD STORAGE PROBLEM

Cloud-based data sharing give rise to following problems:

- How to protect the security of user's personal medical data during its delivery to the cloud?
- How to ensure that medical data stored on the cloud doesn't create privacy problem?
- How to secure the healthcare data stored in a remote cloud?
- How to efficiently retrieve stored medical data for learning and disease prediction?

4. ADVANCE CRYPTOGRAPHY

Two popular public key cryptographies are Paillier and Elliptic Curve. Paillier cryptography calculation had been the establishment of PKC (Public Key Cryptography). ECC, in any case, is developing as a substitution in a few conditions since it gives comparative levels of security contrasted with Paillier cryptography calculation yet with fundamentally diminished key sizes. In National Institute of Standards and Technology (NIST), the exhibition the key size connection amongst ECC and Paillier cryptography calculation is finished by utilizing Table I. Sizes of the key is in bits.

Table 1: ECC and Paillier key comparison [23]

Key Size of ECC	Key Size of Paillier	Key Size Ratio
163	1048	1:7
256	3036	1:12
384	7672	1:20
512	15376	1:30

Because of Paillier key sizes are long and multiple times than ECC key sizes, the length of the general population key and private key is considerably shorter in elliptic bend cryptosystems which gives us quicker handling of circumstances, and lower requests on memory and transfer speed; from some creation it is seen that that ECC is speedier than Paillier for marking and unscrambling, yet slower for signature confirmation and encryption.

5. CONCLUSION

In this paper, review of various works done around privacy of medical data and disease prediction has been presented. We studied various systems proposed in the past and found that the privacy of the medical data can be best ensured with the help of ECC algorithm.

As a future work, we aim to develop a system and analyze the performance of medical data storage and disease prediction with the help of advanced encryption system and machine learning algorithm.

6. REFERENCES

[1] Hung, Y. Zhang, and B. Tai, "Wearable medical devices for telephone healthcare", in Engineering in Medicine and Biology Society, 2004.IEMBS04. 26th Annual International Conference of the IEEE, vol. 2.IEEE, 2004, pp. 53845387.

[2] https://www.patientslikeme.com/.

[3] Gitarja Sandi, I Gusti Bagus Baskara Nugraha, Suhono Harso Supangkat, "Mobile health monitoring and consultation to support hypertension treatment", ICT for Smart Society (ICISS) 2013 International Conference on, pp. 1-5, 2013.

[4] D.P.I.J. Simarmata, "Design and Implementation of Mobile Health Consultation between Doctors and Diabetes Patient", International Conference on Cloud Computing and Social Networking (ICCCSN), pp. 1-4, 2012.

[5] C. Zhang, J. Sun, X. Zhu, and Y. Fang, "Privacy and security for online social networks: challenges and opportunities," Network, IEEE, vol. 24, no. 4, pp. 13–18, 2010.

[6] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 1, pp. 222–233, 2014.

[7] Effy Raja Naru, Hemraj Saini, and Mukesh Sharma, "A recent review on lightweight cryptography in IoT", I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), pp. 887-890, 2017

[8] Kuntinan Plathong, Boonprasert Surakratanasakul, "A study of integration Internet of Things with health level 7 protocol for real-time healthcare monitoring by using cloud computing", In Biomedical Engineering International Conference (BMEiCON), pp. 1-4, 2017

[9] J.-J. Yang, J. Li, J. Mulder, Y. Wang, S. Chen, H. Wu, Q. Wang, and H. Pan, "Emerging information technologies for enhanced healthcare", Computers in Industry, vol. 69, pp. 311, 2015.

[10] L. Griffin and E. De Leastar, Social networking healthcare, in Wearable Micro and Nano Technologies for Personalized Health (pHealth), 2009 6th International Workshop on. IEEE, 2009, pp. 7578.

[11] Yiming Yang, Liang Xiao and Jingbai Tian, "Transmission of clinical information based on HL7 CDA standard", In 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 949 - 952, 2016

[12] Dropbox cloud storage, Online, https://www.dropbox.com

[13] Luca Caviglione, Maciej Podolski, Wojciech Mazurczyk, and Massimo Ianigro, "Covert Channels in Personal Cloud Storage Services: The Case of Dropbox", In IEEE Transactions on Industrial Informatics, Volume: 13, Issue: 4, pp. 1921-1931, 2017.

[14] S. Li, Q. Zhang, Z. Yang, Y. Dai, "Understanding and Surpassing Dropbox: Efficient Incremental Synchronization in Cloud Storage Services", Proceedings of IEEE Globecom, 2015.

[15] J.-J. Yang, J.-Q.Li, and Y. Niu, A hybrid solution for privacy-preserving medical data sharing in the cloud environment, Future Generation Computer Systems, vol. 43, pp. 7486, 2015.

[16] L. M. Kaufman, "Data Security in the World of Cloud Computing," Security & Privacy, IEEE, vol. 7, pp. 61-64, 2009.

[17] V., Chang, & M. Ramachandran, "Towards Data Security with the Cloud Computing Adoption Framework", IEEE Transactions on Services Computing, Volume: 9, Issue: 1, pp. 138-151, 2016.

[18] M. Nassar, A. Erradi and Q. M. Malluhi, "Paillier's encryption: Implementation and cloud applications," 2015 International Conference on Applied Research in Computer Science and Engineering (ICAR), Beirut, pp. 1-5, 2015.

[19] M. S. Hossain, "Cloud-supported cyber-physical localization framework for patients monitoring", In IEEE Systems Journal, Vol. 11, Issue: 1, pp. 118-127, 2015.

[20] M. S. Hossain and G. Muhammad, Cloud-assisted industrial internet of things (iiot) enabled a framework for health monitoring, Computer Networks, vol. 101, pp. 192202, 2016.

[21] Kun He Jing Chen Ruiying Du Qianhong Wu GuoliangXue Xiang Zhang "DeyPoS: Deduplicatable Dynamic Proof of Storage for Multi-User Environments", IEEE Transactions on Computers, Volume: 65, Issue: 12, Dec. 1, 2016.

[22] J. Bethencourt A. SahaiB.Waters "Ciphertext-policy attribute-based encryption" Proc. IEEE Security and Privacy (SP' x2018; 07) pp. 321-334 2007.

[23] M. Chase, "Multi-Authority Attribute-Based Encryption", In Proceeding TCC'07 Proceedings of the 4th conference on Theory of cryptography Pages 515-534.

[24] T. Xu, W. Xiang, Q. Guo, and L. Mo, Mining cloud 3d video data for interactive video services, Mobile Networks and Applications, vol. 20, no. 3, pp. 320327, 2015.

[25] A. Sajid and H. Abbas, Data privacy in cloud-assisted healthcare systems: State of the art and future challenges, Journal of Medical Systems, vol. 40, no. 6, pp. 116, 2016.

[26] M. Jamali S. Samavi N. Karimi S.M.R. Soroushmehr K. Ward K. Najarian "Robust Watermarking in Non-ROI of Medical Images", 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society pp. 1200-1203 Aug. 2016.

[27] Heart disease data set http://archive.ics.uci.edu/ml/datasets/heart+Disease

[28] Christophe Petit, Michiel Kosters, Ange Messenger, "Algebraic approaches for the Elliptic Curve Discrete Logarithm Problem over Prime Fields", https://troll.iis.sinica.edu.tw/pkc16/slides/6-1--Algebraic_approaches_for_the_Elliptic_Curve_Discrete_Logarithm_Problem_over_Prime_Fields.pdf

[29] C. Zhang, J. Sun, X. Zhu, and Y. Fang, Privacy, and security for online social networks: challenges and opportunities, Network, IEEE, vol. 24, no. 4, pp. 1318, 2010.