ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

Data Privacy Preservation through CryptMDB

Prof. Balaji Bodkhe[1], Devangi Raval[2], Dipali Vanjari[3], Jyoti Shirsat[4], Kajal Patil[5]

Professor, Dept. of Computer Engineering, Modern Education Society's College of Engineering, Pune, Maharashtra, India[1]

Student, Dept. of Computer Engineering, Modern Education Society's College of Engineering, Pune, Maharashtra, India[2,3,4,5]

ABSTRACT:. Medical services applications are promising fields of remote devices where patients may be checked remotely. In comparison with wired network, Wireless networks are more unguarded to intruders, alteration, imitation and replaying attacks. It is easy to protect the patients data under transmission and there are many more solutions available to do so, but may not guarantee inside assault wherever the supervisor of the patient information uncovers the patients' credential data. There may be possibility that admin of system can have malicious intent so it becomes essential to consider this scenario in order to maintain privacy and confidentiality of patients data. In this paper, we are going to propose a system in which data is equally and securely distributed among more than one data servers and perform statistical analysis on that distributed data without compromising the patients' privacy. Along with privacy data access speed is another important issue. To overcome this issue we used CryptMDB features which provide security along with fast access.

KEYWORDS: Data Privacy, Paillier Encryption, CryptMDB, Sharemind .

I. INTRODUCTION

Medical services applications are promising fields for remote devices system, wherever patients may be checked through utilizing restorative device systems (WMSNs). Most of the existing systems tend to Protect data under transmission but there is no certain measures to Protect patients data from intruders. There are many security issues in current system such as data stealing, stealing and modification, alter contents with the wrong values. Suppose if the intruder with malicious intent is trying to get access the student information, there are lot many chances for the misuse of data which may lead to severe consequences, like modification in. The patients' data due to lack of security is also possible. The treatment prescribed by the doctors can be hacked which may even lead to death of the patients. Patients are the victims because of the above issues. To prevent these issues, Data privacy preservation through CryptMDB system propose a some new data protocols, where the patient's data is divided into three equal parts by using hash function and distribute them to three database servers. To preserve privacy of patient's data new protocol is proposed on the basis of Paillier cryptosystem which allow doctor to access the patients' data without revealing it to any server, hence it maintains privacy of patient's data. We also use new protocol to preserve privacy of patient's data while performing statistical analysis, without revealing patients sensitive data.

II. LITERATURE SURVEY

A) *Privacy Protection for Wireless Medical Sensor Data(2015)[3] Authors: Xun Yi, Athman Bouguettaya, Dimitrios Georgakopoulos, Andy song and Jan Willemson*

Medical healthcare monitoring architecture is a combination of collection of sensors and sensor network in order to monitor decrepit or chronic patients in their home. The wireless sensor system consist various medical

DOI: 10.15680/IJIRCCE.2019. 0705065

sensors that collects and transmits physiological health indicators to mobile computing devices. This study also presents a prototype for capturing sensor data from wireless sensor nodes.

B) *Understanding Privacy Violations in Big Data Systems(2018)[2] Authors: Jawwad A. Shamsi and Muhammad Ali Khojaye.*

Big data is conductive in analyzing computational issues for predictive concept. In spite of this, they show major interest for parasite privacy. Big data have helped in managing and processing many big-data science issues. Similarly many organizations are using big data for better accuracy of data. This analysis can be done to find many problems and to solve such issues. They do analysis on governments data, any co-orporate data this can be used for identifying attacks. All this attacks can be observed by anonymisation techniques for better privacy of data.

C) *CryptMDB: A Practical Encrypted MongoDB over Big Data(2017)[1].Authors: Guowen Xu, Dongxiao Liu, Hongwei Li, Kan Yang.*

Here we suggest an effective encrypted MDB to obtain users private data to be stored securely by encryption techniques. As CryptMDB is better relational database for computing and fast data access. CryptMDB protects confidential data whenever the query is entered. The biggest problem is whenever the private data is hacked its exibility decreases. So cryptDB provides us to encrypt the query whenever it gets entered into the database. And even provide the data into encrypted format but only to the authorized person. Because authorized person will have a private key from which it can decrypt that data. There is less chances of data loss.

D) *Efficient Paillier crypto processor for privacy-preserving data mining(2016)[7]. Authors: Ismail San, Nuray At, Ibrahim Yakut and Huseyin Polat.*

Provides solution to the performance problem through hardware oriented solution. It gains the insight for solving some computational challenges faced in such kind of applications. Paillier is mostly used as homomorphic encryption which gives surety of data privacy requirements by many data privacy schemes. Paillier Cryptosystem uses paillier crypto processor which uses various cryptographic algorithm. Cryptography is a method which converts the plaintext data into encrypted data. It generates key which is only given to the sender and receiver, so that they can only access the data. For privacy preserving we use algorithms. This algorithm helps to secure users private data. Data mining helps to divide user's large volume of data that split into particular format so they are easily understandable. Pipelining concept is used for parallelism i.e. parallelism is used so that user can do work parallelly and it also saves system time. Thus, this system becomes time efficient.

E) Accelerate the Paillier Cryptosystem in CryptDB by Chinese Remainder Theorem(2018)[6]. Authors: Yau Liu, Shuai Xue.

Chinese Remainder Theorem(CRT) can be use as performance booster for encryption process. like improvement in the system performance under certain condition and also improve CyptMDB performance. Today there are very few methods to increase data encryption or data decryption for this only we use CRT method to improve the performance of CryptDB. Chinese Remainder Theorem(CRT) is nothing but the theorem which knows the remainder of any division of number n in advance, by several number than one can identify uniquely the remainder of division by number n.

F) *CryptDB: Protecting Confidentiality with Encrypted Query Processing(2016)[5]. Authors: Raluca Ada Popa, Catherine M.S. Redfield.*

CryptMDB efficiently run query over encrypted data using a novel SQL aware encryption strategy and this system provides practical and strong level of confidentiality. In this paper, it describes that how queries are encrypted by CryptDB using MongoDB queries. In this section there are two threats : Threat 1 and Threat 2. Threat 1, in which database and administrator are not trusted but in Threat 2, proxy's and user application are trusted. CryptDB gives database server to implement these mongoDB queries on encrypted data almost same as it executes on plaintext data. CryptDB proxy server uses a master or private key(suppose MK or PK) through all over the transmission of data.

G) *Towards Bayesian-based Trust Management for Insider Attacks in Healthcare Software-De_ned Networks(2018)[4]. Authors: Weizhi Meng, Kin-Kwang Raymond Choo,Steven Furnell, Athanasio, V. Vasilakosand Christian W. Probst*

This paper represents a Software Defined Network we spot the attacks caused by insider in healthcare SDN using a trust Bayesian management for such type of environment. Here we have developed a trust based approach, which help us to identify the destructive devices in our medical healthcare environment. Trust inspect the nature of sensor forks. To increase the discovering of execution of one Intrusion Detection System(IDS), or splitting of IDS. IDS are used to identify any act of variance or rules contamination through detection of the beshield network and systems.

III. EXISTING SYSTEM

The security is an utmost necessity for applications of healthcare, especially in case of patient's data privacy. This project discusses the privacy and security issues in healthcare applications using wireless medical networks (WMNs). In these existing system, it discuss about security while data transmission, but it cannot protect the database server's where data is stored into three different servers. As patient's stored data needs to be secured, but it can't get the security so it may get easily modified by inside attacker. These systems is used for securing the communication between medical severs and data server, they used the lightweight encryption protocol and MAC generation protocol.

Disadvantages of Existing System:
Existing system has following Drawbacks

A. Less secure :It doesn't protect data during transmission as it's in human readable form

B. Cannot protect inside attacker :As data is been stored in single database, and security is provided to only only while it transmitted and no certain efforts are there to protect system from inside attack.

C. If any hackers get access to whole database: The data stored on single server, hence if hacker gets access to it whole information is accessible to that hacker.

eq. (3)

IV. PROPOSED SYSTEM

To overcome security issues in an existing system we have investigated Data Privacy Preservation through CryptMDB. As shown in Fig.1 CryptMDB system distribute data among three different database servers by using Sharemind technique.

A. Working:

Main aim of this system is to find compromised data and retrieve it. To find compromised node we use:

1. *Any modification in data base log file:*In first case if any tuple in database get modified then there is a modification in database log file. This can be used to find compromised node.
2. *Multiple login attempts:* If any DB admin login failed and if three continuous login attempt then we can lock that DB admin account.
3. *MAC based query:* If any query to controller (Web Server) is not from registered computer (i.e. Patient, Physician, Researcher) then we lock DB.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

B. Retrieval for Compromised Data:

As we get any above condition through which data get compromised following retrieval process is to follow:

1. If modification in DB log file.

- Decrypted the compromised patient information using Paillier Decryption algorithm.
- Re-encrypt same information using new ($P_{k,}$ S_k) and send this to any Proxy server.
- Send new S_k to respective patient to unloads the data.

2. In case multiple attempts to make login as Patient, Physician, Researcher or DB admin then lock that account for next defined time.
3. If any query from not registered then computer don't accept it.

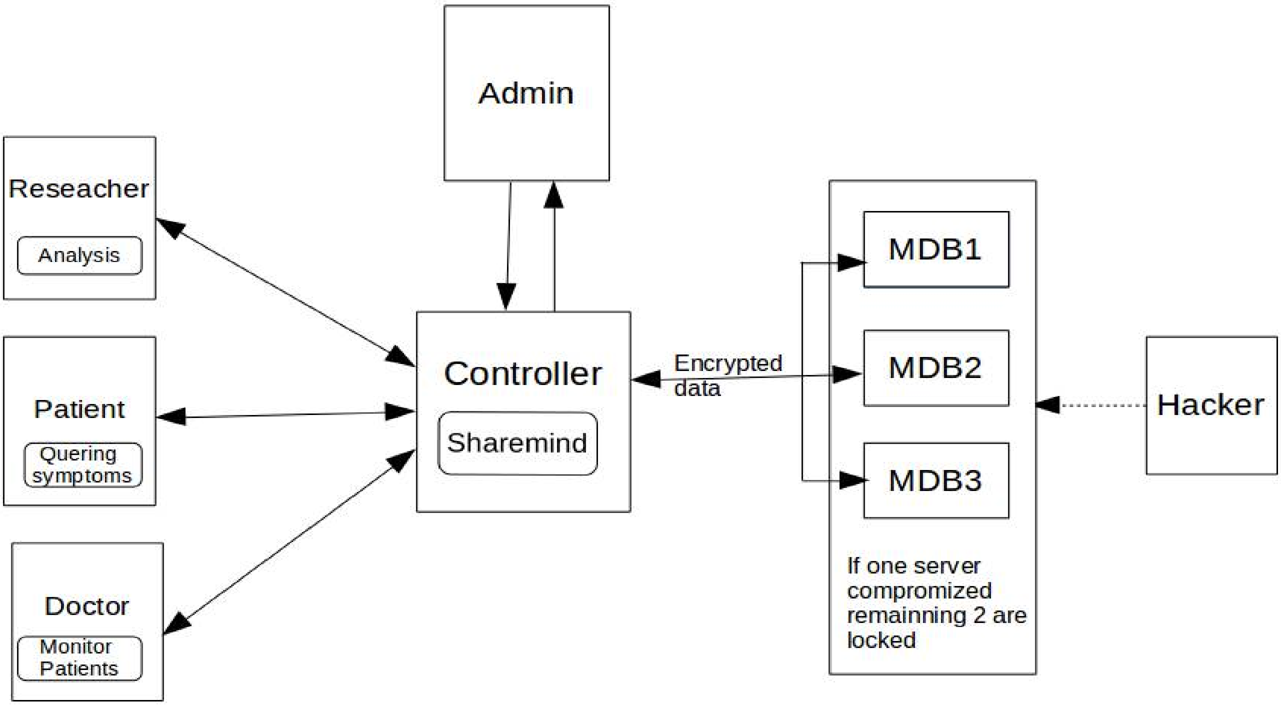

Fig. 1 – System Architecture

V. METHODOLOGY

1) Paillier Cryptosystem:

Paillier cryptosystem is an algorithm or technique created by Pascal Paillier with several interesting properties which is used for encryption. It's a strategy that can be used to obscure information, with a few interesting properties. These assets, when creatively applied, allows the Paillier cryptosystem to be used in different ways other than cryptographic method that simply can't be used.

2) Algorithm:

It includes the following key generation, encryption and decryption .

A) Key generation:

The key generation algorithm works as follows:-

- It chooses two of large prime numbers p and q randomly and independently of each other such that GCD among them is equal to one.
- Compute LCM among P and q
- Randomly Choose one integer g where g $\in\mathbb{Z}^*_{N^2}$ and selection of g such that order of g can be divided by N by checking modular multiplicative inverse.

The public (encryption) key pk is (N,g).
The private (decryption) key sk is (λ,μ).

B) Encryption:

Working of encryption algorithm -

- Let message M be a plaintext to encrypt, where M$\in\mathbb{Z}_N$
- Select any random numbers r where r$\in\mathbb{Z}^*_N$
- Ciphertext will be computed as:

$$C = g^M.r^N(mod\ N^2)$$

C) Decryption*:*

Working of decryption algorithm -

- Let ciphertext C is encrypted text to decrypt, where the ciphertext C$\in\mathbb{Z}^*_{N^2}$.
- Plaintext message will be computed as:

$$m = \left(c^\lambda(mod\ N^2)\right).\mu(mod\ N)$$

3) Advantages:

1. Prevent the inside attack by securely distributing the patient data into multiple data servers.
2. By employing the Paillier cryptosystems statistical analysis can be easily perform on the patient data without compromising the patients' privacy.
3. In architecture we can achieve data storage & data analysis security due to secured distributed database
4. Proposed data retrieval technique allow to retrieve the data compromised server(s)

- **TEST CASES**

1) Registration Test Cases

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter the number in username, middle name, last name field	Number	Error Comes	Error ShouldCome	P
001	Enter the character in username, middle name, last name field	Character	Accept	Accept	P
002	Enter the invalid email id format in email id field	ppgmail,com	Error Comes	Error Should Come	P
002	Enter the valid email id format in email id field	pp@gmail.com	Accept	Accept	P
003	Enter the invalid digit no in phone no field	99999	Error Comes	Error Should Come	P
003	Enter the 10 digit no in phone no field	9871997997	Accept	Accept	P

2) System Test Cases

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Sharemind Technique	Database	Divide database into 3 server	Accept	P
002	Paillier Cryptography Encryption	Database	Store data using Paillier Cryptography	Accept	P
003	Doctor request for data	System	Accept	Accept	P
004	Researcher request for data	System	Accept	Accept	P
005	Hacker try to hack the data	Detection of Hacker	Accept	Accept	P
006	Hacker get the information from 1 distributed server	Not collected all data	Accept	Accept	P

VI. CONCLUSION

In this paper, we have explored the issues in existing system regarding data security and privacy in data collection, data storage, queries and provide a complete solution for medical data. In this system we achieve reliable and secure communication by using Paillier cryptosystem algorithm, AES algorithm.

We forth put a new data collection protocol by using Sharemind technique, which divides the data into three different multiple servers and stores it. We have used CryptMDB which provides security along with large storage and fast access features of MongoDB.

ACKNOWLEDGMENT

We would like to thank all the authors of literatures which we have surveyed. We sincerely thank to the management of Modern Education Society's College of Engineering, Pune for supporting and allowing the survey and implementation of Data Privacy Preservation through CryptMDB.

REFERENCES

[1] Guowen Xu, Dongxiao Liu, Hongwei Li,Yang. "CryptMDB: A Practical Encrypted MongoDB over Big Data." IEEE ICC (2017)

[2] Jawwad A. Shamsi, Muhammad Ali Khojaye. "Understanding Privacy Violations in Big Data Systems ".IEEE (2018)

[3] Xun Yi, Athman Bouguettaya, Dimitrious Georgakopoulos, Andy Song. "Privacy Protection for Wireless Medical Sensor Data." IEEE Transactions on Dependable and Secure Computing: 369-380 (2015)

[4] Weizhi Meng, Athanasic V. Vasilakos, Kim-Kwang Raymond Choo.

[5] Raluca Ada Popa, Catherine M.S. Redfield, Hari Balakrishnan. "CryptDB: Protecting Confidentiality with Encrypted Query Processing". Association for Computing Machinery (ACM) 2016.

[6] Yau Liu, Shuai Xue."Accelerate the Paillier Cryptosystem in CryptDB by Chinese Remainder Theorem". International Conference on Advanced Communications Technology (ICACT)

[7] Ismail San, Nuray At, Ibrahim Yakut and Huseyin Polat. "Efficient paillier crypto processor for privacy-preserving data mining". Security and Communication Networks (2016)

DOI: 10.15680/IJIRCCE.2019. 0705065