Data Privacy Preservation through CryptMDB

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

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

***Abstract* : Dramatic increase in medical data made it to enter into era of big data, such a large volume of data need to be stored in such a way that it could be accessible to its authorized user by preserving confidentiality. Relational databases cannot address user's demands for fast data access and fast calculations because the data cannot be processed in distributed way. To overcome this problem, non-relational database such as MongoDB have been emerged up and been applied in various scenarios. In addition, the physiological data of a private square measure deeply touchy. Hence, security may be a foremost necessity of welfare applications, notably on account of patient protection, if the patient features a humbling unhealthiness. In this paper, we are focusing on the identification of attacker's in healthcare Software Defined Network's (SDN). Based on this survey, we try to develop a trust based management system from which we figure out malicious devices in a healthcare environment and also proposes a practical encrypted MongoDB(i.e. CryptMDB). As CryptMDB achieves better efficiency than existing relational database in terms of data access and calculating. The principle commitment of this paper is to disperse patient's data safely in various data server's by the Paillier cryptosystems to perform measurable investigation on the patient data.**

***Keywords:* MongoDB, Big Data, CryptDB, Software-Defined Networking, Trust Computation and Management, Data Base Security, Privacy Protection.**

I. INTRODUCTION

With rapid development in network and information technology data becomes primary concern into our everyday life [1]. To handle such large volume of data NoSQL databases are required. MongoDB is a type of NoSQL database, which models all data in form of documents and work on concept of collection [2][3]. Since MongoDB is NoSQL database it provides many storage advantages along with fast access, regardless of security measures [4]. Software defined network (SDN) can be used as a defend in medical network against attacks. There are many security issues in existing systems such as data-stealing, stealing and updating, storing the wrong values. Privacy protection is widespread where data is stored without any safety which is vulnerable to the hackers, as they want to use patient's private data. Suppose if the hacker is trying to hack the patient details, there are many chances for the misuse of data which may lead to severe consequences. The data can also be modified by the hackers due to lack of security. The treatment prescribed by the doctors can be hacked which may even lead to death of the patient. Patients become the victims because of the above issues [5]. To prevent these issues, data privacy preservation through CryptMDB system is proposed. This system is a system which is used to check the malicious activities and produces electronic reports to a management station [6]. It consists of AES algorithm, Paillier algorithm for encryption and Sharemind-technique for classification. The data is distributed stored among the three multiple servers. For instance, if the patient's sugar level is monitored 10 times per day then the data is updated in the database which is present in the one of the multiple server. Likewise, the values for blood pressure, heart beat, and temperature are also noted at regular intervals. The AES algorithm encrypts the patients details before storing it in the multiple servers and then it uses Paillier algorithm to re-encrypt the data [7].After that the data is stored in MongoDB server. If the hacker tries to modify or use the patient's data, a notification is sent to respective user and admin. The system also detects the server to which the hacker tried to hack amongst the multiple servers and locks down other two servers' as soon as it detects the compromised server.

III. LITERATURE SURVEY

Sr. No.	Paper Name	Author Name	Published Year	Advantages
1.	Sharemind: a framework for fast privacy-preserving Computations [7]	Dan Bogdanov, Sven Laur1, and Jan Willemson	2016	SHAREMIND—is a machine for secured way of preserving data processing that depends on share computing techniques. In multi party computing environment which recognize estimate functions.

2.	Towards Bayesian–based Trust Management for Insider attacks in Healthcare Software –Defined Networks[5]	Weizhi Meng, kim Kwang Raymond Choo, Steven Furnell Athanasios V. Vasilakos and Christian W. Probst	2018	SDN- we identify the insider attacks in healthcare SDN's by applying a trust Bayesian management for such type of environment.
3.	Privacy Protection for Wireless Medical Sensor Data[8]	Yi, Xun, Et al.	2016	By applying encryption techniques, try to protect data and server. As, patient's data is split into three servers and they are stored.
4.	Understanding Privacy Violations in Big Data Systems[16]	Jawwad A. Shamsi, Muhammad Ali Khojaye	2018	Big data is conductive in analyzing computational issues for predictive concept. In spite of this, they show major interest for parasite privacy.
5.	Privacy Preservation Data Analysis in Mental Health Research.[10]	Jingquan Li, Xueying Li	2015	Privacy is a ground laying pre- requisite for health research. Hence, in this paper the confidentiality of patient's data is preserved.
6.	MongoDB with Privacy Access Control[1]	Shweta Siriah, Bhushan Deshpande, Deepak Asudani	2018	A monitor is put for execution, which has been planned to work for forth put security. This monitor act as mediator between MDB user and MDB server which has access to keep control by recording data.
7.	Data Modelling for Discrete Time Series Data Using Cassandra and MongoDB[2]	Dharavath Ramesh, Ashay Sinha, Suraj Singh	2016	Cassandra model is best for storing and analyzing huge quantity of data in series like discrete time series knowledge. In the manner the data is introduce into the databank in order it gives fast access to data when queried.
8.	MongoDB NoSQL Injection Analysis and Detection[3]	Boyu Hou, Kai Qiam, Lei Li, Yong Shi, Lixin Tao, Jigang Liu	2016	Provides defense method to NoSQL database systems,to prevent injections or any other type of attack happening. They examines the maturity of security measures from MongoDB, a typical NoSQL database system with aspects both attack and defense at code level.

9.	Active Trust: Secure and Trustable Routing in Wireless Sensor Networks[4]	Yuxin Liu, Mianxiong Dong, Kaoru Ota, Anfeng Liu	2016	Active Trust significantly improves data route success probability and optimize network lifetime. Comprehensive theoretical analysis and experimental result indicates that performance of Active Trust scheme is better than previous study.
10.	CryptMDB: A Practical Encrypted MongoDB over Big Data[6]	Guowen Xu, Dongxiao Liu, Hongwei Li, Kan Yang	2017	Here we suggest a 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.
11.	Accelerate the Paillier Cryptosystem in CryptDB by Chinese Remainder Theorem[11]	Yau Liu, Shuai Xue	2018	It uses CRT(Chinese Remainder Theorem) to accelerate encryption process and improves the system performance under certain condition and also improve CyptMDB performance.
12.	CryptDB: Protecting Confidentiality with Encrypted Query Processing[12]	Raluca Ada Popa, Catherine M.S. Redfield	2016	CryptMDB efficiently run query over encrypted data using a novel SQL aware encryption strategy and this system provides practical and strong level of confidentiality.
13.	Achieving Authorized and Ranked Multi-Keyword Search over Encrypted Cloud Data[13]	Hongwei Li, Dongxiao Liu, Kun Jia, and Xiaodong Lin	2015	ARMS(Authorized and ranked multi-keyword search scheme) is efficient and more superior than existing approaches in terms of computational overheads and functionalities.
14.	Toward Exploiting Access Control Vulnerabilities within MongoDB Backened Web Applications[14]	Shuo Wen, Yuan Xue, Jing Xu, Hongji Yang, Xiaohong Li, Wenli Song and Guannan Si	2016	System exploits access control vulnerabilities within MongoDB back-end application and to solve the sophisticated data model,MongoDB accesses operation mode precisely represents the MongoDB action performed in web application has been introduced.

15.	Efficient Paillier crypto processor for privacy-preserving data mining[15]	Ismail San, Nuray At, Ibrahim Yakut and Huseyin Polat	2016	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.

III. PRELIMINARIES

In this section, we introduce the CryptDB architecture. Encryption tools are served as basic of our system.

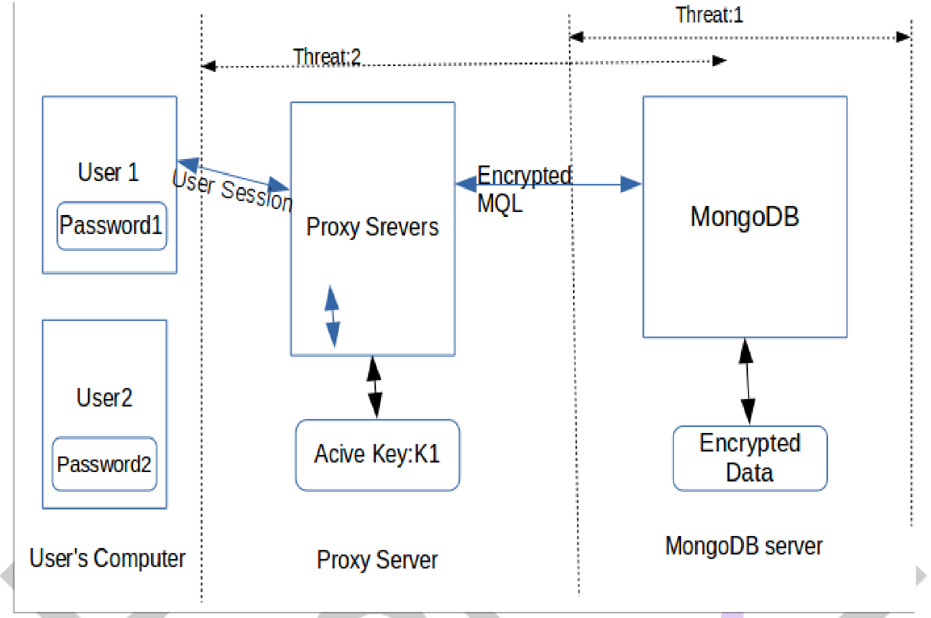

Figure 1. Block Diagram

1. CryptMDB Architecture:

As shown in figure(1), there are main three parts of cryptMDB system: Proxy servers, MongoDB server's and User computers. Initially an encryption tool encrypts data provided by user before storing it into MongoDB server. When user want to search particular contents of database, they need to send MQL to Crypt MongoDB proxy servers. At proxy server pre-encrypted tools re-write MQL queries and send to the MDB sever. As soon as MongoDB server receives that encrypted MQL queries it took efforts to execute MQL in order to match corresponding cipher text which will be delivered to proxy server . Eventually, the proxy server attempts these cipher texts and sends them to corresponding authorized users. But it does not give any access to tender data of particular users, which make sure that user's sensitive data cannot be disclosed to any crypt MongoDB server. Because in crypt MDB every user contain their own distinct key to encrypt their data. Although crypt MongoDB protects confidential data.

2. Compromised MongoDB (Threat 1):

MDB server is malice but inquisitive. On one hand , it executes MongoDB query language given by crypted proxy server, on the other hand it possibly attempts to submit the content of authorized user data and it acquire the association users knowledge. Besides, substitute server is assume to trustable in Crypt MongoDB. Therefore this section includes MDB software which binds the data. In this paper, we gainstay this threat by MDB server to implement MQL queries over cipher data in crypt MongoDB. The entered data will be encrypted firstly than these encrypted data will be stored in MDB. MDB server compares cipher texts when it incur the queries requested from proxy server. So it cannot get access to original data. Thus, MDB server can't access user's private data.

3. Arbitrary Threats:

CryptMDB server and proxy server will get compromised in case when arbitrary threats occur. Arbitrary threat (Threat 1) are more hazardous in comparison with threat1 as in case of arbitrary threats malicious user can directly get access to the data by utilizing keys. To intercept user's data from reveal to malicious user, we acquire distinct keys to encrypt each user data. These key becomes active only when corresponding user logging in to CryptMDB. Thus, other users data remain safe as even in case when attackers get access to entire system as they are able to decrypt only current active users data.

IV. KEY FEATURES

A. Confidentiality :

For threat MongoDB server should be honest but curious, while executing Mql queries MongoDB server can utilize the computing power to infer users information. But in CryptMDB Proxy server encrypts all the data before storing it in MongoDB .Users Mql queries are also encrypted before sending it to MongoDB server. After executing encrypted Mql queries over encrypted data MongoDB server returns cipher text to corresponding user without any seduction in information plain text.

B. Resistance to Threat :

Different keys are adopted to encrypt individual user's information. User's key only be activated at time when user login to MongoDB server. Thus, in case when MongoDB server and proxy server are compromised by attacker and try to get plain text from cipher text they can only steal current user data whereas other user's data remains safe.

V. ADVANTAGES

1. Practical approach to prevent the inside attack by securely distributing the patient data in multiple data servers.
2. Employing the Paillier cryptosystems to perform statistical analysis on the patient data without compromising the patients' privacy.
3. In Proposed system, Due to secured distributed database architecture we can achieve data storage & data analysis security.
4. Proposed data retrieval technique allow to retrieve the data compromised server(s).

VII. CONCLUSION

We have investigated the safety and privacy problems with help of cryptMDB and given an entire resolution for privacy-preserving and protect patient's data. Privacy is much important fundamental requirement in healthcare field, the idea of using cryptMDB is it utilizes homomorphic asymmetric cryptosystem for encrypting patient's data and SDN helps in decoupling of network control and helps in identifying insider attacks by using Bayesian approach. For the privacy of the patient's information, we tend to projected a brand new information assortment protocol that splits the patient information into 3 numbers and stores them in 3 multiple information servers, severally. As long joined information server isn't compromised, the privacy of the patient's information will be preserved. Using cryptMDB we can also achieve strong privacy and it is much better than relational database in terms of calculations and data access.

ACKNOWLEDGEMENT

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 regarding Data Privacy Preservation Using CryptMDB.

REFERENCES

[1]Shweta Siriah, Bhushan Deshpande, Deepak Asudani. "MongoDB with Privacy Access Control". International Journal of Research and Review 2018(IJRR).
[2]Dharavath Ramesh, Ashay Sinha, Suraj Singh. "Data Modelling for Discrete Time Series Data Using Cassandra and MongoDB". Recent Advances in Information Technology (RAIT) (2016)
[3]Boyu Hou, Kai Qiam, Lei Li, Yong Shi, Lixin Tao, Jigang Liu. "MongoDB NoSQL Injection Analysis and Detection". IEEE (2016)
[4]Yuxin Liu, Mianxiong Dong, Kaoru Ota, Anfeng Liu. "Active Trust: Secure and Trustable Routing in Wireless Sensor Networks". IEEE (2016)
[5] Weizhi Meng, Athanasic V. Vasilakos, Kim-Kwang Raymond Choo. "Toward Bayesian-based Trust Management for Insider Attacks in Healthcare SDN". IEEE (2018)
[6]Guowen Xu, Dongxiao Liu, Hongwei Li, Kan Yang. "CryptMDB: A Practical Encrypted MongoDB over Big Data." IEEE ICC (2017)
[7] Dan Bogdanov, Sven Laur, and Jan Willemson. "Sharemind: A framework for fast privacy preserving computation "(2018)
[8]Yi, Xun, et al. "Privacy Protection for Wireless Medical Sensor Data." IEEE Transactions on Dependable and Secure Computing: 369-380 (2016).
[9]D. He, S. Chan and S. Tang. "A Novel and Lightweight System to Secure Wireless Medical Sensor Network." IEEE Journal of Biomedical and Health Informatics: 316-326 (2014).
[10] Jingquan Li, Xueying Li. "Privacy Preservation Data Analysis in Mental Health Research." IEEE (2015).
[11]Yau Liu, Shuai Xue. "Accelerate the Paillier Cryptosystem in CryptDB by Chinese Remainder Theorem". International Conference on Advanced Communications Technology (ICACT).
[12]Raluca Ada Popa, Catherine M.S. Redfield, Hari Balakrishnan. "CryptDB: Protecting Confidentiality with Encrypted Query Processing". Association for Computing Machinery (ACM) 2016.
[13]Hongwei Li, Dongxiao Liu, Kun Jia, and Xiaodong Lin. "Achieving Authorized and Ranked Multi-Keyword Search over Encrypted Cloud Data". IEEE ICC (2015).
[14]Shuo Wen, Yuan Xue, Jing Xu, Hongji Yang, Xiaohong Li, Wenli Song and Guannan Si. "Toward Exploiting Access Control Vulnerabilities within MongoDB Backened Web Applications". IEEE (2016).
[15]Ismail San, Nuray At, Ibrahim Yakut and Huseyin Polat. "Efficient paillier cryptoprocessor for privacy-preserving data mining". Security and Communication Networks (2016)
[16] Jawwad A. Shamsi, Muhammad Ali Khojaye. "Understanding Privacy Violations in Big Data Systems ".IEEE (2018).