An Analytical Study on Data Security using Blockchains and AI

Kahandal Kalyani[1] Sonawane Simran[2] Shah MohammedSalik[3] Prof. S. K Shinde[4]
[1,2,3,4]Department of Computer Engineering
[1,2,3,4]KJCOEMR, Pune, India

Abstract— Immediate and effective treatment of the patients is the central goal of the Doctor. The identification of the ailment is one of the most important first steps that the doctor performs towards improving the state and wellbeing of the patient. This is done through various tests and correlating the symptoms and the results of the disease. When there isn't a strong correlation, the doctor can refer the previous cases with the same parameters to help the patient rather than perform trial and error on the patient until a particular treatment is deemed fruitful, which can be a highly painful ordeal for the patient. Therefore, there are data vendors that accumulate and disperse medical data amongst various hospitals and clinics, but most of the vendors and aggregators have a low level of trust which hinders data sharing between them. Therefore, the Blockchain platform can be of use here as it can guarantee the security of the data, and combined with Artificial Intelligence an effective access control mechanism is formulated that can guarantee that unauthorized personnel do not get access to sensitive Public Health Records.

Keywords: Blockchain, Hash key Generation, Key Validation

I. Introduction

A blockchain is a digital, public ledger that records online transactions. Blockchain is the core technology for cryptocurrencies like bitcoin. A blockchain is similar to a bank log, but open and approachable to everyone who uses the cryptocurrency.

The blockchain was designed to be used as a book-keeping platform. Blockchain is a system consisting of a decentralized network that is used as a public ledger for all the databases and information that the company has. It can hold an enormous amount of data. Its features make the best platform for businesses to use it for storage and data management purposes.

As a system and decentralized network and as a public ledger for all the databases and information that the company has it can provide high level of security. It can hold an senormous amount of data. Its security features make the best platform for businesses to use it for storage and data management purposes. It aims to change the way we store, manage and transfer data.

Blockchain is a distributed system where nodes are not necessarily trustworthy. AI gains little from being distributed over a network, therefore, Artificial Intelligence is best used for problems which require noticing patterns inferring rules of behavior, predicting eventual outcomes, determining underlying causes. The one use where these could come together would be in distributing the data a blockchain needs to contain.

Today, Bitcoin is pushing the limits of nodes; memory dedicated to this distributed ledger takes too much space. Casual users are not willing to dedicate that much of their machines memory to serving the worldwide network.

It is possible that an AI system could distribute the data such that each node would contain a reasonable amount of the whole blockchain while at the same time the whole blockchain would be somewhere on the network. In pieces, but all intact and distributed securely.

Running an AI sub-system with enough smarts to keep track of a large growing distributed ledger sounds daunting proposition, but it is conceivable that each node needs do nothing more than mimic the way neurons in our brains find memories. Any single node need not be very smart; the collection altogether acts like a smart entity.

Blockchain technology allows patients to allot their own rules for their medical data, for example, authorizing researchers to access parts of their data for a fixed duration of time. With blockchain technology, patients can associate with other hospitals and collect their medical reports as per their convenience.

Blockchain healthcare use cases are being identified every day, and with them, the healthcare system can be completely modernized. Many healthcare and blockchain companies are right now working on or have already released blockchain-based systems to develop healthcare record sharing for both experts and patients. By allocating the patient's health history, tracking the pharmaceuticals, and improving payment options, blockchain is becoming a beneficial tool for healthcare by remodeling the industry worldwide.

Support vector machine (SVM) is capable of data assignment and thereby finds its applications in the retail industry for summary like detecting customer spending habits, deviation detection and customer behavior position.Secure SVM, a secured modified version of the SVM algorithm was developed and utilized for training over the privacy-conserving SVM scheme using blockchain-based encoded data. Blockchain design and architecture helps to build a protected and dependable data-sharing platform among multiple retail stores of same the same brand, where IoT data is permitted and then documented in a shared ledger. The removal of a trusted party and introduction of reinforcing security assures the confidentiality of the sensitive data for each data provider as well as the SVM model parameters. There are a lot of feature abstraction and template production that is used for the purpose of training data for modeling Secure SVM.

Administered Learning Algorithms like Boosting, Bagging, Gradient Boost, AdaBoost, Random Forest, Extra Trees can be used to exposing Bitcoin Blockchain invisibility in retail payments. Bitcoin is a cryptocurrency whose transactions are saved on shared, openly available ledger. As Bitcoin provides a high degree of anonymity with an entity's real-world identity hidden behind a pseudonym, supervised learning algorithms find great applications to reveal the anonymity of buyers from different transactions by clustering bitcoin addresses. The attackers have devised a mechanism to identify possible owners of a bitcoin cluster by predicting the category of a yet-unidentified cluster.

K-means Clustering is used in payments to identify malevolent action. As the operator is accepting bitcoin as a form of payment, the transformation of bitcoin in different apps is set to change the payment scheme. As bitcoin uses blockchain technology, grant, allocation, processing data between different parties over a network, it becomes essential to recognize abnormal nodes which may be malevolent and muddled in illegal activity. The K-mean paradigm allows the use of clustering technique to check and accept transactions between legitimate users and reject it otherwise. A blockchain can be clustered in groups based on its behavioral pattern. The behavior is analyzed and the parameters selected along with the time taken for one transaction and the amount involved from one node to another node. The reason for selecting this parameter is that usually the transaction amount is the most important and predominant feature of a node and for doing this the algorithm used is K-means.

The algorithm is implemented after extracting the sequence data to represent node behaviors, and then clustering the nodes into categories. After analyzing the flock behavior, the classical behavior models for each category can be selected as behavior figures, to identify strange behavior models that do not conform to any template. Furthermore, flock behavior models into categories that may both lead to profound judgment into the blockchain network and help maintainers manage and organize the nodes.

II. Literature Survey

This section of the literature survey eventually reveals some facts based on thoughtful analysis of many authors work as follows.

1) W. Dai [1] explains that the ability to handle an activity on large amounts of data efficiently is critical, as such data can inform decision making in both public section and private section and the output is beneficial in increased productivity and reduced bureaucracy. There are a number of hindrances in conventional data trading platforms is due to the buyer and broker who are dishonest. Proposed a secure data trading platform (SDTP) on Ethereum& Intel's Software Guard Extensions (SGX) and carry out an in-depth analysis. In coming time research will involve source data tracking to further enhance the security of seller's raw data.
2) G. Magyar [2] introduces blockchain as secured, cryptographically immutable, write once, in addition to containing blocks of records, where the blocks are connected together using a primary key referencing mechanism. The data available in blockchain can be authorized or unauthorized. In an authorized blockchain the security, validity, and integrity of the blockchain are generally certified by a well-known legal entity and in unauthorized blockchain the security and integrity of the blockchain are secured by the majority of the unconventional peers. Thus, the proposed technology solves essential issue of retrieving data without threatening personal privacy and it also opens new opportunities for automatic personal monitoring devices.
3) Dr. M. Kumar [3] elaborates on the era of information technology as nowadays everyone's life is dependent on it. Thousands of cyber-attacks taking place around the world attackers are constantly evolving by using elegant and stealthy techniques to target the victim. To keep any organization shielded and out of harm's way the network monitoring and log auditing is one of the most chief processes. Network security devices can come with sufficient security but as the attackers keep changing their attacks as they have many master plans, security devices also have to be regulated accordingly. The authors in the proposed paper not only provide a safe log storage platform for long time but also gives the potential for deep analysis of archived log record to better understand the complex attack pattern.
4) Z. Guan [4] narrates the speedy development of computer and information technology due to which various types of data are being produced with extraordinary speed by people or smart devices. A huge volume of data, often termed big data, is of great value in various applications. Government agencies maintain data about personal income, tax information, etc financial institutions maintain banking transactions, mortgage loans thus the data owners don't want to share their data due to privacy. The authors proposed two data sealed, honest and efficient data trading schemes depend on the hash chain micropayment approach and the blockchain technique.
5) M.Chowdhury [5] explains sensitive private information such as electronic medical records and academic records is categorized as personal data. Personal information is generally arranged across many data systems. The data-consumer with whom the data is divided often needs to check the authenticity of the shared document/record by communicating with the document/certificate issuing authority. In proposed technology, the data is shared through a blockchain network which is a framework that will assure authenticity. The proposed work decreases the complexity time for data sharing, and make better decision-making process and minimizes the overall cost.
6) M.Shen [6] estimates that smart cities will include more and more Internet-of-Things (IoT) infrastructures ensuring a huge amount of data gathered from different IoT devices such as manufactoring, energy transmission, and agriculture. Paper presents a novel privacy-preserving SVM training scheme called a secure VM which handles data privacy and data integrity by using the blockchain methods to build an SVM training technique in multi-part scenarios where data is collected from multiple data providers. In coming years the researchers sanctioned building a wide range of privacy-preserving ML training techniques on multi-part encrypted datasets.
7) Uchi.Uchibeke [7] extracts most of the new big data are available on Cloud Computing Technology where the development of many computation services to accommodate the drastic grow the volume and velocity of big data are present. Benefits, such as cost-effective and reliable, scalable and distributed computing of large datasets are useful to businesses and in academics. A private and permissioned Hyperledeger blockchain architecture is presented to control access management and ensures data transparency and traceability for secure

data sharing with auditable and data self-sovereignty for the holder.

8) S. Ramamoorthy [8] presents the important task for the data centers is to maintain the data storage devices which stock the data collected from the multiple users. Data can be used for any business applications they can take usefulness of massive storage space provided by the cloud computing environment to store the data. Examples of cloud storage space in real-time are Amazon Simple Storage Systems (S3), EMC Mozy, Rackspace. High-level confidentiality is maintained by the cloud storage space. Blockchain Technology is one of the recent technique in the field of data security this technology makes use of strength of hash function and its unique values to secure data. Thus, cloud networks should get approval from other group users. It is not possible to revitalize the same hash value for the attackers.
9) W. Liang [9] introduces a technique that can decrease the communication load of node consensus, simplify the consensus flow in data transmission, upgrade the security of data transmission. With the speedy and advancements in the industrial Internet of Things (IoT), production efficiency has been greatly boosted. The proposed technique uses blockchain-based dynamic secret sharing mechanism through a reliable trading center is recognized using the power blockchain sharing Model. The proposed approach can upgrade the transmission rate and packet receiving rate by 12% and 13%, respectively and good superiority in sharing management.
10) P. Urien [10] approaches small objects that are polled by controllers that push their transactions to the ethereum blockchain. For blockchain transactions trust issues is one of the most important factors. Such Transaction is based on the ECDSA signature which is based on 32 bytes secret keys. To prevent the data to be hacked they have built the system over a JC3.04 standard platform is designed by interface CryptoCurrency Smart Card (CCSC). The proposed paper deals with the integrated sensor data in ethereum transactions and deployment of the CCSC application at low cost and less power object powered by an open hardware platform and second is deployment in the cloud of CCSC java cards plugged to RACS servers.

III. Conclusion

This survey paper has elaborated on the related research that has been done in the previous years on this paradigm. The various different techniques that have been used for the preservation of the privacy of the patients Public Health Records in the data vending approach have been studied in detail and identified that most of the access control mechanisms suffer from the problem of automation which has been improved by our proposed methodology. For the purpose of Future work, the presented technique will be elaborated further in detail in the upcoming editions.

References

[1] Weiqi Dai, Chunkai Dai, Kim-Kwang Raymond Choo, Changze Cui, Deiqing Zou, and Hai Jin,” SDTE: A Secure Blockchain-based Data Trading Ecosystem” , IEEE Transactions on Information Forensics and Security. DOI 10.1109/TIFS.2019.

[2] Gábor Magyar,” Blockchain: solving the privacy and research availability tradeoff for EHR data” 2017 IEEE 30th Jubilee Neumann Colloquium • November 24-25, Budapest, Hungary, 2017.

[3] Dr. Manish Kumar1, Ashish Kumar Singh, Dr. T V Suresh Kumar3,” Secure Log Storage Using Blockchain and Cloud Infrastructure” 9th ICCCNT 2018 July 10-12, IISC, Bengaluru Bengaluru, India 2018.

[4] Zhangshuang Guan, Xiaobei Shao, and Zhiguo Wan,” Secure, Fair and Efficient Data Trading without Third Party Using Blockchain ” ,IEEE Confs on Internet of Things, 2018.

[5] Mohammad JabedMorshed Chowdhury∗, Alan Colman∗, Muhammad AshadKabir, Jun Han∗ and Paul Sarda,” Blockchain as a Notarization Service for Data Sharing with Personal Data Store” , 17th IEEE International Conference On Trust, 2018.

[6] Meng Shen, Xiangyun Tang, Liehuang Zhu, Xiaojiang Du, Mohsen Guizani,” Privacy-Preserving Support Vector Machine Training over Blockchain-Based Encrypted IoT Data in Smart Cities ” IEEE Internet of Things Journal. ,10.1109/JIOT.2019.

[7] UchiUgobameUchibeke, Sara HosseinzadehKassani, Kevin A. Schneider, Ralph Deters,” Blockchain access control Ecosystem for Big Data security”, IEEE, Confs on Internet of Things, 2018.

[8] S. Ramamoorthy, B.Baranidharan,” CloudBC-A Secure Cloud Data access Management system” 978-1-5386-9371-1/1900c, IEEE, 2019.

[9] W. Liang, M. Tang, J. Long, X. Peng, J. Xu and K. C. Li” A Secure Fabric Blockchain-based Data Transmission Technique for Industrial Internet-of-Things” , ,IEEE Transactions on Industrial Informatics DOI 10.1109/TII.2019.

[10] Pascal Urien,” Towards Secure Elements for Trusted Transactions in Blockchain and Blockchain IoT (BIoT) Platforms.” SEC 1: Elliptic Curve Cryptography Certicom Research, May 21, 2009.