Securing Data in Distributed System using Blockchain and AI

Abhijeet Sontakke[1], Aditya Gaikwad[1], Neha Ovhal[1], Kalyani Nagawade[1], Sarita Patil[2]

[1]Student, Dept. of Computer Engineering, G. H. Raisoni COE & Management Pune, Maharashtra, India
[2]Professor, Dept. of Computer Engineering, G. H. Raisoni COE & Management Pune, Maharashtra, India

Abstract - *In hospitals in developing countries when an individual with symptoms corresponding to a deadly disease such as hepatitis, influenza, cancer, etc. is encountered, the staff can either suggest an expert in that particular field or if the disease hasn't been identified, the staff could recommend the patient to a general physician. Most of the time, the general physician will subject the patient through a series of trial and error to identify and eliminate the improbable disease. This process is a very painful ordeal experienced by the patient as the doctors try to determine the cause of the symptoms, one disease medication at a time. This problem has largely been overcome by developed countries by a Data Vendor. This paper analyses all the past work on providing the security for the data in the network where especially data brokers are being involved. The best way to provide this is, introducing of block chains and Artificial intelligence. This research article analyses past methodologies and evaluate them to introduce a newer version of maintaining data integrity in network.*

***Key Words*: Block chain, Artificial intelligence, Intrusion detection system, Network threats, Access control mechanism.**

1. INTRODUCTION

Data and its security have been paramount since the day human beings started accumulating it. Since the earliest stone age era, humans have evolved to process the information around us by various sensory inputs received by different sensory organs. This has led to the development of various arts such as paintings, music, textures, etc. This thirst for the unknown has made us go into space and advance a significant amount in technology over the years.

All of this data created needs to be stored somewhere where it cannot be modified or destroyed, as these are valuable lessons learned throughout the lifetime of an individual or a group which can help the species to grow forward. This has led to the humans collecting and storing a lot of data on various different topics and was accelerated due to the invention of the printing press at which allowed the information to take the form of books that could retain it for a very long time.

Another revolution in the storage, retrieval, and access of data happened when the internet was conceived. Electronic storage was already invented by then but the internet added another element to this as the internet allowed the various computers and computing devices all over the world to connect to each other and share information. This was designed to facilitate the exchange of information between the researchers over a large distance to eliminate the need to be physically present at the location to utilize the resources. Due to its initial success, the internet was opened to the public and various different services that used the internet as the backbone started flourishing.

The internet started growing exponentially with a lot more users and machines being connected every day. People started using the platform more and more and this led to an increased number of users interacting online. With social media and educational portals, the internet grew to astronomical sizes and the data being produced every day grew to a massive size. With the growth of data and the users online, it created a nourishing environment for people to learn and share valuable skills and information all over the world. The major drawback of open access to everyone on the internet was that there are also some individuals with malicious intent that can ruin the experience of another person purely for personal gain.

A lot of users on the internet have valuable and sensitive personal information that is stored in the databases and various organizations to have their internal data that is confidentially stored electronically. This increases the likelihood of an attacker gaining access to this information that would lead to compromised security as well as a huge loss for the organization. This is problematic as there are no alternatives for storage and the convenience offered by the database. Therefore, there is an utmost need to provide a mechanism for controlling the access to the sensitive data through which only the trusted employees and other members of the organization can access the data, based on their hierarchy.

This brings forth the blockchain paradigm, being proposed by a group of scientists in the late 1990s the technique was initially developed for use in a digital notary as it is an excellent choice for tamper-proofing a document. The paradigm was largely unused until it was utilized for the creation of the world's first cryptocurrency. Due to its strong tamper-proof and distributed nature, it is an apt choice for a cryptocurrency. The blockchain is one of the most secure applications and is capable of providing very high security to the data stored and can, therefore, be utilized in an application to safeguard the data and provide an efficient as well as an effective access control mechanism for the sensitive data of an organization.

This paper dedicates section 2 for analysis of past work as literature survey and section 3 concludes the paper with feasible statement of the literature study.

2. LITERATURE SURVEY

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

S. Yu states that due to rapid pace of technology and the introduction of Internet of Things or IoT, there has been a swift increase in the number of intelligent devices being connected to the internet. These devices are capable of generating a large amount of data due to the fact that it is connected and is interacting with the internet. A large number of devices generate equally large amount of data which cannot be processed efficiently. Therefore, the authors propose an effective technique based on blockchain that can provide a low-cost alternative to create economic value for the IoT data generated. A major drawback of this methodology is that it has the potential to be misused by uploading large amounts of malicious data.

R. Wang elaborates on the foundation of network security construction which is the PKI or The Public Key Infrastructure. The researchers also commented on the reliability and the robust security offered by the blockchain platform. Therefore, the authors amalgamated both the methodologies to strengthen the Public Key Infrastructure by a permissioned blockchain that converts the PKI into a privacy-aware PKI. This is crucial as the implementation of a permissioned blockchain also improves the efficiency of the configuration and certificate application. The major drawback is that this technique has been a very specialized approach towards the blockchain paradigm.

C. Ehmke explains that the innovative paradigm of Blockchain has been popular and has seen extensive usage recently. The blockchain was utilized for financial applications and that is how it gained immense popularity and limelight. The blockchain was readily picked up by a plethora of researchers and implemented in various different fields, which has greatly helped in bringing increased security to numerous applications. Due to the fact that the blockchain paradigm requires a user of the blockchain to download the whole chain to gain an overview. To ameliorate this effect, the authors have implemented a scalable and lightweight blockchain protocol.

R. Wang introduces the video surveillance system as an irreplaceable tool that can be used to efficiently manage and survey big cities. When a video surveillance system is installed it can easily transmit environment information remotely, this is highly useful as the person does not need to travel long distances and physically be present in the location for the management. Due to a large-scale increase in the monitoring standards with the inclusion of IoT and real-time monitoring, it is susceptible to attacks. Therefore, the authors developed a system for video surveillance based on permissioned blockchains and Convolutional Neural Networks for a seamless and secure system. A Major drawback in the system is that large scale testing of the System has not been performed and will be done in the upcoming researches.

J. Lou states that there has been a lack of a key management feature in the Named Data Networking, which is utilized to name each and every object by the producer and also digitally sign it. There are some disadvantages of the conventional approach such as lack of trust between the sites as well as the high chances of failure observed in the centralized architecture if the main node fails. Therefore, the authors in this paper propose an efficient key management scheme based on blockchain for the Named Data Networking paradigm. The blockchain increases the trust between the sites as well as the decentralized architecture is highly useful in overcoming failure. The drawback of the proposed scheme is that it has not been evaluated extensively for its feasibility in reducing the NDN cache pollution.

S. Wang explains that there has been a very fast development of cryptocurrency in recent years, which has led to detailed scrutiny of the paradigm. This has uncovered a lot of irregularities in the paradigm such as the Smart Contracts that have been the cause of “The DOA Attack” which has resulted in a huge loss. Therefore, the authors have presented a comprehensive and systematic review of the smart contracts in the blockchain paradigm. The authors have presented a six-layer architecture for smart contracts for increasing the security of the system. The authors have not implemented a formal verification which can provide confidence.

Y. Xu introduces the concept of decentralized storage that is based on the blockchain framework. The blockchain is one of the most innovative concepts that can be used to design a highly secure decentralized framework. The authors have proposed section blockchain protocol, which aims to eliminate the storage problem that is encountered in certain devices. The proposed methodology is highly resilient to failure due to the decentralized architecture, as well as, it has the ability to withstand heavy loads and optimization gracefully due to the implementation of the Blockchain paradigm.

M. Marchesi in his keynote speech details the rapid development and ongoing researches going on in the field of blockchain. This is due to the increased attention to this paradigm and increased demand in this sector. The author indicates that this increased pressure on a nascent framework has led to an increase in security lapses that are evident in the various different incidents on the Ethereum platform and the cryptocurrency exchanges.

That being said the speaker also highlighted the immense opportunities that can be utilized by using the blockchain paradigm such as the Blockchain tokens that can be used to implement a reward and penalty-based scheme for developers.

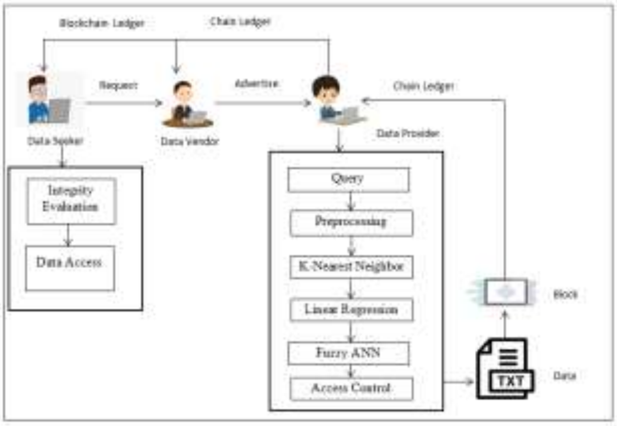
A. Maksutov elaborates on the paradigm of Blockchain and its uses. The authors have proposed an innovative concept for the detection and identification of various money laundering schemes that use the blockchain framework for their nefarious activities. The proposed methodology has been used to deanonymizing the transactions and tracking the coin join transactions, which allows the authors to evaluate user participation. All of this information is used to determine if the transactions are being fraudulent or used to launder money.

F. Wessling states that the addition of blockchain to existing platforms is problematic as it is different from building the applications from scratch by incorporating the Blockchain into the application. The authors determine the amount of blockchain required for various different implementations, this is done by analyzing the attributes of blockchain such as anonymous, trustless and immutable, etc. The authors have outlined the various different processes that utilize various different elements of the blockchain technology that can be implemented based on the specific application and use case of the application.

J. Wang explains that most of the applications based on crowd sensing gather a huge amount of data by using pervasive smartphone users to provide the data. But most of the time the users are not compensated enough for their contribution to the system. Therefore, the authors have presented an innovative framework for a privacy-preserving reward and penalty scheme that rewards the users for contributing to the large data sensing paradigm using the trustless and secure blockchain. The major drawback of this paper has been that the authors have not discussed the solutions for a possible collusion attack.

S. Pandey introduces the benefits of utilizing the Blockchain paradigm for its security and decentralized architecture. The author states that there has been a jump in the number of researches is going on in this field and there has been increased interest in implementing the blockchain framework to make existing systems resilient and secure. To this effect, the authors have formulated an ingenious and practical simulation tool for planning, stability, and design of the systems and applications as well as networks in a blockchain environment. The BlockSIM is an opensource and comprehensive solution for all the Blockchain simulation needs. The major drawback is that the authors have not modeled the internet latency that would affect the accuracy of the simulation.

3. SYSTEM OVERVIEW DIAGRAM



4. CONCLUSION

In this paper there have been some related works that have utilized the Public health record database and techniques to achieve various different approaches and to identify their flaws and shortcomings. There have been various methodologies that have been proposed by a plethora of authors each offering a unique technique to Public Health Record management. This has influenced our approach drastically and helped and enabled us to propose a secure and efficient Public Health Record management system that uses the Blockchain Paradigm. The Blockchain is used due to its inherent nature of being resilient to changes and tampering, this is utilized to provide an effective Access control Mechanism that can help restrict the leakage of valuable sensitive data of the patients in the Public Health Records. The methodology discussed will be implemented in the upcoming researches.

REFERENCES

[1] S. Yu, K. Lv, Z. Shao, Y. Guo. J. Zou and b. Zhang, “A High-Performance Blockchain Platform for Intelligent Devices”, 1st IEEE International Conference on Hot Information-Centric Networking, HotICN, 2018.

[2] R. Wang et al, “A Privacy-Aware PKI System Based on Permissioned Blockchains”, IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 2018.

[3] C. Ehmke, F. Wessling and C. Friedrich, “Proof-of-Property – A Lightweight and Scalable Blockchain Protocol”, ACM/IEEE 1st International Workshop on Emerging Trends in Software Engineering for Blockchain, 2018.

[4] R. Wang et al, “A Video Surveillance System Based on Permissioned Blockchains and Edge Computing”, IEEE International Conference on Big Data and Smart Computing, BigComp, 2019.

[5] J. Lou, Q. Zhang, Z. Qi and K. Lei, “A Blockchain-based key Management Scheme for Named Data Networking”,

Proceedings of 2018 1st IEEE International Conference on Hot Information-Centric Networking, HotICN 2018.

[6] S. Wang L. Ouyang et al, “Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends”, IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019.

[7] Y. Xu, “Section-Blockchain: A Storage Reduced Blockchain Protocol, the Foundation of an Autotrophic Decentralized Storage Architecture”, 23rd International Conference on Engineering of Complex Computer Systems (ICECCS), 2018.

[8] M. Marchesi, “Why Blockchain Is Important for Software Developers, and Why Software Engineering Is Important for Blockchain Software”, International Workshop on Blockchain Oriented Software Engineering (IWBOSE), 2018.

[9] A. Maksutov et al, “Detection of Blockchain Transactions Used in Blockchain Mixer of Coin Join Type”, IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), 2019.

[10] F. Wessling et al, “How Much Blockchain Do You Need? Towards a Concept for Building Hybrid DApp Architectures”, ACM/IEEE 1st International Workshop on Emerging Trends in Software Engineering for Blockchain, 2018.

[11] J. Wang et al, “A Blockchain-based Privacy-Preserving Incentive Mechanism in Crowdsensing Applications”, IEEE Access, 2018.

[12] S. Pandey et al, “BlockSIM: A practical simulation tool for optimal network design, stability, and planning”, IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2019.