Secure E-Documents Storage using Blockchain

Ayush S. Ghanghoria[1], A Sahaya Anto Raja[2], Vivek J. Bachche[3], Ms. Neha Rathi[4]

[1,2,3]Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, Maharashtra, India
[4]Assistant Professor, Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, Maharashtra, India

---

Abstract - *Making of false documents and certificates by forging it from the original ones have become more popular nowadays. Also, those documents are stored in their respective offices in a database which is a Centralized Database. Now if we need to access them then we should be physically present in that particular office. Also, these offices are subject to natural disasters like floods, earthquakes, etc. which if attacked by any of these calamities will lead to the loss of all the documents permanently. The blockchain method serves as a solution for safe storage of the documents in the form of e-documents on a blockchain network. The immutability property will help the document from being replaced by another document and also will result in permanent entry of a particular document in Distributed Databases. Also, the blockchain property of distributed ledger system will help to keep the documents unmodified and unaffected forever.*

***Key Words*: Centralized Database, forging, Blockchain, immutability, distributed ledger system.**

1. INTRODUCTION

Storage of documents has become a very important thing in our current lifestyle. Also, with this rising importance of documents, the theft and forgery of documents is also increasing day-by-day. So, there is a very great need for the protection and authentication of documents to keep it from being destroyed, manipulated or forged into some other document. Also, some documents are so important in a way that if they are lost, they cannot be regenerated or restored. Also, because of losing the document, one can easily lose his/her ownership over the property which was being authorized to that particular person through the document. Apart from all this, the person has to personally visit the respective office for restoring their document and this process takes time. Because of this, importance of document keeping and its storage, the bribing in our society also has increased with a great increase in forgery and creation of any educational or property related document by giving particular amount of money. This increases frauds and thefts as well as bribing in our society. All the above-mentioned problems can be eradicated by using blockchain as a system of storage for storing documents. This will be a means for easy recovery of document, complete authentication of data and easy accession of data as well. All of these systems will help us to greatly reduce frauds, bribes and theft of document and also will help to preserve the ownership of a particular person over his/her own document.

2. BACKGROUND

Document storage currently, in this era, is done by storing it in a centralized database system. If in this centralized DB system one change is made to the database, then every other people/office connecting to that database for getting information will get the false results. Also, there is deletion and updating of data in the centralized DB system, thus making the changed information irrecoverable. Thus, the forgery of documents and frauds can be easily done in the current era. These problems are dealt in our model by using a decentralized database, instead of centralized databases. Decentralization of data can prove to be helpful in the following ways:

- To make a change in the information at least 51 percent of nodes information has to be changed.
- Since blockchain doesn’t have delete or update features, so the data/document can be easily recovered from any node.

The concept of blockchain was proposed by Satoshi Nakamoto in 2008. Blockchain is an online ledger that provides decentralized and transparent data sharing. With distributed recordings, all transaction data (that is stored in nodes) are compressed and added to different blocks. Data of various types are distributed in distinct blocks, enabling verifications to be made without the use of intermediaries. All the nodes then form a blockchain with timestamps. The data stored in each block can be verified simultaneously and become inalterable once entered. The whole process is hospitable to the general public, transparent, and secure. Currently, there are a minimum of four sorts of blockchain networks — public blockchains, private blockchains, consortium blockchains and hybrid blockchains.

Public blockchains: A public blockchain has absolutely no access restrictions. Anyone with an Internet connection can send transactions thereto also as become a validator (i.e., participate within the execution of a consensus protocol) Usually, such networks offer economic incentives for those that secure them and utilize some sort of a symbol of Stake or Proof of Work algorithm. Some of the most important, most known public blockchains are the bitcoin blockchain and therefore the Ethereum blockchain.

Private blockchains: A private blockchain is permissioned. One cannot join it unless and until invited by the network administrators. Participant and validator access is restricted.

Hybrid blockchains: A hybrid blockchain features a combination of centralized and decentralized features the precise workings of the chain can vary supported which portions of centralization decentralization are used.

3. E-DOCUMENTS

Usually, any file or paper is prepared for its official use in its printed form. Originally, any computer data were considered as something internal — the final data output was always on paper. However, the event of computer networks has made it in order that in most cases it's far more convenient to distribute electronic documents than printed ones. This constitutes the hardcopy of a document/paper. But it is possible to visualize the same content on our mobile/computer screens. This kind of document that can be visualized on our device screens are called as E-documents. Because of E-Documents we can have both genuinity of the document as well as save the storage space used for conventional documents and also save paper. Inspite of all these advantages, it still has some drawbacks. However, using electronic documents for final presentation instead of paper has created the problem of multiple incompatible file formats. Even plain text computer files aren't free from this problem. This is because most text editors have no fonts of other languages save English. Also on various platforms, various symbols are used for the representation of the same character. Say, for example the newline character. Even more problems are connected with complex file formats of various word processors, spreadsheets, and graphics software. To counter the situation, many software companies distribute free file viewers for his or her proprietary file formats (one example is Adobe's Acrobat Reader). To eliminate such problems based on different platforms, many companies provide free document viewers to view their documents. The specialized electronic articles in physics use TeX or PostScript formatting languages.

4. METHODOLOGY

A. *System Design*

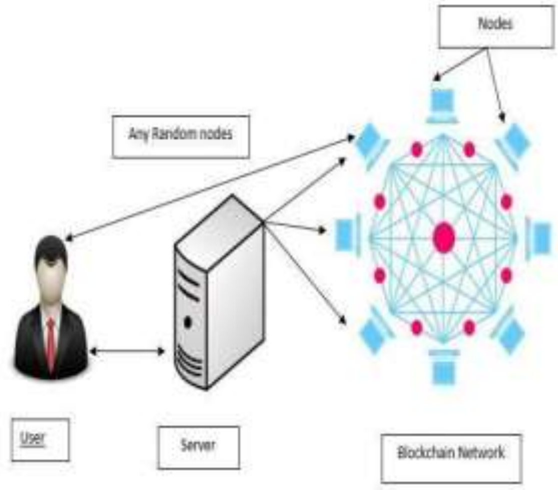
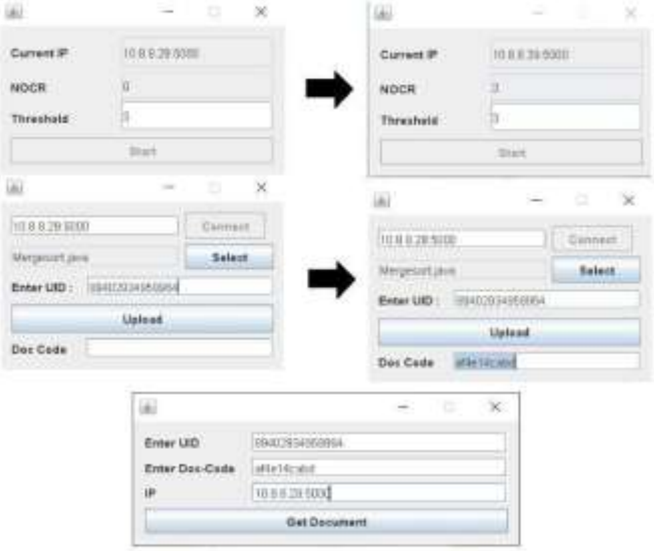


Fig. 1: Proposed System

The proposed system uses immutability property of blockchain to keep the documents in them secure and always available over the internet. Storage of documents is done in the blockchain network.

B. Implementation



Current IP 10.8.8.29:5000
NOCR 0
Threshold 3

Current IP 10.8.8.29:5000
NOCR 3
Threshold 3

10.8.8.29:5000
Mergesort.java
Enter UID : 89402934958954
Doc Code

10.8.8.29:5000
Mergesort.java
Enter UID : 89402934958954
Doc Code af4e14cabd

Enter UID 89402934958954
Enter Doc-Code af4e14cabd
IP 10.8.8.29:5000

Fig. 2: Implementation

Construction of the block-chain network: Initially each and every node in the blockchain system will be separate and independent entity until a connection is made between them. Every node will be having their own blockchain-based database. These blockchain-based databases will be having only the "genesis-block" before the P2P connection is established between them. To establish this connection between all the nodes of the blockchain network, we require a third party hardware, which we have termed here as the bottle-server. This server will be open for the nodes to connect with itself. The bottle server will be having a particular threshold amount up to which it will receive incoming connections. The bottle-server will also make sure that the connecting nodes are authorized nodes through some verification process. As the incoming nodes are being received, their IP addresses are taken into account by the bottle-server. After the threshold is reached, the interconnection between every nodes is done by the bottle-server thus making P2P connections between every nodes. After the P2P connection is established between each pair of nodes, synchronization is done among all of their blockchains for an even reflection of changes made to the blockchain.

Storage of Digital documents: Only the authorized nodes forming the blockchain are allowed to enter/transfer data into the blockchain network. No other computer can do so because they won't be having the privilege of adding data to the blockchain. A soft copy of the original document is made, which after its perfect verification is sent into the blockchain network by the authorized node. This e-copy of certificate will be added into every node's blockchain due to the synchronization between each and every nodes in the blockchain network. After the e-document is sent, the owner of the document is provided with a secret document-code which is formed by SHA-256 hashing algorithm. This document-code can be used later to obtain the e-copy of the document.

Accessing the document: For accessing the document, the user will access the API which is used for data retrieval from the blockchain network. This API will take as input 2 parameters during document uploading:

- Aadhar number (or any unique ID).
- Document-code obtained

This API will first connect to the bottle-server and from the bottle-server it will get the IP address of any random node in the blockchain network. Then an OTP is generated by the bottle-server and this OTP is provided to the API and a particular node in the blockchain network.3. After this the OTP from the API and the node is matched. If both are same the document automatically gets downloaded into the user's machine through the API.

5. RESULT

We have created a situation where we can make e-document storage easy and provide 100 percent availability to the documents anywhere and anytime. Also through verification we have preserved the genuinity of the document and kept it from being tampered by any outsider.

6. CONCLUSIONS

Data security is one of the major features of blockchain technology. Blockchain is a large and open-access online ledger during which each node saves and verifies an equivalent data. This blockchain-based system will provide completely secure storage of the document over the internet by keeping it safe from being tampered by any third person. Also it will provide 100 percent availability of e-document as and when required by the owner of that particular e-document.

7. FUTURE SCOPE

The current system does not have any analyzer to analyze what is written in the document and to whom it belongs. This information can be currently interpreted by a human being. But in future, the nodes can be made smart through NLP and given the ability to understand right from wrong, thus raising a security warning whenever it finds any fault in the currently received document

REFERENCES

[1] Xiuping Lin, "Semi-centralized Blockchain Smart Contracts: Central-ized Verification and Smart Computing under Chains in the Ethereum Blockchain", Department of Information Engineering, National Taiwan University, Taiwan, R.O.C., 2017.

[2] Yong Shi, "Secure storage service of electronic ballot system based on block chain algorithm",Department of Computer Science, Tsing Hua University, Taiwan, R.O.C., 2017.

[3] Zhenzhi Qiu, "Digital certificate for a painting based on blockchain technology", Department of In-formation and Finance Management, National Taipei University of Technology, Taiwan, R.O.C., 2017.

[4] Ajit Kulkarni, (2018), "How To Choose Between Public And Permis-sioned Blockchain For Your Project", Chronicled, 2018.

[5] Jonathan Alexander, Steven Landers and Ben Howerton (2018). Netvote: A Decentralized Voting Network

[6] Salanfe, Setup your own private Proof-of-Authority Ethereum network with Geth, Hacker Noon, 2018

[7] Gong Chen, Development and Application of Smart Contracts, https://www.fisc.com.tw/Upload/b0499306-1905-4531-888a-2bc4c1ddb391/TC/9005.pdf