Immutable and Privacy Protected E-Certificate Repository on Blockchain

Rekha Kashyap, Karan Arora, Asad Azam, Megha Sharma

Abstract—Fake education certificates or fake degree is one of the major concerns in higher education. This fraud can be minimized if there is a tamper-proof and confidential registry of certificates wherein not one but multiple certified authorities verifies and stores the issued certificate in immutable repositories with proper privacy maintained. Secondly, there should be a mechanism for retrieving the authentic certificate without much cost and time. Blockchain is an immutable, shared, distributed ledger without the control of a single centralized authority that fits very well for the discussed use case. The proposed work, PrivateCertChain, has implemented the idea for university having multiple affiliated colleges, by deploying and verifying digitally signed e-certificate on Ethereum Blockchain. Multiple affiliated colleges can serve as the miners for verifying the signature of the issuer. For privacy concerns, the content of the certificate will be hashed and this hashed value will be stored in Blockchain along with the roll number of the certificate holder. Once the transaction hash is generated, it will be converted to QR code. The QR code is shared with the respective owner of the certificate and it will also serve as the credential of the certificate. Thus, anyone having the credential can view the authentic certificate which is kept on the blockchain, by scanning QR through the dedicated application designed for verification. The proposed solution can be a foolproof mechanism against all frauds as it guards for integrity, confidentiality, authenticity, and privacy of educational certificates.

Keywords—Blockchain, Smart Contract, ethereum, Privacy, Confientiality, e-Certifcate, e-Degree, Immutability Introduction

I. INTRODUCTION

In the education sector, certificate plays a crucial role as it is an official tag to determine one's knowledge. Certificate issued to learners is used by stakeholders interested in the evidence of an individual's learning. The certificate determines the suitability of an individual as it serves as proof of past performance in the eyes of probable employers, education institutes or any other stakeholder. Proper confidentiality, integrity, privacy, and authenticity of the certificate are most important for trust in educational certificates. The computer security field has always studied all possibilities to secure these certificates but is often constrained by the cost incurred for the same. The certificate issuing process has developed a lot over time from handwritten to semi-printed certificate, and finally digital certificates with a centralized database, but still, most of the universities operate in their centralized systems behind closed doors. Although centralized databases are maturing to solve security concern, they were centrally controlled needs trust mechanism with a very high cost of security.

Revised Manuscript Received on February 06, 2020.

Dr. Rekha Kashyap, Inderprastha Engineering College, AKTU
Karan Arora, Inderprastha Engineering College, AKTU
Asad Azam, Inderprastha Engineering College, AKTU
Megha Sharma, Inderprastha Engineering College, AKTU

Blockchain is an immutable, secured, shared and distributed, ledger and has facilated the way of recording and tracking resources It has offered a solution against the possibility of attack to centralized systems as there is no centralized authority which saves or controls the data. It is an emerging technology, helpful in protecting resources that can be tangible (e.g., money, houses, cars, lands) or intangible (e.g. copyrights, digital documents, and intellectual property rights) [1]. In the majority of the databases, the aim is to keep the data correct for the current moment behaving as a snapshot of a moment in time. Blockchain databases are able to keep information that is relevant now, but also all the information that has come before. Blockchain technology stores data using cryptographically secured hashing algorithm making it computationally infeasible to modify or compromise the data once stored. This very property has led people to call a blockchain database immutable. The data stored can provide a real-time view while maintaining the ever-growing archives.. Blockchain being a shared database does not directly support the privacy of the records, the existing blockchain framework is used for correctness and availability but not privacy.

A blockchain can be made private by having security measures installed thus ensuring privacy, which like a centralized database; can be write-controlled and read-controlled. That means the protocol can be set up so only private participants can write into the blockchain. The objective of this work is to implement privacy enabled, private blockchain-based e-Certificate repository on Ethereum, that ensures against the issuance of fake certificates and at the same time guarantees the confidentiality of data. It also allows authorized stakeholders to verify the data in a timely and cost-effective manner.

II. RELATED WORK

A blockchain is an immutable and distributed data structure that is replicated and shared among the members of a network. As a technology, it got introduced with Bitcoin, a cryptocurrency that promises a market capitalization of 180 billion dollars as of January 2018 [2][3][4] and offered a solution for the double-spending problem for digital transaction. With the increasing interest in the blockchain technology, many new platforms and applications have been proposed to utilize the advantages offered by this technology. Various papers have been written to highlight the benefits of this technology in banking[5], healthcare[6], real-estate[7] IoT[8], etc.

Owing to growing concerns over use of fraudulent certificates in the education sector we have proposed a privacy-preserving e-Certificate repository using the benefits of blockchain.

Much literature has discussed suitability of blockchain technology for digital certificates or educational certificate repository, but we could not find any implemented work related to e-certificates ensuring confidentiality registered in the literature. Some of the work carried out for similar area is discussed next.

Aisong Zhang and Xinxin Ma[10] have proposed a "Decentralized Digital Certificate Revocation System Based on Blockchain". This system proposed a decentralized digital certificate revocation system with collaborative management of digital certificate revocation lists by multiple CAs and can invalidate the digital certificate in special cases to protect the user's information. Alexander Yakubov, Wazen M. Shbair, Anders Wallbom, David Sanda[11], proposed Blockchain-Based PKI Management, facilitating secure information exchange over the internet eliminating single-point-of-failure and rapid reaction to CAs shortcomings.

They have also developed a PKI management framework based on blockchain for issuing validated certificates with the proper revoking mechanism. Murat Yasin Kubilay, Mehmet Sabır Kiraz and Hacı Ali Mantar [12], proposed Certificate Transparency Based on Blockchain. The work identifies and and computes the hash values in the form of merkle tree and stores the merkle root in the blockchain. The work is different as it discussed the immutability of file storage without any provision for the privacy of file storage. Joon sun, Joon Wu, Joo Han[13], proposed the "Accredited certificate authentication system, and accredited certificate authentication method based on blockchain". The solution mandates having an e-wallet, for the stakeholder wanting to get the certificates verified and also does not talk about the privacy of the certificates. Ahmed Kosba[9], proposed "The blockchain-based model of cryptography and privacy-preserving smart contracts". The goal of this work is similar but their approach is very different, it is based on, on-chain privacy using sealed auction wherein bidders are involved in generating keys for the encryption. This model can't fit in an educational system. Our work, **PrivateCertChain**, proposes an Ethereum blockchain enabling privacy-preserving repository for storing e-certificates with multiple authorities validating the certificates. The verifiers will have the credential in the form of QR code to enable the access to the SHA-256 hashed certificates, through dedicated application designed for same task.

III. BACKGROUND

A. *Ethereum*

Ethereum[14] is a software platform designed to make Distributed Applications and Smart Contracts. It is a public, open-source platform which promises application building without control or interference from third party. It also supports its own crypto currency in the name of Ether. It is Blockchain-based distributed software platform that allows developers to build and deploy decentralized applications. Ethereum uses KEECAK-256 hashing algorithm, to hash the content of the block. The KEECAK-256 algorithm converts an infinitely large file to a 256 bit code based on SHA-3 principles.

B. *Smart Contracts*

The general purpose computation that take place in a blockchain or distributed ledger are being referred as "Smart

Contract" as it is kept as a record of all transactions that neither can be changed nor denied in the future[17]. It can be developed as a digital protocol that enables enforcement of the rules laid in the contract, by all parties involved in the contract.

It also does not require the support of third party for the enforcement or for resolution of disputes. The credibility of all transactions are maintained by the platform. These transactions are track- able and irreversible. The aim of smart contracts is to provide security that is superior to traditional contract law and to reduce other transaction associated with contracting.

C. *Solidity*

Solidity was developed as a contract-oriented programming language with all possible constructs and features to enable for efficient coding of Smart Contracts[15][16]. The core contributors of Solidity are Gavin Wood, Christian, Alex, Yoichi Hirai and few others who developed it for various blockchain platforms including Ethereum.

D. *ECDSA*

In cryptography, the Elliptic Curve Digital Signature Algorithm(ECDSA), offers a variant of the Digital Signature Algorithm (DSA) which uses Elliptic curve cryptography [19]. ECDH is an anonymous key agreement protocol that allows two parties, each having an elliptic curve public-private key pair, for deriving a shared secret and establishing it over an insecure channel. This shared secret may be directly used as

a key, or to derive another key. The key or the derived key can then be used to encrypt subsequent communications using a symmetric key cipher. They are widely used in blockchain technology as they are computationally superior as compared to other crypto algorithms and shorter keys are required for blockchain platform which is supported in ECDSA, without comprising the security.

E. *Merkle Tree*

Merkle tree is a data structure used by several Blockchain.

Every block stores all the transaction data in the form of a Merkle tree. The data structure, hashes the content of child

node and combine it into to form the header of parent node.

This technique of combining the child nodes headers and

adding it to the header of the parent node, continued iteratively till we reach the final node, which is the root node. Thus, the root node will contain information about all of the nodes present in the tree. Thus, any change made in any node will change all its parent nodes and definitely leading to change in the root node.

F. *SHA-2 Hashing*

SHA-256 stands for Secure Hash Algorithm – 256 bit and is a type of hash function commonly used in Blockchain. A hash function is a type of mathematical function which turns data into a fingerprint of that data called a hash.

It's like a formula or algorithm which takes the input data and turns it into an output of a fixed length, which represents the fingerprint of the data.

A hash function will give the same hash for the same input always no matter when, where and how you run the algorithm. Equally interestingly, if even one character in the input text or data is changed, the output hash will change. Also, a hash function is a one-way function, thus it is impossible to generate back the input data from its hash.

IV. PROPOSED WORK

The paper has proposed, **PrivateCertChain**, as an implementation of the blockchain-based e-Certificate repository over Ethereum. The work claims to be an effective solution for the existing problem of fake educational certificates endangering the very credibility of the education system. The very property of Ethereum blockchain converts the certificates into the immutable history of transaction records i.e. provides undoubted authenticity by restricting the tempering of records. To ensure against unauthorized issuance of degree, each degree is digitally signed by the authorized issuing authority and miners will verify the signature before validating the degree and adding it to the chain.

For making the certificates secured against unauthorized access and for preserving the privacy of degree, the content of certificate is hashed using SHA-256, a hashing algorithm and then stored on the blockchain along with the roll number of the respective certificates. The verifiers unlock the degree using the credential shared by the owner/claimant of the degree.

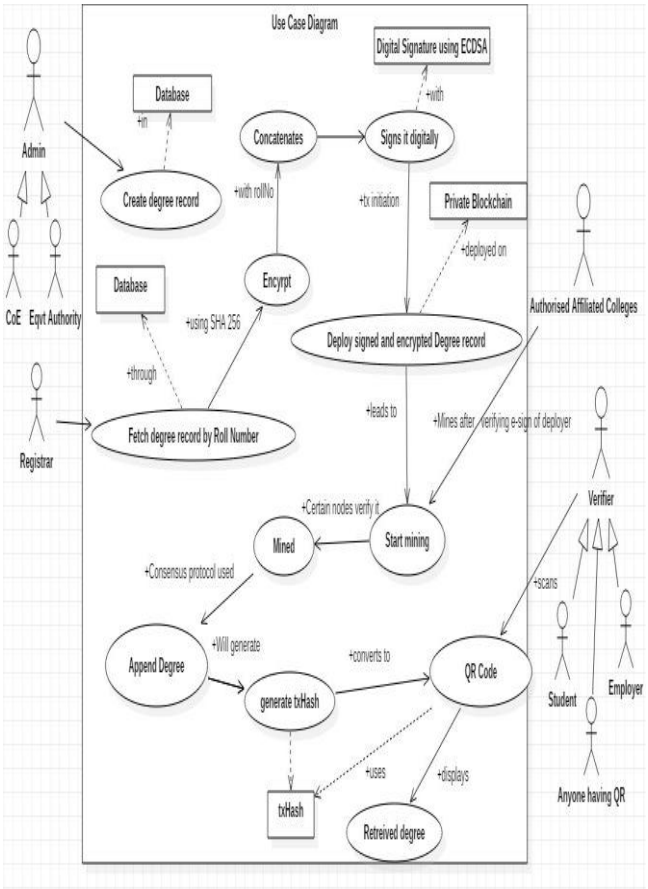

Fig:1(Use Case Diagram)

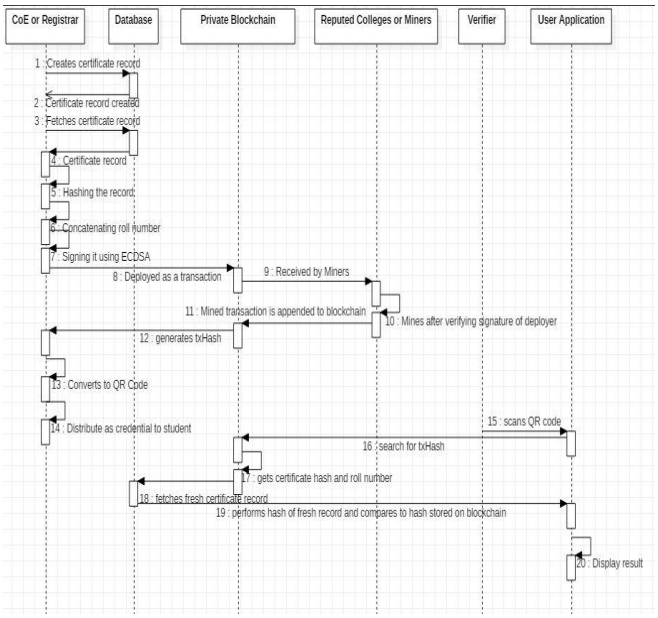

Fig:2(Sequence Diagram)

This process is described in detail in the following section in three phases.

A. *Deployment of e-Certificates*

The authorized Issuing Authority (Controller of Examinations), controls the deployment process. Digitally signed and cryptographically secured e-certificate are created using the following steps.

1) ***Saving the degree in the database:*** The process is similar to the current practice wherin the responsible stake holder, who is Controller of Examinations will save the e-Certificate in the database(refer fig 1 and fig 2) , which is also known as world state.

2) ***Hashing and signing the e-Certificate:*** The content of degree is fetched and hashed using the SHA-256 hashing algorithm. This hash is concatenated with the roll number of the e-certificate holder and signed by Issuing Authority's Digital Signature using ECDSA and private key.

3) ***Deployment of e-Certificate on Blockchain:*** The signed e-certificate record is deployed to the private blockchain. This is the initiation of transaction.

B. *Validation of e-Certificates*

Multiple authorities take part in the validation process. In our project, certified colleges affiliated to the university act as validators and miners as follows:

1) ***Validating Digital Signature of Issuer using ECDSA Algorithm:*** The deployed contract is authenticated by the validators. The process involves verification of digital signature of the initiators(deployer, registrar in our case). If signature verifies to be genuine, the verified transactions are relayed in the network for mining.

2) ***Mining the e-Certificate:*** Miners solve the mining puzzle and relay the block to reach consensus on the Blockchain. If mined with sufficient consensus, the transaction is appended to the Blockchain in the Merkle tree and becomes an immutable transaction.

It then becomes the part of the hashchain and txHash(transaction hash) will be issued(refer fig 1).

3) ***Transaction processing on Blockchain:*** Once the hash of the transaction, txHash is generated, issuing authority converts it into QR code. This QR code is shared with the student and it serves as the credential used to verify and access the certificate. This completes the transaction processing stage.

C. *Verification of e-Certificates*

For verification, an application with a user-friendly interface is developed for verification by stakeholders using the credential shared by the claimant of the certificate using the following steps.

1) Verification of a certificate can be done with ease by just scanning a single QR(credential) through a dedicated application.

2) Claimant(possible owner of the certificate), in need of verification, shares the credential with the verifier. It only allows the stakeholder to view and verify the certificate.

3) Any stakeholder who wishes to verify the authenticity of the certificate will scan the QR code. QR code is having txHash embedded in it.

4) txHash will fetch the SHA-256 hash and its corresponding roll number from the blockchain. The record corresponding to the fetched roll number is fetched now. This record is hashed with SHA-256.

5) On successful match of the two hashes i.e., the hash stored in Blockchain and the fresh hash computed from the record stored in database matches, it can be claimed that the certificate is valid and not tampered/altered and complete certificate record is shown to the verifier. If hashes do not match, it can be claimed that the certificate is fake or altered/tampered on the database.

6) The privacy and confidentiality are ensured as no one can access the certificate without the consent of the certificate holder. Anyone can have access to ledger or repository, but all the information will be in hashed form, ensuring privacy and confidentiality(refer fig 2).

V. EXPERIMENT AND RESULT ANALYSIS

We tested our implementation by deploying it on Ethereum's official test network: Ropsten. We deployed 100 transactions to issue certificates of 100 students. Average of each certificate deployment as transaction's computational and gas cost is provided in Table 1. Deployment of each certificate is a transaction and is broadcasted only once by the Issuing Authority (denoted by prefix 'R' in table). Transaction by a verifier is broadcasted once per student per verification, i.e., a total of 100 times (for 100 students, verification is carried once for each student). Since it is tested on Ropsten, no real money was spent in form of Ethers.

Entity: Transaction	Gas Used(100 transactions)
R: Issue Certificate	4171000
V: Verify Certificate	1020000

Table 1. Costs in gas on Ropsten test network.

Action	Average time in seconds
Registrar: Issuing Certificate	2
Miners: Mining the Certificate	120-180
Ledger: Block Confirmation	480-600
Verifier: Verifying Certificate	10

Table 2. Computational time analysis for various stakeholders.

VI. CONCLUSION

The proposed idea to prevent certificate fraud by storing e-Certificate records on Blockchain is to make the certificate immutable, private and authenticated in a decentralized manner by multiple authorities. It also opens the door for the transparent working of universities and creates an ecosystem where stakeholders can participate in the authentication process, thus creating immense trust in the educational Certificates. In this work, we have proposed and implemented, PrivateCertChain, which is an e-Certificate deployment system that uses Smart Contracts and hashing to enable and secure, cost-efficient certificate deployment while guarantying its security, privacy, and confidentiality of all certificates. With Ethereum private blockchain, we can send hundreds of transactions per second onto the blockchain, using every aspect of the smart contract to distribute the load on the blockchain. For Universities with more students, measures will be taken to withhold greater throughput of transactions per second.

REFERENCES

1. M. Pilkington, "Blockchain Technology: Principles and Applications," in Research Handbook on Digital Transformations, 2016. [online]Available: https://ssrn.com/abstract=2662660, (accessed on February 13,2019)
2. I. Eyal, A. E. Gencer, E. G. Sirer, and R. V. Renesse, "Bitcoin-NG: a scalable blockchain protocol," in 13th Usenix Conference on Networked Systems Design and Implementation (NSDI'16), Berkeley, CA, USA,2016, pp. 45-59
3. CoinMarketCap.Com, "Crypto Currency Market Capitalization," [online] Available: https://coinmarketcap.com/currencies/, (accessed August 15, 2018)
4. S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," 2009.[Online] Available: http://www.bitcoin.org/bitcoin.pdf ,(accessed February 13, 2019)
5. Universa, (2016), Blockchain is Reshaping the Banking Sector. [Online]. Available: https://medium.com/universablockchain/blockchain-is-reshaping-the-banking-sector-fd84f2fc475
6. I. Kar. (2016). Estonian Citizens Will Soon Have the World's Most Hack-Proof Health-Care Records. [Online]. Available: http://qz.com/628889/this-eastern-european-country-is-moving-its-health-recordsto-the-blockchain
7. D. Oparah. (2016). 3 Ways That the Blockchain Will Change the Real Estate Market. [Online]. Available: http://techcrunch.com/2016/02/06/3-ways-that-blockchain-will-change-the-real-estate-market
8. Prasad Joshi(2017). The Blockchain of Things: Why it is a major game changer for Internet of Things. Available: http://www.forbesindia.com/blog/health/the-blockchain-of-things-why-it-is-a-major-game-changer-for-internet-of-things
9. A.E. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, Hawk: The blockchain model of cryptography and privacy-preserving smart contracts, in IEEE Symposium on Security and Privacy, SP 2016, San Jose, CA, USA, May 22–26,2016.IEEE Computer Society,2016,pp.839-858.

10. Aisong Zhang and Xinxin Ma.(2018) Decentralized Digital Certificate Revocation System Based on Blockchain. Available: https://www.researchgate.net/publication/327325392_Decentralized_Digital_Certificate_Revocation_System_Based_on_Blockchain
11. Alexander Yakubov, Wazen M. Shbair, Anders Wallbom, David Sanda.(2018). Blockchain-Based PKI Management. Available: https://orbilu.uni.lu/bitstream/10993/35468/1/ blockchain-based-pki.pd
12. Murat Yasin Kubilay, Mehmet Sabır Kiraz and Hacı Ali Mantar.(2018).[online]Available: A New PKI Model with Certificate Transparency Based on Blockchain
13. Joon sun,Joon Wu, JooHan.(2018). Accredited certificate authentication system. Available: https://patentimages.storage.googleapis.com/09/1e/63/8d2501530060bb/WO2018008800A1.pdf
14. Vitalik Buterin Ethereum white paper made simple. Available: https://blockchainreview.io/wpcontent/uploads/2013/02.01._final_Ethereum-White-Paper-Made-Simple.pdf
15. M.Rouse.(2018).Solidity.[Online]Available: https://whatis.techtarget.com/definition/Solidity
16. Allison, Ian.(2016).[Online]Available: "PwC blockchain expert pinpoints sources of ambiguity in smart contracts".
17. Nick Szabo. Use Cases for Business & Beyond. Available: https://www.perkinscoie.com/images/content/1/6/v2/164979/Smart-Contracts-12-Use-Cases-for-Business-Beyond.pdf
18. National Institute of Standards and Technology.(2001).AES Encryption and Related Concepts[online]Available: https://townsendsecurity.com/ites/default/files/AES_Introduction.pdf
19. Scott Vanstone, The Elliptic CurveDigital signature Algorithm, in 2013 IEEE International Conference on Machine Intelligence and Research Advancement, Katra, India 21-23 Dec. 2013
20. D Emmons.(2018)White paper and Demo: UX for Authenticated & Verified ERC20 Payments Using MetaMask and EthSigUtil [online] Available: https://medium.com/coinmonks/whitepaper-and-demo-ux-for-authenticated-verified-erc20-payments-using-metamask-and-ethsigutil-7a146afcd65e