UNIVERSITY POLITEHNICA OF BUCHAREST
FACULTY OF AUTOMATIC CONTROL AND COMPUTERS
COMPUTER SCIENCE DEPARTMENT

RESEARCH REPORT

Blockchain platform for issuing and verifying Higher Education diplomas

Student:
Dorin-Liviu Heroi

Thesis advisor:
Alexandra Cernian

TABLE OF CONTENT

TABLE OF FIGURES

1 INTRODUCTION

Blockhain has been all over the news in the last two years and it has gotten the attention of multiple developers and researchers, bringing it to multiple fields. From a historic point of view, it has been brought to the public's attention with the bitcoin cryptocurrency, by a person going by the name of Satoshi Nakamoto. The blockchain is a peer-to-peer network, used as a distributed ledger and, as the name states, is represented by a chain of blocks, where blocks are batches of transactions that, in popular blockchains, are hashed and encoded into a structure named Merkle tree. Transactions can represent multiple things, from currency exchange, to specific function calls. These blocks are linked together based on cryptography principles. The main advantages of the blockchain are its distributed nature and immutability. The distributed nature assures that if one point of contact fails, the system still holds and there is not only one single point of failure, as seen in centralized systems. Its immutability represents the fact that once a block has been created and added to the chain, it cannot be modified.

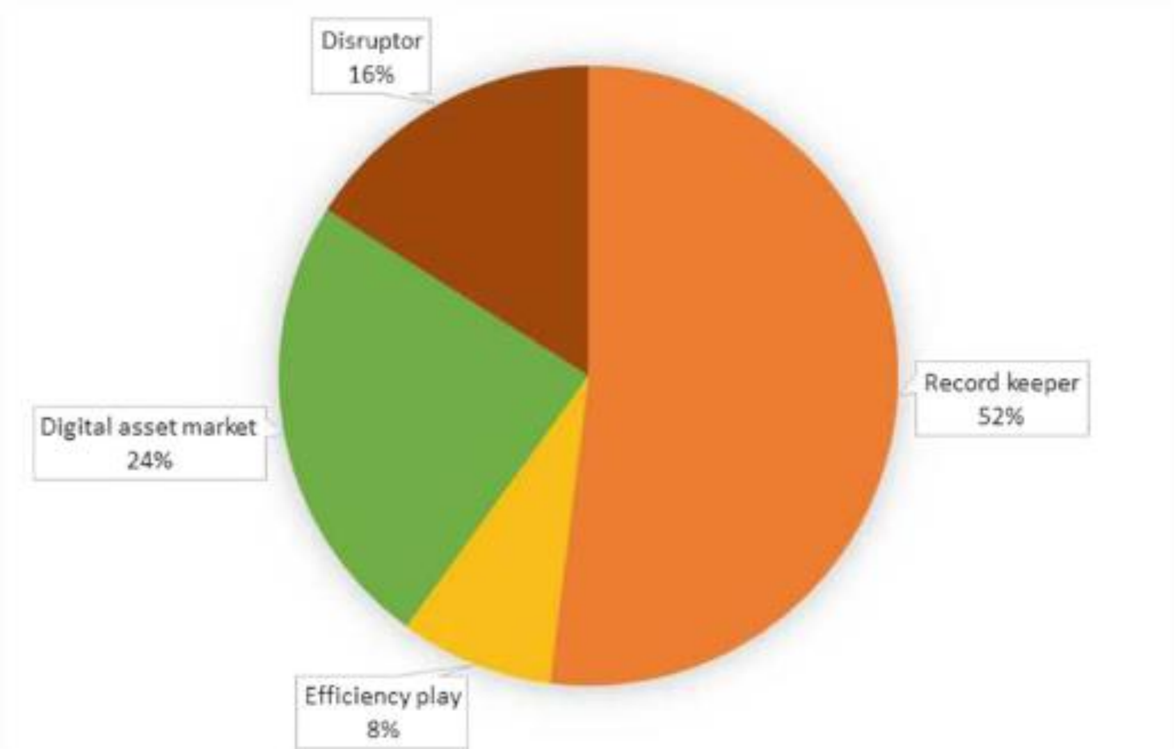
Another important aspect is smart contracts, which is a method of automatic processing of a transaction. This means that it can control the transaction according to the terms of the contract, such as making certaing payments without the need of a middleman, make custom verifications on transactions. Smart contracts are an extremely powerful tool, further reducing the level centralization and human interaction and verification. This is in most part due to the fact that their execution is controlled and cannot be interfered with.

Given these advantages, we can picture the use of blockchain in a different context than the one that is has been usually used in, which is finance. We can use it, accompanied by a distributed system of storing data, to store diplomas issued by HEI(higher education institutions) and make it so that a third party can verify the validity of the respective document. This also facilitates both the institution's issuing process, making it much less paper centric, meaning less costs and quicker deployment times, and the students' access to their documents. The students' privacy is also taken care of as the verification can be done only after the student consents to hand their diploma and its hash.

2 MOTIVATION AND CONCEPTS

Given that blockchain is a rising technology, research has been made over the fields it has propagated to. Given the discussed topic, we can observe in more detail in [1], that the amount of articles regarding both diplomas and blockchain is scarce. This signifies that there is not yet much research put into this usage of the blockchain, which can mean both that the rules followed when developing such project are less rigid and a greater degree of freedom is allowed, testing and trying things that had not been tried beofre, but also the small number or even the inexistence of standards or good practices, which may result in a long time spent trying to come up with something new that has a certain basis.

Even though there are not many, in some papers, such as [2], there is also specified the openess that students show when speaking about the adoption of such system. This early market research strenghtens the motivation of the project and justifies that its development would not be in vain.



1 The share of applications within the educational domain classified into four blockchain initiatives, [3]

As this pie chart shows, efficiency play (poorly named – but basically document verification) comprises only a small percent which means that implementations in this area are welcome.

2.1 Types of blockchain implementations

When talking about blockchain, we can divide it in two: public and private.

The public part is the part most people have heard of and are familiar with. In this category, we can confidently say that the main participants are Bitcoin and Ethereum. If we were to put some main, brief labels on these participants, we would label them as transparent, due to the fact that the transactions made are available to everyone that has access to the chain of transactions and taxable, as transactions that are taking place need to also be accompanied by a fee. This fee is put in place for multiple reasons. One of the reasons is that due to the fact that transactions are proven by workers performing proof-of-work, they have to be incetivised and they are paid with currency from these fees. Note that the implementation may differ from one blockchain to another. Another reason is to prevent malicious intent of denial of service by spamming with transactions. It is important to mention that transparency is not limited to public blockchains, as there are also public blockchains, such as Monero, that do not have transparency in transactions.

The private part is more enterprise oriented and it is not as popular its public counterpart, but also presents some advantages over it. One of these private implementations is Hyperledger Fabric and is mainly reffered to as a „permissioned blockchain”, due to its architecture that allows only specific participants to take part in the blockchain. Also, there can be allowed only certain permissions to specific participants.

In this project, it was decided that the use of a private blockchain would be more beneficial, from multiple points of view, given throughout the document. The idea behind it is that only specific institutions, the issuers, can take part in the blockchain.

2.2 Hyperledger Fabric

We can observe in the table presented in [1] that there is a note regarding the use of the private blockchain Hyperledger, specifing that the majority of blockchains that are used in such projects are public blockchains. Given the fact that the project intends to use this type of blockchain implementation, its scarce use may bring more importance to the project by discovering new challenges in such development, but also new opportunities and advantages as oposed to a public blockchain. One obvious advantage is the fact that there is no cost associated to any type of process on Hyperledger, as opposed to a transaction on a public blockchain. This means low costs, while also preserving the advantages of a blockchain, such as decentralization, but one of the main advantages when choosing this solution is the permissioned architecture of the blockchain. Unlike public blockchains, where you take part in a open network of anonymous participants, here you can establish trust in a network where participants are known. Given the requirements of privacy that are omnipresent regarding students, another advantage is that there can be made both public and confidential transactions, to suit certain needs at specific moments.

An important ideology that this project is based on and takes advantage of a Hyperledger benefit is that the world of blockchain is new and is in constant development. Every detail of it can change for the better so the more customizable and modular it is, the better. While code on a popular private blockchain, such as Ethereum is written in Solidity, a language specific to

Ethereum, Hyperledger is more accesible, being open for customization in popular languages such as Python or Java. This means that if someone wants to make an addition or an adjustment, it is easily possible, without first learning and understanding Solidity.

3 IMPLEMENTATION

Talking about storing accomplishments, there are multiple implementations present. Out of them, we remind of BADGR and Mozilla Open Badges. Their purpose is to store accomplishments in such a way that they are easily recognizable and proven to be true. This facilitates the task of proving skills to an employer or an educational institution. Based on this idea, there have also been blockchain interpretations and implementations. In [4] it is spoken about EduCTX's idea of storing a type of information based on ECTS(European Credit Transfer System). The idea would be that a student would have a wallet where he can have stored credits of all the courses he has attended to and also certifications he has obtained.

From this project's perspective, EduCTX's inclusion of different courses is unnecessary, so here only the problem of storing diplomas is addressed.

3.1 Requirements

There have been identified multiple requirements that this project needs to accomplish to be taken into consideration in a real-life scenario. Some have been selected from requirements available in [5], due to their relevance.

1. None but authorized departments from the accepted institutions are allowed to issue diplomas
2. Diplomas should be confidential on chain
3. Processes of issuance and verification should not imply any technical skill
4. The possibility of revoking issued documents
5. Diplomas should be issued in a digital format

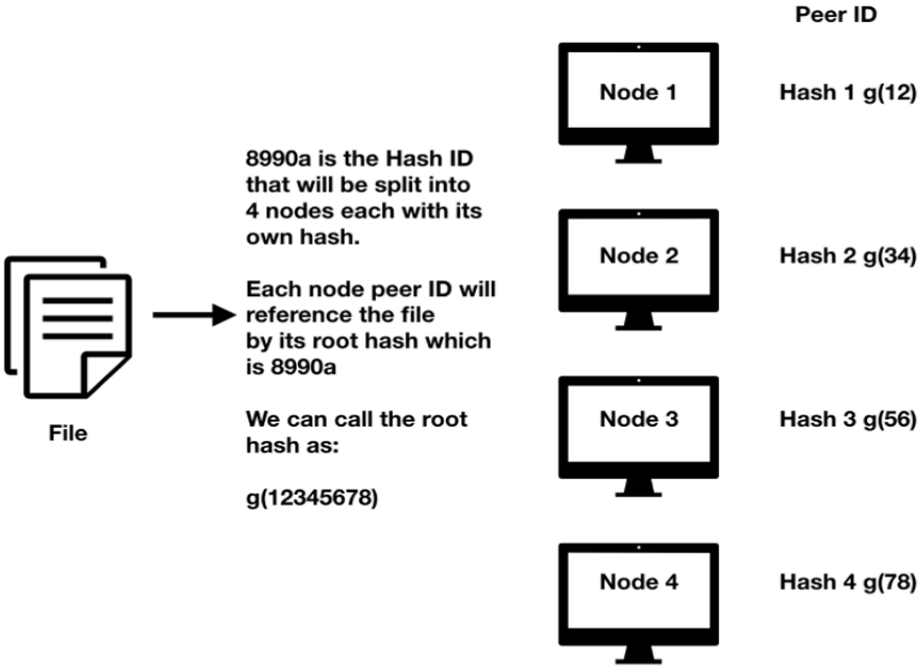
The first point implies that even though there may be multiple approved entities taking part in the network, only some of them, the ones that have specific rights should be allowed to issue diplomas for students. If any other member signs documents, the transaction should be rejected. The second point sheds light on the needed privacy that should be present in the project. This is achieved by storing only the hash on the blockchain and storing the diploma in an encrypted format on the IPFS network. The third point requests that the main parts of the project should be accessible to people that do not have high technical skills. The processes of issuing and verification should be as straight forward as possible. The fourth points present the possibility of diploma revocation, which is discussed later on. This too should be an easy process to execute. The fifth and last requirement states that the diploma should be issued in a digital format, which is obvious, because the entire project is in a digital format.

3.2 Revocation of documents

In some situations, given specific misconduct or even human mistakes made when issuing the diploma, an institution may be required to revoke an issued diploma. While the blockchain is immutable, there have been multiple tries to implement such system. Vidal, Gouveia and Soares show in [6] some bitcoin specific implementations using UTXO (unspent transaction output) but also Blockcerts's approach, that uses a certificate revocation list, which centralizes the solution, therefore making it not useable, given the project's requirements of decentralization. With decentralization in mind and also the fact that it should be usable on multiple types of blockchain, they propose another approach, that uses UTXO, such as the first approach. The takeaway from this approach though, is the inclusion in such solution of a JSON explaining the reason and date it has been cancelled. [7] also talks about a method to prevent revocation spam, but it is not applicable in this case, as the blockchain is permissioned.

3.3 Storing documents

While the processes of issuing and verification may come together with the concept of blockchain, storing files on such blockchain is not the best approach. One alternative is using IPFS (InterPlanetary File System). IPFS is a peer-to-peer decentralized network that lets users back up files and websites by hosting them across numerous nodes.


Peer ID
Node 1
Hash 1 g(12)
8990a is the Hash ID that will be split into 4 nodes each with its own hash.
Node 2
Hash 2 g(34)
Each node peer ID will reference the file by its root hash which is 8990a
File
We can call the root hash as:
g(12345678)
Node 3
Hash 3 g(56)
Node 4
Hash 4 g(78)

2 IPFS storage mode *[8]*

One important aspect is that one student can have multiple diplomas. From this, results the fact that a list of all the files' hashes is also need so that it is known what to search for. This list of files will be available in every student's wallet, so they can easily retrieve the needed documents.

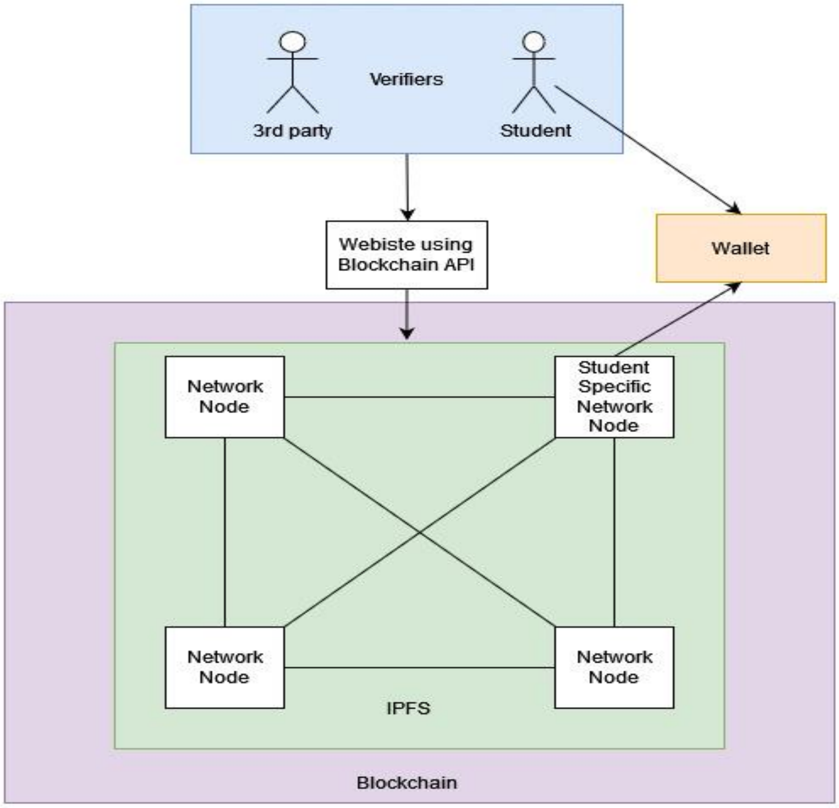
3.4 IPFS Privacy

Storing diplomas on a distributed systems like this one may represent a privacy issue, as anyone could access your diploma. The proposed solution is that a public-private key is generated by every student. Their public keys are requested and would be used by the institution. The diplomas would be encrypted with their public key and they could retrieve the diploma from the IPFS network and decrypt it with their own key.

4 FINAL PROPOSED ARCHITECTURE

All things considered, a final proposed architecture is presented below and should be able to support a workflow like so:

An institution issues diploma in digital format for every student and stores it as a .pdf file. Then, students create a public-private key pair and send the public key to the institution, while safely storing the private key. The institution uses the API, which hashes diploma, then uploads the hash on the Hyperledger blockchain. The API also encrypts diploma with the specific student's public key, then uploads it on the IPFS network. The returned hash ID is then stored to the student's wallet. From this point on, if a student has access to his private key, he also has access to his wallet, where he can get the hash IDs of the documents that are issued to him. With those hashes, he can request to download the encrypted diploma from the IPFS network. The encrypted diploma can be then decrypted with his private key and hashed. Lastly, the hash can be validated on the blockchain. This last step can be made by anyone who possesses the hash and who knows what it represents. A student can also choose to share the diploma with a third party so the third party can check by itself.



3 Final architecture

5 DEVELOPMENT PLAN FOR NEXT SEMESTERS

As the first semester has reached it end, research has been made and the problem and one solution to solve it has been defined. In the following semesters, the proposed development path is proposed like so:

- Familiarize and develop the Hyperledger fabric blockchain with the afore specified requirements
- Familiarize and develop the IPFS network for diploma storage with the afore specified requirements
- Create an easy method with which students can generate their wallets
- Create and polish the API so that the interaction between an application (such as a website) and the blockchain and IPFS network is facilitated

As the project is only at the beginning and no development work has been done yet, estimating a time period in which these steps would be done would be difficult. Even so, a rough estimate can be given and that is that the development would be made in the following two semesters. In the last semester, there will be polishing of the overall project and final tests.

BIBLIOGRAPHY

[1] M. A.-Y.-O. Renato Q. Castro, "Blockchain and Higher Education Diplomas," *Eur. J. Investig. Health Psychol. Educ.,* pp. 154-167, 2021.

[2] N. L. G. F. P. Guendalina Capece, "Blockchain Technology: Redefining Trust for Digital Certificates," *Sustainability,* 2020.

[3] M. T. M. H. S. M. Aida Kamisalic, "A Preliminary Review of Blockchain-Based Solutions in Higher Education," in *Learning Technology for Education Challenges*, 2019, pp. 114-124.

[4] M. H. K. K. M. H. Muhamed Turkanović, "EduCTX: A Blockchain-Based Higher Education Credit Platform," *IEEE Access,* 2017.

[5] B. R. E. S. S. S. K. B. S. Jerinas Gresch, "The Proposal of a Blockchain-Based Architecture for Transparent Certificate Handling," *Lecture Notes in Business Information Processing,* vol. 339, pp. 185-196, 2019.

[6] F. G. C. S. Fernando Richter Vidal, "Revocation Mechanisms for Academic Certificates Stored on a Blockchain," in *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, Seville, Spain, 2020.

[7] WebOfTrustInfo, "A Decentralized Approach to Blockcerts Credential Revocation," 2017. [Online]. Available: https://github.com/WebOfTrustInfo/rwot5-boston/blob/master/final-documents/blockcerts-revocation.md. [Accessed 27 1 2022].

[8] V. Tabora, "Using IPFS For Distributed File Storage Systems," 22 6 2020. [Online]. Available: https://medium.com/0xcode/using-ipfs-for-distributed-file-storage-systems-61226e07a6f. [Accessed 29 1 2022].