A Validation System for Academic Records based on Blockchain

Gabriel Fernández-Blanco, and Tiago M. Fernández-Caramés

GTEC, Faculty of Computer Science, Universidade da Coruña, 15071 A Coruña, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
Correspondence: g.fblanco@udc.es

DOI: *https://doi.org/10.17979/spudc.000024.17*

Abstract: Academic certificate forgeries and the lack of resources to tackle them is a problem that implies big costs to society. It harms the figure of the certificates and impairs the trust of academic institutions. The solution proposed in this paper is aimed at recording and verifying the academic merits through a decentralized application or Dapp. Such an application is supported by a smart contract deployed in the Ethereum blockchain with a simple frontend or interface. The application makes use of a decentralized storing system based on IPFS in order to consume the data that are not managed by the blockchain. To assess the performance of the developed system, the latency of its most common operation (read operations) is measured as the number of requests per second increase. The obtained results show that the system is really fast, being some nodes able to respond to more than 1,000 requests in less than one millisecond.

1 Introduction

The falsification of academic degrees continues to be a widespread and recurring problem that affects even the most developed countries. Traditional technological solutions suffer from vulnerabilities inherent to their design, such as the presence of centralized elements in their architecture (introducing single points of failure) or their incapability to detect unauthorized modifications. Although there are many solutions to verify the authenticity of physical or digital documents (A. Rustemi et al., 2023), they are not aimed at preventing frauds in the academic field. For instance, there are cases in which a student may have passed some exams or subjects due to a favorable treatment, which is a type of fraud committed even by political leaders.

To avoid such a kind of situations, the application proposed in this paper provides a solution for preventing academic title fraud by fully applying decentralized and distributed registry technologies. Thanks to the security and privacy provided by blockchain networks such as Ethereum, academic files become resistant to falsification or fraudulent modifications, since any change is recorded permanently and can be validated by each blockchain node.

Moreover, the proposed application stores the student records on a blockchain in a secure way. Such files consist of the marks received during a degree for the different subjects (or, ideally, for all the performed activities). There is a subset of administrators that are responsible for creating the transactions that update the blockchain. At any time, students can download and send their academic file to any interested entity, which can be easily verified through the developed application.

The proposed system is similar to the one proposed by the University of Zurich (J. Gresch et al., 2018). In addition, there are few solutions to this problem that fully decentralize its operation (A. Rustemi et al., 2023) A. Pfefferling and Kehling (2021). Working entirely on a decentralized architecture guarantees persistence on the academic information that is supported by the nodes of a P2P (Peer-to-Peer) network, which avoids inappropriate modifications to the files or reliance on central servers (T. M. Fernández-Caramés and Fraga-Lamas, 2023). To provide such a feature, the system presented in this paper makes use of a database based on InterPlanetary File-System (IPFS), a purely decentralized protocol.

2 Design and Implementation

The proposed application consists of the subsystems shown in Figure 1, which are the following:

- **Blockchain**: record data are represented by their hash code. Such a hash is stored in the blockchain, supported by the rest of the data that are stored outside the blockchain (in the decentralized database). With this strategy, the size of the information stored in the blockchain is significantly reduced, but requires keeping the raw information outside the chain (off-chain). This increases transaction speed and cost savings, since operational costs decrease (i.e., the gas spent in Ethereum-based applications).
- **Decentralized Database (OrbitDB)**: This database stores the private information of teachers/professors and students (i.e., each academic record, AR), as well as their public keys. The application interacts with this database in order to login students and administrators, as well as for detecting new changes in the ARs. The information stored in OrbitDB is transparent by default, so developers have to take privacy-protection actions to protect users' data privacy.

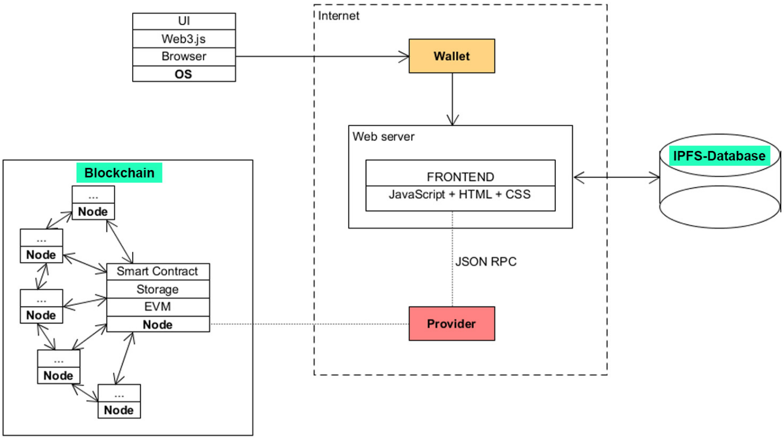

Figure 1: High-level view of the different subsystems.

The proposed application consists of the following actors (whose main interactions are shown in Figure 2):

- **Students**. Once registered, each student is able to see the constant evolution on his AR through the periodic exams and courses. Whenever he/she wants, he/she can download his/her AR (typically in a PDF file) and share it with any third-party, like an organization interested in hiring the student and in validating the student academic merits.
- **Teachers/Professors/Administrators**. They can be seen as permissioned teachers/professors, who are responsible for updating the ARs. For example, after the exams, teachers/professors will register the achieved marks of the students in the decentralized database. After such changes, the validity of these modifications needs to be approved by an administrator (e.g., by the head of the department or a person of the university/school administration). These updated records are introduced in a 'pending record

list' in which are stored all the new ARs and their changes. After confirming these last changes, the validator (i.e., the person that approves the registered changes) will perform a transaction on the blockchain to update the ARs. This update can be performed individually for each AR or in batch (i.e., for several ARs at the same time) to accelerate the process.

- **Third-party**. Any external user can verify the validity of an AR. This is simply performed by introducing the provided AR in the app, which will show almost instantly whether the introduced AR is part of the blockchain or not (by comparing its hash with all of the hashes stored in the blockchain).

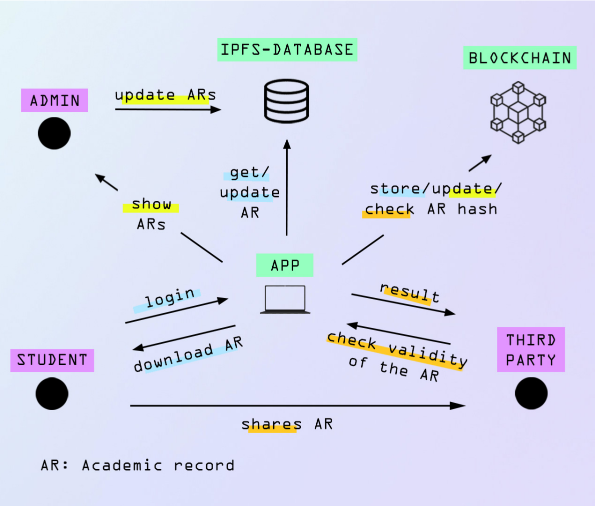

Figure 2: User interaction.

3 Results

Reading data from the blockchain is one of the most common operations in blockchain applications. Hence, it is important to see if the system maintains acceptable response times with high loads of read-operations. For this matter, a novel tool [3] was used to 'flood' the developed Ethereum-based application with read-only JSON-RPC requests such as `eth_call` (it executes a new message call immediately without creating a transaction on the blockchain) or `eth_getBalance` (it returns the balance of a specific account address).

Figure 3 illustrates the results of an `eth_call` test, measuring the evolution of the latency (the time it takes to receive the response from the call) as the number of requests per second (rps) grows. A customized network was created as a test environment with 12 nodes, 4 of them acting as validators/miners. However, note that node quantity is irrelevant when reading data from the blockchain, as transactions are actually not emitted.

As it can be observed in Figure 3, for the worst case evaluated (i.e., for a maximum of 2,000 requests per second), latency is really low (in the order of milliseconds). Obviously, higher loads above 2,000 rps will derive into a relevant increase in latency, but such loads are not usual in a typical academic data verification environment. Thus, the proposed provides all the advantages of decentralized systems in terms of security and trust with an equally acceptable operating time.

It is worth noting that Figure 3 shows a U-shape for each node curve, having more latency at 10 rps than for 100 rps. This is essentially due to the inner workings of the client cache and the 'warming-up' effect of the nodes. However, from 100 rps the nodes behaves as expected, having a smooth increase of the latency as more rps are being executed. Since such an initial behaviour is inherent to the selected tool, it has been preserved in the Figure so that future researchers understand the obtained outputs.

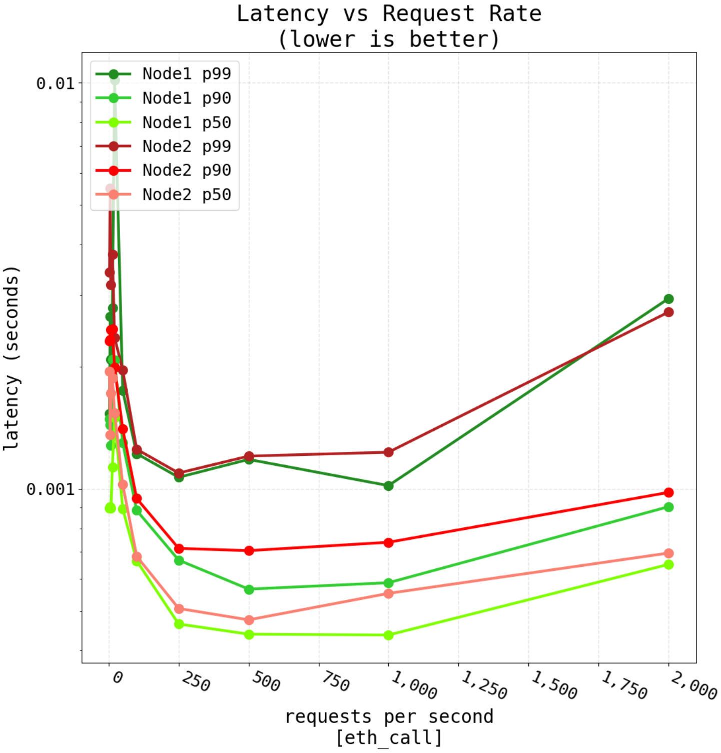

Figure 3: Latency results for the tested blockchain network.

4 Future work

The obtained results allow for concluding that the developed system provides a robust solution through smart contracts that reveals a great potential. For instance, the system can be used for automating a large number bureaucratic procedures such as title issuance or payments. Also, it could bring other new types of learning, as pay-as-you-go classes that employ smart contracts. Even, it could allow organizations to have greater collaborations with other institutions, so the verifiable records could be recognized as complete Curriculum Vitaes.

However, this technology is not without its challenges. It requires a careful design by specialised professionals, as its peculiarities involve new security vulnerabilities. Moreover, it is up to see its potential adoption, and how it would complaint with current data-protection laws as GDPR (General Data Protection Regulation) (i.e., the 'right to be forgotten' it is still hard to address due to its immutable nature). Nevertheless, there are more and more approaches to solve these problems such as the use of cryptography to separate the user's data from its physical-sensitive information.

In addition, the system would need to be improved to increase its energy-efficiency, since ecology is one of the foremost social debates in developed countries. Since the Bitcoin energy consumption has been widely criticized, there exists more interest in the literature that includes strategies and techniques to create green blockchains and to reduce blockchain energy consumption (N. Lei et al., 2021). Therefore, it is possible to use consensus protocols that improve massively electrical consumption over Proof-of-Work such as Proof-of-Authority. Also, it is necessary to study the feasibility of making use of an edge/fog-computing network of low-consuming devices (M. Sánchez-de la Rosa et al., 2023) in the developed Ethereum-based network, which may further decrease energy costs.

5 Conclusion

This paper has presented a decentralized tamper-proof academic record validation system that makes use of a blockchain (Ethereum) and a decentralized storage system (IPFS). The performance of the system has been tested and the obtained results show that the latency of the system is suitable for most of the potential applications that want to secure data or prevent fraud, while harnessing the benefits of decentralized technology (i.e., data privacy, security and immutability).

Acknowledgment

This work has been funded by CITIC that is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS), by the Xunta de Galicia (by grant ED431C 2020/15), and by grants PID2020-118857RA-100 (ORBALLO) and TED2021-129433A-C22 (HELENE) funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR.

Bibliography

A. Pfefferling and P. Kehling. Design disclosure for blockchain-based application used in public education certificates with electronic hashes. *Hochschule Mittweida*, 2021.

A. Rustemi, F. Dalipi, V. Atanasovski, and A. Risteski. A systematic literature review on blockchain-based systems for academic certificate verification. *IEEE Access*, 11:64679–64696, 2023.

Flood:Github-Repository. Load testing tool for benchmarking EVM nodes over RPC, 2023. URL *https://github.com/paradigmxyz/flood*. [Online; accessed 4-September-2023].

J. Gresch, B. Rodrigues, E. John-Scheid, S. Kanhere, and B. Stiller. The proposal of a blockchain-based architecture for transparent certificate handling. *Springer International Publishing*, pages 185–196, 2018.

M. Sánchez-de la Rosa, C. Núñez-Gómez, M. Caminero, and C. Carrión. Exploring the use of blockchain in resource-constrained fog computing environments. *Software: Practice and Experience*, 53:971–987, 2023.

N. Lei, E. Masanet, and J. Koomey. Best practices for analyzing the direct energy use of blockchain technology systems: Review and policy recommendations. *Energy Policy*, 156: 112422, 2021.

T. M. Fernández-Caramés and P. Fraga-Lamas. Design of a fog computing, blockchain and IoT-based continuous glucose monitoring system for crowdsourcing mHealth. *Multidisciplinary Digital Publishing Institute Proceedings*, 4:5757, 2023.