Towards Using Blockchain Technology to Prevent Diploma Fraud

Qiang Tang

Luxembourg Institute of Science and Technology (LIST)
4362, Esch sur Alzette, Luxembourg
qiang.tang@list.lu

Abstract. After its debut with Bitcoin in 2009, Blockchain has attracted enormous attention and been used in many different applications as a trusted black box. Many applications focus on exploiting the Blockchain-native features (e.g. trust from consensus, and smart contracts) while paying less attention to the application-specific requirements. In this paper, we initiate a systematic study on the applications in the education and training sector, where Blockchain is leveraged to combat diploma fraud. We present a general system structure for digitized diploma management systems and identify both functional and non-functional requirements. Our analysis show that all existing Blockchain-based systems fall short in meeting these requirements. Inspired by the analysis, we propose a Blockchain-facilitated solution by leveraging some basic cryptographic primitives and data structures. Following-up analysis show that our solution respects all the identified requirements very well.

1 Introduction

In the education sector, diploma forgery is a long-standing problem [4]. With paper-based diplomas, forgery costs a little and it has been a prevalent business in many developing countries. Even in the developed countries where reputation is highly emphasized, diploma fraud is also not rare. It was reported in 2007 [14] that, Marilee Jones, the dean of admissions at the MIT, had fabricated her undergraduate degree and was forced to resign after working nearly three decades at the institute. The situation becomes better when Internet becomes ubiquitous and information sharing is made easy. However, until today, it remains difficult to validate diplomas. One of the major obstacles is the lacking of a universal platform between diploma issuers and diploma verifiers. In parallel to academic diplomas, fraud in industrial qualifications and working experiences is also a problem in the same vein.

To tackle the diploma forgery problem, many diploma verification solutions have been proposed and implemented. Most of them are institution-based, namely an institution offers an online system for diploma verifiers to validate users' diplomas. Exceptionally, BADGR [5] and Open Badges [13] offer unified solutions for managing users' entire educational history by collecting the digital certifications acquired by them at different academic institutes. Despite all the

efforts, these solutions have not been widely used, and diploma fraud stays as a serious problem today.

1.1 Blockchain in a Nutshell

Following Nakamoto's seminal work [15], the concept of Blockchain has become very popular in the society as it rapidly becomes the key enabling technology for the variety of cryptocurrency systems, including Bitcoin [15] and the altcoins. It is worth noting that Blockchain represents one special case of the distributed ledger technologies (DLTs), which are decentralised databases that rely on independent computers to record, share and synchronize digital transactions. Regardless the different forms, a DLT promises similar properties to those from a Blockchain. For simplicity reasons, we choose to use the term Blockchain in this paper and our discussion generally applies to DLT.

Take the Bitcoin Blockchain as an example, repeatedly, a certain number of new data entries (e.g. transactions) will be packed into a new block by a miner and appended to the existing (longest) chain. The new block includes the hash value of the last block of the current chain, and is formed with some specific features, e.g. a proof of work (PoW) needs to be carried out so that the hash value of the new block contains some number of consecutive zeros. After its formation, the new block will be broadcast to the whole network, and it will be accepted by other network nodes after everything being validated. Blockchain systems act as the key foundation platform for *smart contracts*, which facilitate automated execution of software programs in a verifiable manner. One of the notable examples is Ethereum, which is the second largest cryptocurrency system after Bitcoin and gains the popularity because of its powerful smart contracts functionality.

Benefiting from its unique decentralised architecture, Blockchain possesses the following useful properties.

- *Decentralisation and Democracy.* Everyone can potentially act as a miner and has the same privilege to generate blocks and approve blocks to the Blockchain. This is generally true for systems employing the PoW as the consensus mechanism in the permissionless scenario, while it can be different in other cases. Regardless, Blockchain eliminates a single trusted party and anchors the trust base to multiple parties.
- *Integrity, Immutability and Consistency.* If an attacker or a group of colluded attackers does not dominate the consensus process, e.g. informally, in the case of Bitcoin Blockchain more than 51% of the computing power is at the hands of semi-honest miners, then it will not be able to modify the existing blocks that have been agreed on by the consensus. In other words, when being a majority, semi-honest miners can guarantee a consistent view for the Blockchain users and assure that no malicious attackers (including dishonest miners) can manipulate the blocks on the longest chain. Note that the trust assumption based on 51% semi-honest miners is only a theoretical bound, while it has been shown that 25% colluded miners can disrupt the operations of Bitcoin Blockchain [10].

– *Transparency, auditability and Disintermediation.* Comparing to existing information systems, Blockchain offers more transparency towards not only the data of blocks but also the origination of these blocks. In a permissionless Blockchain, everything is totally transparent to the world, while it is transparent to the authorized entities in other cases. Transparency naturally leads to auditability, and it also help eliminate many intermediaries in practice particularly when smart contracts feature is equipped.

Numerous Blockchain systems and applications have been proposed so far, and we refer the readers to the abundant surveys (e.g. [3, 19, 21]) and observatory reports for more information (e.g. those from the EU Blockchain Observatory & Forum).

1.2 Emerging Blockchain-based Solutions

In the US, MIT is running a Digital Certificates project which uses Blockchain as a key infrastructure [7]. In EU, to support the digital single market, the Connecting Europe Facility (CEF) programme is funding a set of generic and reusable Digital Service Infrastructures (DSI). Among all, there exists a Blockchain DSI (the European Blockchain Services Infrastructure (EBSI)) which aims to accelerate the creation of cross-border services, where *diploma validation* is one of the selected use cases. Funded under EU's Horizon 2020 research and innovation programme, the QualiChain project is dedicated to verifying educational credentials based on Blockchain [9] and the EDSSI project is a similar project [8]. In Cyprus, University of Nicosia [16] tried to digitize and decentralize their internal processes and have issued their first academic certificates as a proof of concept. In France, BCDiploma [1] shares the same goal towards a global certification network of higher academic institutions. In Switzerland, Gresch et al. [12] proposed a blockchain based system for managing diplomas called UZHBC (University of ZuricH BlockChain), and Schär and Mösli [11] did a similar project through University of Basel's Center for Innovative Finance and BlockFactory Ltd[1].

Besides these initiatives, other researchers have also promoted Blockchain to solve the fraud issues in diplomas, e.g., Turkanović et al. [22], Serranito et al. [18], Tariq et al. [20], Brinkkemper [2]. Instead of focusing on technical solutions, Olivier et al. [17] investigated the business models for Blockchain-based diploma management solutions.

1.3 Contribution and Organisation

Most existing digitized diploma management systems have tried to directly transform the paper-based ancestors into digital systems. These Blockchain-based systems move a step further to provide more guarantees on security and interoperability. However, a comprehensive modeling of digitized diploma

[1] https://blockfactory.com/

management is still missing today. In particular, security and privacy requirements have not been systematically studied.

In this paper, we close the gap by initiating a systematic study on digitized diploma management systems. Our main contribution is two-fold. Firstly, we present a general system structure and identify both the functional requirements (e.g. data included in a diploma and the time-stamping of diplomas) and the non-functional requirements from both the security and privacy perspectives. We also analyse the existing systems and show that they all fall short to meeting the identified requirements. Secondly, we follow the security/privacy-by-design principles to propose a Blockchain-facilitated diploma management solution. By relying on some basic cryptographic primitives (e.g. digital signature and hash function) and data structures (e.g. hash tree), we show that the proposed solution satisfies all the identified requirements.

The rest of this paper is organized as follows. In Section 2, we describe the system architecture and identify the functional/security/privacy requirements. In Section 3, we analyse existing (Blockchain-based) diploma management systems with regard to the identified requirements. In Section 4, we describe our new Blockchain-facilitated diploma management solution. In Section 5, we demonstrate that our solution addresses the identified requirements well. In Section 6, we conclude the paper.

2 System and Threat Modeling of Diploma Management

We assume a general system architecture for digitized diploma management systems, as shown in Fig. 1. The players fall into three categories including user (who is first a student and then becomes a graduate after being issued the diploma), diploma issuer, diploma verifier, and an intermediary platform. In our modeling, we assume there is one diploma issuer, which serves any number of users and diploma verifiers. The intermediary platform might be optional when the diploma issuer and the diploma verifiers directly interact with each other. In Section 4, we argue that, by introducing Blockchain as the intermediary platform, we can solve the interoperability problem when diploma verifiers need to validate diplomas issued by different diploma issuers.

When setting up a digitized diploma management system, there should be an *Initialisation Phase* for all players to set up their parameters. The details of this procedure will depend on the specific solution, and we will not focus on it at this point. As indicated in the general system architecture, from the perspective of a user u, the workflow consists of four phases.

1. *Diploma generation.* User u registers and studies at an institute, which will act as the diploma issuer to issue a diploma to him once proper qualification is achieved.
2. *Diploma outsourcing.* If an intermediary platform is employed, the diploma issuer stores some information about user u's diploma on the platform, which will then take care of the following-up diploma verification activities.

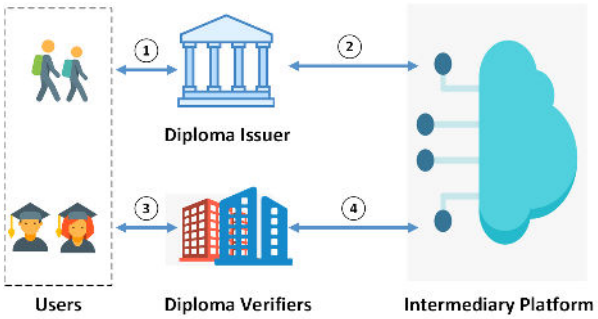

Fig. 1. General System Architecture

3. *Diploma usage.* User u presents some elements of his diploma to an organisation, who will then acts as the diploma verifier to validate these elements.
4. *Diploma verification.* The diploma verifier interacts with the intermediary platform to validate the elements received from user u. If there is no platform employed, then the interaction is directly with the diploma issuer.

Clearly, the first two phases will happen only once (except that diploma may be updated), while the last two phases can happen as many times as user u wants. Next, we elaborate on the functional and non-functional (i.e. security and privacy) requirements for a digitized diploma management system.

2.1 Functional Requirements

Regarding the functionality, we would like to emphasize a major difference between paper-based diploma and the fully digitized one. In the paper-based case, a diploma usually contains a minimum set of attributes. When the diploma verifier needs more information, for either confirming the diploma's validity or demanding supplementary data, it will interact with the diploma issuer with/without the involvement of the user. However, in the digitized case, we would like to digitize the whole process and eliminate the burden of frequent issuer-verifier communications. To this end, we highlight the following aspects.

In the *data* aspect, a diploma should contain all the necessary information for a diploma verifier to comprehensively evaluate the user's capabilities. The potential attributes could include at least:

- Name of the issuer
- Nature of the study
- Quality of the study
- Date of issuance
- Name of the user
- Gender of the user
- Birth date of the user

– Photo
– Identity number
– Courses taken and evaluation results
– Activities and achievements
– Teachers

The "Gender of the user" and "Birth date of the user" attributes are necessary to bind the diploma to the user and remove ambiguities among users who share the same name. In a country with a large population, the combination of these attributes may still not be able to uniquely identify a user, so it may be required to include other attributes, such as "Photo" and "Identity number". Although a diploma could contain a long list of attributes, we would like to emphasize that the user may choose to reveal only a selected set of attributes to a specific diploma verifier (in step 3 of the workflow).

In the *time-stamping* aspect, a diploma's issuance should be timestamped by some third party so that the diploma verifier can use this as a second factor to assure when the diploma has been issued. Suppose diplomas are only signed by the issuer's signing key based on some signature scheme, a diploma issuer could issue a diploma now and claim it has been issued 10 years ago. The timestamp anchor is crucial to prevent such fraudulent activity from an issuer or an attacker which has compromised the issuer (detailed in Section 2.2).

In paper-based systems, it is prohibitively costly to prevent fraudulent activities from users and issuers because fraud detection will require a lot of man power. In contrast, *efficiency and cost* has been advocated as an advantage for digitized systems, e.g. the aforementioned Blockchain-based solutions. We emphasize that the cost issue needs some special attention when an intermediary platform is employed, as has been attempted by Olivier et al. [17].

2.2 Security Requirements

When designing a digitized diploma management system, we first need to consider the following threats which also exist its paper-based ancestor.

– *Fake diploma issuer*. An attacker may pretend to be a legitimate diploma issuer and issues/sells diplomas for profits.
– *Diploma forgery*. An attacker may try to impersonate the legitimate diploma issuer to forge diplomas on its behalf.
– *Diploma issuer fraud*. A malicious diploma issuer can try to generate diplomas for users who are not officially registered. For example, the issuer can generate diplomas for non-existing users to exaggerate its performance, or can generate diplomas for celebrities to gain popularity.

Besides these, we need to tackle many new threats emerging in the digitized systems.

– *Diploma issuer corruption*. When an attacker corrupts the diploma issuer and obtains its credentials, then this attacker may try to forge diplomas or do other harmful things (e.g. privacy leakages).

– *User corruption*. When an attacker corrupts a user, as an immediate effect, it will result in privacy breaches. Besides, it may also cause security issues, e.g. blackmailing and denial of service.
– *Intermediary platform corruption*. When an intermediary platform is employed, there is a risk that it will be corrupted. If this happens, the diploma verification service may be disrupted. It is worth noting that Blockchain possesses its own security vulnerabilities, therefore it is inappropriate to assume it to be corruption free.
– *Diploma data confidentiality*. An attacker may try to obtain certain information in a user's diploma while the user wants to keep such information confidential against the attacker.
– *Diploma data integrity*. Once a diploma is issued, it should not be altered by any entity other than the diploma issuer. In case the any change needs to be made to the diploma, it should be done with the consents from the diploma issuer, the user and the time-stamping third party. An attacker may try to modify the diploma without all the consents.

Note that there are other threats against digitized systems in general, e.g. denial of service attacks. We omit them in our discussion.

2.3 Privacy Requirements

Privacy is not a serious concern in traditional paper-based systems due to the absence of effective information transmission media. However, things can change dramatically due to the introduction of intermediary platforms (e.g. Blockchain) and automated information processing and sharing capabilities (e.g. Blockchain's smart contracts). It is worth emphasizing that automated information processing and sharing brings serious privacy concerns with respect to privacy regulations such as EU's GDPR. Next, we categorize the privacy concerns as follows.

– *User privacy*. Some user may not only want to keep her diploma private (i.e. diploma data confidentiality) but also want to hide the fact that his/her diploma has been verified. For example, if an employer notices that an employee's diploma has been verified by its competitor, then it may infer that this employee has applied a job there.
– *Diploma issuer privacy*. In practice, a diploma issuer may choose not to reveal certain information about the issued diplomas, e.g. a university may not want to reveal the precise number of diplomas to the general public.
– *Diploma verifier privacy*. For a diploma verifier, its requests may reveal a lot of information, e.g. how many applicants it has, who are these applicants, when these applicants have submitted their applications, and so on. In practice, such information could reflect business secrets so that the diploma verifier may want to hide it as much as possible.

2.4 Summary

In digitized diploma management systems, some traditional security concerns (e.g. fake diploma issuer, diploma forgery, and diploma issuer fraud) still exist, but it is much easier to address them with cryptographic tools such as digital signatures. On the other hand, many new security and privacy concerns emerge, partly due to the involvement of third-party intermediary platform. In particular, enriching diploma with more detailed data attributes amplifies the consequences of any security or privacy breach.

As mentioned above, combating fraud in paper-based systems is tedious and prohibitively costly, while it becomes much easier in digitized systems because this is the motivational design purpose. However, a digitized system introduces new cost, with respect to credential management, data storage and verification. Beside security and privacy, this aspect will affect the practicality of new digitized systems.

3 Analysis of Existing Systems

In this section, we analyse the existing digitized diploma management systems with respect to the identified requirements in Section 2. Due to the fact that most of these solutions have been described at high level and focused on the workflow and smart contract design, our analysis stays at high level as well. Regarding the security and privacy requirements, no rigorous formulation and discussion have been provided for these solutions, so that our analysis results are only qualitative without in-depth cryptographic analysis.

3.1 Meeting the Functional Requirements

Table 1 indicates how well the existing solutions meet the *data* and *time-stamping* aspects of the functional requirements from Section 2.1. For the former, most solutions only mention the term "diploma" (or, "certificate") while providing no detail about it. In contrast, some solutions (e.g. [13, 7]) assume that any capability assertion can be contained in a diploma. For the latter, most solutions meet the requirement to some extent, benefiting from their adoption of Blockchain as the intermediary platform.

Table 1. Analysis of Functional Properties

	Data Aspect	**Time-stamping Aspect**
Schär and Mösli [11]	Not detailed	Achieved via Blockchain
Gresch et al. [12]	Not detailed	Achieved via Blockchain
Turkanović et al. [22]	Not detailed	Achieved via Blockchain
Serranito et al. [18]	Not detailed	Achieved via Blockchain
Tariq et al. [20]	Include many attributes	Achieved via Blockchain
MIT's Digital Certificates [7]	Can be any capability assertion	Achieved via Blockchain
Open Badges [13]	Can be any capability assertion	Not mentioned

The *efficiency and cost* aspect is hard to analyse, and efficiency is often taken for granted as a result of employing smart contracts. From the perspective of business models, Olivier et al. [17] made some dedicated investigation from both the qualitative and quantitative aspects. They show that a sustainable business model relies on a lot of factors, such as market share, technology maturity, and acceptance of users and employers. Below, we emphasize three types of cost that have not been taken seriously in the existing systems.

- *Diploma storage cost*. When a diploma is issued digitally, e.g. by signing a PDF file, the issuing is fairly efficient and costs almost nothing. However, unlikely the paper diploma which will be kept by the user, the PDF file and/or its digest will need to be stored, e.g. on the Blockchain platform, and this will incur additional storage costs.
- *Diploma verification cost*. For Blockchain-based solutions, verification is done through a function call to the smart contract and is usually very efficient. The cost will depend on the smart contract cost and the underlying Blockchain platform.
- *Sustainability cost*. When an intermediary platform is employed, sustainability becomes a concern. For example, the platform might get bankrupted and then the solution needs to be migrated to a new platform. Moreover, the platform's business model may change and subsequently affect the cost and efficiency of operations. In some cases, this may even disrupt the solution.

3.2 Meeting the Security Requirements

Table 2 summarizes how the existing systems meet the security requirements from Section 2.2. Aligned with their major motivational objective, most solutions meet the requirements on *Fake diploma issuer* and *Diploma forgery*. In addition, most solutions ask the diploma issuer to sign diplomas and only store the hash values on Blockchain, therefore, they can meet the requirements on *Diploma data confidentiality* and *Diploma data integrity*. In comparison, other requirements are either completely ignored (marked with ?) or somehow partially addressed (marked with Yes$^?$).

Table 2. Analysis of Security Properties

	[11]	[20]	[22]	[18]	[12]	[13]	[6]
Fake diploma issuer	Yes	Yes	Yes	Yes	Yes	?	Yes
Diploma forgery	Yes	Yes	Yes	Yes$^?$	Yes$^?$	?	Yes
Diploma issuer fraud	?	?	?	?	?	?	?
Diploma issuer corruption	?	Yes$^?$	Yes$^?$	?	?	?	?
Intermediary platform corruption	?	Yes$^?$	Yes$^?$	?	?	?	?
Diploma data confidentiality	Yes	Yes	Yes	Yes	Yes	?	Yes
Diploma data integrity	Yes	Yes	Yes	?	Yes	?	Yes

For the *Diploma forgery* requirement, the solutions from [18] and [12] do not explicitly say whether a diploma will be signed or not. Therefore, an attacker may be able to forge diplomas if it somehow gains access to the network. For the *Diploma issuer corruption* and *Intermediary platform corruption* requirements, the solutions from [20] and [22] have introduced the role of "auditor"/"observer"/"accreditation body" into their solution so that it may somehow help meet the requirements.

3.3 Meeting the Privacy Requirements

Most existing solutions have security as their major objective, which is typically achieved by combining cryptographic primitives (e.g. digital signature) and Blockchain features. In contrast, privacy has been largely ignored and make these solutions vulnerable in reality. In Table 3, we indicate how the solutions meet the privacy requirements from Section 2.3.

Table 3. Analysis of Privacy Properties

	[11]	[20]	[22]	[18]	[12]	[13]	[6]
User privacy	?	Yes$^?$	Yes$^?$	Yes$^?$	?	?	Yes$^?$
Diploma issuer privacy	?	Yes$^?$	Yes$^?$	Yes$^?$	?	?	Yes$^?$
Verification requester privacy	?	Yes$^?$	Yes$^?$	Yes$^?$	?	?	Yes$^?$

The solutions from [11–13] do not provide privacy guarantee because they employ permissionless Blockchain, namely Ethereum. It is worth noting that *User privacy* cannot be achieved even though only the hash value of a diploma is stored on the Blockchain. The reason is simply because the hash value can uniquely identify the diploma and the user. In fact, EU's GDPR has already pointed out this kind of vulnerability.

4 New Blockchain-facilitated Solution

In this section, we first introduce the diploma format for our solution and then motivate the usage of Blockchain for diploma management. Finally, we describe our solution in detail.

4.1 Diploma Format

In order to support the selective disclosure of diploma attributes and avoid attribute-specific signatures, we propose to organise the attributes in a binary tree structure. For description simplicity, let's assume that a diploma can include N attributes where N is an power of 2. We denote the attributes as $attr_1, attr_2, \cdots, attr_N$. For the sake of privacy protection as we will explain below, each attribute is accompanied with a salt value, denoted as $a\text{-}salt_i$ for $attr_i$.

Let H be a cryptographic hash function, we can construct the attribute hash tree as follows.

1. Associate each leaf node with an attribute and a salt value, i.e. $Node_i$ is associated with $attr_i$ and $a\text{-}salt_i$. For each $Node_i$ $(1 \leq i \leq N)$, compute its value as $Hash_{[i]} = \mathsf{H}(attr_i || a\text{-}salt_i)$.
2. For internal node $Node_{[1,2]}$ which has $Node_1$ and $Node_2$ as its children, compute its value as $Hash_{[1,2]} = \mathsf{H}(Hash_{[1]} || Hash_{[2]})$. Do the same for internal nodes $Node_{[3,4]}, Node_{[5,6]}, \cdots, Node_{[N-1,N]}$.
3. For internal node $Node_{[1,4]}$ which has $Node_{[1,2]}$ and $Node_{[3,4]}$ as its children, compute its value as $Hash_{[1,4]} = \mathsf{H}(Hash_{[1,2]} || Hash_{[3,4]})$. Do the same for internal nodes $Node_{[5,8]}, Node_{[9,12]}, \cdots, Node_{[N-3,N]}$.
4. Continue as above until reaching the root node $Node_{[1,N]}$ which has the value $Hash_{[1,N]} = \mathsf{H}(Hash_{[1,\frac{N}{2}]} || Hash_{[\frac{N}{2}+1,N]})$

For illustration purpose, this tree construction process is shown in a toy example with four attributes in Figure 2.

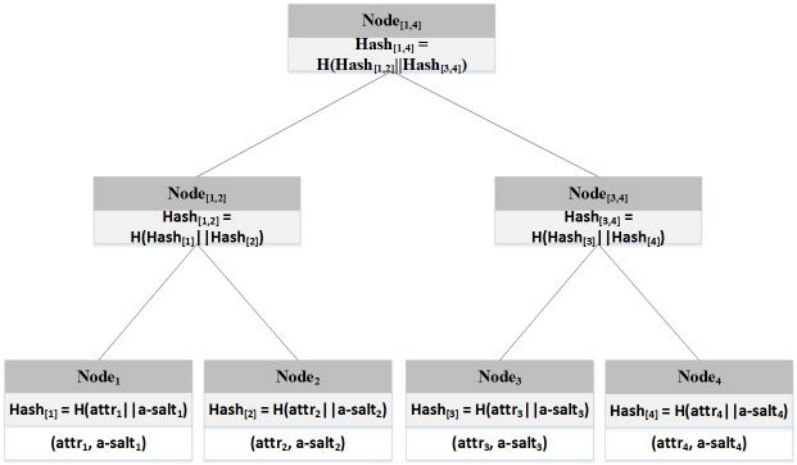

Fig. 2. Hash Tree for Diploma Attributes

As most existing solutions do, it is natural to ask the diploma issuer to sign a user's diploma. However, this is hardly sufficient to satisfy the desired security requirements in our threat model. For example, the diploma issuer can commit a *Diploma issuer fraud* without any barrier. Therefore, in our solution, we additionally require the user to sign the diploma as well. By doing so, the aforementioned diploma issuer's fraudulent activities can be prevented. Suppose the user and the diploma issuer possess key pairs (PK_u, SK_u) and (PK_I, SK_I) respectively, for a secure digital signature scheme (**Sign**, **Verify**). Then the user's diploma will be in the following format.

$$diploma_u = (attr_1, a\text{-}salt_1, \cdots, attr_N, a\text{-}salt_N, Hash_{[1,N]}, \mathbf{Sign}(Hash_{[1,N]}, SK_I), \mathbf{Sign}(Hash_{[1,N]}, SK_u))$$

For the sake of notation clarity, we use dig_u to denote the root value $Hash_{[1,N]}$ and the two signatures in the diploma $diploma_u$. It will be used to prepare information for Blockchain in Section 4.3.

4.2 An Attempt without Intermediary Platform

Similar to paper-based diploma management solutions, a digitized solution can also be issuer-based or institution-based without any intermediary platform. In this case, the issuer needs to deal with everything that is necessary for addressing every diploma verifier's verification request. With this old-style design, we can easily observe the following challenges for the diploma issuer.

- The diploma issuer needs to be always online in order to deal with the potential request from any relevant diploma verifier. This stands for a high availability requirement, and furthermore implies a strong cyber-threat resilience requirement because the issuer is a lucrative cyber-attack target.
- The issuer needs to dedicate a lot of efforts to support diploma verification, e.g. maintaining a database for diplomas, deploying user interfaces for diploma verifiers, maintaining its signature key pairs and keeping a log of key update history.
- The issuer needs to leverage on some third party service to time-stamp diploma issuance. Note that the time-stamping operation is crucial to prevent fraudulent activities of the issuer itself.

Considering the diversity of all possible diploma issuers, it is to be expected that they will adopt various different systems of their own choices. These systems will have very different interfaces and workflows. It in turn poses a significant challenge to the diploma verifiers which are forced to use all these different systems.

4.3 Blockchain-facilitated Diploma Management Solution

Based on our discussion in Section 4.2, it is clear that a (trusted) third party is required to provide the time-stamping service. Furthermore, it will be ideal if this third party can help address the challenges facing both the diploma issuer and diploma verifiers. In the following, we describe a solution by involving a Blockchain platform as the intermediary to provide the time-stamping service and facilitate other operations. An additional advantage of Blockchain is that it facilitate interoperability between different solutions. For instance, these solutions can use the same Blockchain platform (e.g. a consortium Blockchain operated by education agencies) and the same smart contracts. In this case, interaction with different solutions will be easy for the diploma verifiers.

We assume the Blockchain platform stays neutral from the diploma issuer and the users so that they will not be able to influence the standard operations of the platform. To this end, the diploma issuer or any user should not own any

mining node of the Blockchain platform. For simplicity, we further assume that the Blockchain platform provides two smart contracts. $Contract_{issue}$ allows the diploma issuer and users to interact with the Blockchain platform, to store and update data on the platform. $Contract_{verify}$ allows diploma verifiers to interact with the Blockchain platform, to retrieve data from the platform. The system structure of our Blockchain-based solution is shown in Fig. 3.

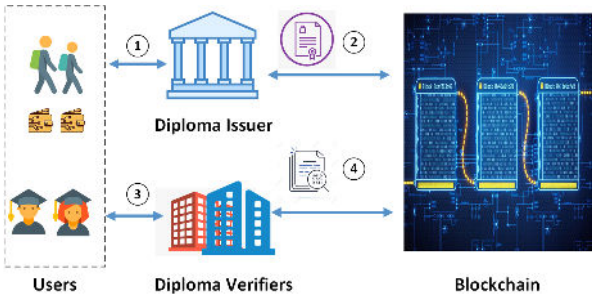

Fig. 3. Blockchain-facilitated System Architecture

Initialisation Phase. The diploma issuer and users should run the $Contract_{issue}$ smart contract to store their public key certificates on the platform, and also update these certificates when the signing keys have been updated. Every user should obtain a wallet that is compatible with the Blockchain platform, and the wallet should allow the user to securely store his data and to interact with the Blockchain (e.g. execute smart contracts). Similarly, the diploma issuer should also obtain a wallet. At the same time, the issuer should maintain a local database to store data from all its users.

Referring to the general workflow from Section 2, the detailed instantiation for our solution are as follows.

1. *Diploma generation.* After a user u has passed the qualification tests, the diploma issuer issues diploma $diploma_u$ according to the format defined in (1). In addition, a salt value $d\text{-}salt_u$ is chosen to protect $diploma_u$. Use u stores $diploma_u$ and $d\text{-}salt_u$ in his wallet, and the diploma issuer stores the same data in its local storage. In case the user's wallet is lost, the diploma issuer can help him recover his data based on its local storage.
2. *Diploma outsourcing.* At a certain time, the diploma issuer batches the issued diplomas and store some information about them on the Blockchain as follows.
   (a) The diploma issuer collects the diplomas for all relevant users. For simplicity, we assume there are l diplomas and use the subscripts $1 \leq i \leq l$ to distinguish them. To hide the precise value of l, the diploma issuer chooses an integer $L \geq l$ and generates some fake diplomas $diploma_i$ for $l + 1 \leq i \leq L$.

(b) For each leaf $diploma_i$ ($1 \leq i \leq L$), the diploma issuer computes a hash value as $Hash_{[i]} = \mathsf{H}(dig_i \| d\text{-}salt_i)$, where dig_i is defined at the end of Section 4.1.

(c) Finally, the diploma issuer runs the $Contract_{issue}$ smart contract to store $Hash_{[i]}$ ($1 \leq i \leq L$) on the Blockchain. In addition, it also generates a signature for the concatenation of these hash values (i.e. $Hash_{[1]} \| \cdots \| Hash_{[L]}$) and stores it on the Blockchain.

3. *Diploma usage.* When the user u wants to present some attributes of his diploma to a diploma verifier, he should present the corresponding salt values for these attributes as well. In addition, he should also provide the necessary values of relevant leaf/internal nodes, so that the diploma verifier can compute the value of the root node of the attribute hash tree. Finally, he should provide the two signatures in his diploma. For instance, referring to the toy example Fig. 2, if user u wants to present $attr_1$, then he should provide the following values.

$$(attr_1, a\text{-}salt_1, Hash_{[2]}, Hash_{[3,4]}, \mathbf{Sign}(Hash_{[1,4]}, SK_I), \mathbf{Sign}(Hash_{[1,4]}, SK_u))$$

4. *Diploma verification.* The diploma verifier runs the $Contract_{verify}$ smart contract to retrieve the up-to-date public keys of the diploma issuer and user u. Then, it performs some preliminary verification on the received data as follows.

   (a) It uses the received values from user u to construct the value of the root node for his diploma.

   (b) It then uses the retrieved public keys from Blockchain to verify the received signatures from user u.

   If the above verification passes, the diploma verifier runs $Contract_{verify}$ to retrieve the values stored by the the diploma issuer in step 2.(c). Then, it continues the verification as follows.

   (a) It verifies the signature on the concatenation of hash values.

   (b) It finally checks that the diploma it receives from user u matches with one of these hash values.

   The above two verification guarantees that the diploma information presented by user u is up-to-date (in case that the attributes in the diploma has been updated after its issuance).

By design, the above solution has addressed the functional requirements, from both the *data* aspect and the *time-stamping* aspect. Moreover, we have minimized the amount of data on Blockchain and simplified the interaction between the diploma issuer/verifier and the Blockchain platform. Due to the smart contracts only perform data storage and retrieval and there is nothing new in comparison to the standard ones, so that we skip a detailed description on them.

5 Security and Privacy Analysis of the new Solution

Before going into the security and privacy properties, we first summarize the assumptions underlying our analysis.

- The Blockchain is neutral so that its operations will not be influenced by the diploma issuer and users. Some authority should regularly audit the Blockchain platform and its transactions resulted from the diploma management solution.
- The public keys of the diploma issuer and users should be certified by some authority which acts as a root of trust. How to determine the authority is beyond the scope of this paper and will be decided when the proposed solution is actually deployed.
- The diploma issuer should faithfully carry out its duty. For example, it should properly manage its key pair: updating the key pair upon revocation and putting the new public key on the Blockchain immediately. It should stick to the workflow to generate diplomas and manage the information on Blockchain. Except for these legitimate tasks, the diploma issuer may abuse the system for some additional benefit.

It is worth noting that different types of Blockchain platform can be adopted for our solution. It is natural to assume that a group of higher education agencies can set up a consortium Blockchain and dedicate it to diploma management. The Blockchain can be further made permissioned so that access to the platform needs to be authorised. On the other hand, a permissionless Blockchain like Ethereum can also work. However, we note that it is difficult to regulate a permissonless Blockchain, particularly when it spans across different jurisdictions. Furthermore, the cost and efficiency aspects may be less predicable.

5.1 Security Analysis

When designing our solution, we have adopted a security-by-design approach by taking into the security requirements from Section 2.2. Next, we briefly explain how these requirements have been addressed by our solution.

- *Fake diploma issuer*. In our solution, the diploma issuer is required to sign its diplomas and its signing key is certified by some authority and stored on the Blockchain. Therefore, this threat is mitigated because a fake issuer will not possess a certified public key.
- *Diploma forgery*. As long as the signature scheme is secure, it will be computationally hard for an attacker to forge a diploma.
- *Diploma issuer fraud*. Sicne a diploma needs to be signed by both the user and the diploma issuer, this threat is automatically mitigated. Note that if the diploma issuer colludes with users then it can issue seemingly legitimate diplomas. How to mitigate such fraud is beyond the scope of this paper.

– *Diploma issuer corruption*. In our solution, when an attacker corrupts the diploma issuer and obtains its credentials, then this attacker may try to forge diplomas for users who collude with the attacker. In any case, a forgery can be detected by the diploma issuer because some information of the diploma should be stored on the Blockchain. Once detected, the forged diplomas can be made illegal. Besides forging diplomas, we do not see any other threats when the attacker obtains the signing key.
– *User corruption*. In our solution, we have required every user to use a digital wallet for managing his diploma information. In reality, most of the wallets usually possess very strong security guarantees (e.g. designed based on trusted computing technologies) as they are often used to secure crypto-currencies. Therefore, this threat can be mitigated very well with existing products.
– *Intermediary platform corruption*. When the underlying Blockchain platform is carefully selected, the risk of platform corruption is pretty low. This is particularly true if we choose a platform which is operated by a consortium of trusted entities (e.g. higher education agencies). Nevertheless, there is still a risk of corruption in theory. That is why we made the assumption that platform auditing should be carried out regularly so that the threat can be minimized to an acceptable level.
– *Diploma data confidentiality*. Salt values have been used for protecting attributes in diplomas so that a user can reveal only the selected attributes to the verifier. Furthermore, salt values are also used to protect diplomas when their information is stored on the Blockchain. Coupled with the cryptographic hash function, confidentiality is guaranteed.
– *Diploma data integrity*. Due to the fact that the root value of the attribute hash tree is required to be signed by both the diploma issuer and the user, no modification can be made without the consents from both these entities.

5.2 Privacy Analysis

In the design, we have tried to minimise the information disclosure to both the diploma verifier and the Blockchain platform, through organising the diploma attributes in a tree structure and integrating attribute-wise and diploma-wise salt values into the computation. Next, we briefly explain how the requirements from Section 2.3 have been addressed by our solution.

– *User privacy*. For a user u, the only information stored on the Blockchain is $Hash_{[u]} = \mathsf{H}(dig_u \| d\text{-}salt_u)$. The value of $d\text{-}salt_u$ is only known to user u, the diploma issuer and the diploma verifiers chosen user u. Therefore, it is computationally infeasible for any other entity to link this hash to user u. Furthermore, the diploma verifier will retrieve a batch of hash values to verify user u's information, so that any other entity will not know which hash value is targeted for verification. This provides a second level of protection.
– *Diploma issuer privacy*. In our solution, by carefully choosing the value for L in the *Diploma outsourcing* phase, the diploma issuer can hide the precise value of l.

– *Diploma verifier privacy*. By design, our solution hides the specific diploma a diploma verifier is querying from the Blockchain platform. If the diploma verifier possesses multiple wallets, then it will be difficult for an attacker to link them. Moreover, the diploma verifier may also choose a trusted third party to interact with the Blockchain platform on its behalf. This can also hide the diploma verifier's activities from the potential attackers.

6 Conclusion

In this paper, we have systematically studied digitized diploma management systems, from the system structure to requirements and to Blockchain-based solutions. The Blockchain-native time-stamping feature is crucial for our solution to prevent fradulent activities from diploma issuers, while the integrity and immutability features make Blockchain an ideal platform to store diploma-related data and track key updating history. In comparison, the smart contract feature seems less important to us. This leads to an observation that thinner variants of existing Blockchain platforms (e.g. Ethereum) could suffice for our needs. It is an interesting future work to investigate this further and implement our solution for a detailed performance study.

References

1. Blockchain Certified Data 2021. Bcdiploma, 2021.
2. F. L. Brinkkemper. Decentralized credential publication and verification : a method for issuing and verifying academic degrees with smart contracts, June 2018.
3. B. Butijn, D. A. Tamburri, and W. Heuvel. Blockchains: A systematic multivocal literature review. *ACM Comput. Surv.*, 53(3), 2020.
4. H. Clifton, M. Chapman, and S. Cox. śtaggeringt́rade in fake degrees revealed, 2018.
5. Inc. Concentric Sky. Badgr service, 2021.
6. Digital Credentials Consortium. Building the digital credential infrastructure for the future, 2021.
7. Digital Credentials Consortium. Digital academic credentials, 2021.
8. EDSSI Consortium. EDSSI – European Digital Student Service Infrastructure, 2021.
9. QualiChain Consortium. Decentralised qualifications' verification and management for learner empowerment, education reengineering and public sector transformation, 2021.
10. I. Eyal and E. G. Sirer. Majority is not enough: Bitcoin mining is vulnerable. In Nicolas Christin and Reihaneh Safavi-Naini, editors, *Financial Cryptography and Data Security - 18th International Conference*, volume 8437 of *LNCS*, pages 436–454. Springer, 2014.
11. F. Mösli F. Schär. Blockchain diplomas: Using smart contracts to secure academic credentials. *Beiträge zur Hochschulforschung*, 41(3):48–58, 2019.
12. J. Gresch, B. Rodrigues, E. J. Scheid, S. S. Kanhere, and B. Stiller. The proposal of a blockchain-based architecture for transparent certificate handling. In W. Abramowicz and A. Paschke, editors, *Business Information Systems Workshops - BIS 2018 International Workshops*, volume 339 of *Lecture Notes in Business Information Processing*, pages 185–196. Springer, 2018.
13. IMS Global Learning Consortium Inc. Open badges 2.x, 2021.

14. T. Levin. Dean at M.I.T. Resigns, Ending a 28-Year Lie, 2007.
15. S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf, 2008.
16. University of Nicosia (UNIC). Blockchain certificates (academic others), 2021.
17. M. Oliver, J. Moreno, G. Prieto, and D. Benítez. Using blockchain as a tool for tracking and verification of official degrees: business model. In *29th European Regional Conference of the International Telecommunications Society (ITS): "Towards a Digital Future: Turning Technology into Markets?"*, 2018.
18. D. Serranito, A. Vasconcelos, S. Guerreiro, and M. Correia. Blockchain ecosystem for verifiable qualifications. In *2nd Conference on Blockchain Research & Applications for Innovative Networks and Services, BRAINS 2020*, pages 192–199. IEEE, 2020.
19. M. Swan. *Blockchain: Blueprint for a New Economy*. O'Reilly, 2015.
20. A. Tariq, H. B. Haq, and S. T. Ali. Cerberus: A blockchain-based accreditation and degree verification system, 2019.
21. H. Treiblmaier and T. Clohessy. *Blockchain and Distributed Ledger Technology Use Cases*. Springer, Cham, 2020.
22. M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić. Eductx: A blockchain-based higher education credit platform. *IEEE Access*, 6:5112–5127, 2018.