Academic Achievement Recognition and Verification Using Blockchain

Axel Curmi and Frankie Inguanez(✉)

Malta College of Arts, Science and Technology, Paola PLA9032, Malta
{axel.curmi.a100445,frankie.inguanez}@mcast.edu.mt
http://ict.mcast.edu.mt

Abstract. Falsification of certificates is a growing concern and the verification process can be a lengthy and challenging one. In this research, we are proposing a distributed ledger-based solution for the storage and verification of academic qualifications. An entity that would want to verify certificates can make use of our API service that would, in turn, scan a certificate, find the matching certificate template, extract the necessary data and verify it from a blockchain stored copy. In this research, we also propose an improved manner of verifying the ownership of a blockchain public address which also does not allow a user to present an address of a third party, this being one of the common security concerns of similar solutions. We also calculate the possible costs to adopt this system in all EU countries taking into consideration different gas prices, which is a determining factor to the transaction cost of a blockchain network. We conclude that a blockchain based certificate verification system addresses various issues related to document forgery and is a viable solution even with the current state of technology.

Keywords: Blockchain · Smart contracts · Certificates · User-access control · OCR

1 Introduction

Academic background and merit misinformation in CV have become a problem as people are more inclined to provide wrong information to seek advantage over the ever-growing competition. In addition, such information has become more difficult and bureaucratic to validate due to the ever-increasing security and privacy policies adopted by organisations. Falsification of such information not only sheds bad light on graduates, but also damages the reputation of the providing institution. Therefore, automated, easy and instant validation is required.

In this research, distributed ledger technology is used to publish academic achievement information on a peer-to-peer distributed network, known as the blockchain, such that this crucial information is protected thanks to advanced cryptographic techniques. Provided that both the academic institution and certificate holder have public blockchain addresses, a smart contract is used to

U. Schwardmann et al. (Eds.): Euro-Par 2019 Workshops, LNCS 11997, pp. 153–165, 2020.
https://doi.org/10.1007/978-3-030-48340-1_12

publish the recognition across parties, this equivalent to the graduation ceremony. Thus, the purpose of storing the certificate information on the blockchain is purely to serve any verification requests in a fully automated and instant manner, and not as a datastore to support such a system, so a mix of off-chain and on-chain data is needed. When a 3^{rd} party receives a certificate and needs to verify its authenticity they can use our proposed API through which they can send a scanned image of the physical certificate together with the public blockchain address of the certificate holder. OCR and regular expression patterns make it ever more possible to extract the necessary information from existing documents, thus, further automating and improving the process flow. The extracted information would include the institution, for which the public blockchain address would have been previously registered. If the extracted data matches what is found on the blockchain then the certificate is verified instantly. One of the common challenges in having a blockchain based system is the verification of ownership for a blockchain public address, which is mostly of a concern in this scenario with regards to a certificate holder should an individual want to impersonate another to claim ownership of their achievements. We propose an adaptation of other research to solve this problem, which will be addressed further in Sect. 3. We also study the financial viability of our solution by identifying the low cost for registering the smart contract and evaluating our solution with the total number of tertiary qualifications in Europe. Given that there is no real rush to have certificates published instantly, these can be staged over a prolonged period in order to reduce the gas price and thus the actual financial cost for publishing certificates. The verification of academic achievement is technically free, in terms of blockchain transaction costs, yet computational power is needed for the extraction of data from scanned certificates as well as for bandwidth so to offer a good quality of service.

This paper is structured as follows, in Sect. 2, we present the Literature Review. The proposed solution is showcased in Sect. 3, the Research Methodology. In Sect. 4 the results are presented and discussed in detail, with concluding arguments and recommendations in Sect. 5.

2 Literature Review

2.1 Problems of Printed Certificates

Even though printed certificates are still preferred and seen as the most secure form of certificates, paper documents have a few notable disadvantages to keep in mind such as [5]: 1) Not being immune to forgery; 2) Awarding bodies are the single point of failure, meaning that certificates can still be valid, however the ability to validate them would be lost; 3) Secure certificates are costly (passports, routinely cost between €20 and €150); 4) No way to revoke the certificate without having the owner relinquish control; 5) Verification process is time consuming.

2.2 The Blockchain

A blockchain can be used to minimise the authority an intermediary has within a centralised service, such as validation of academic certificates [8]. For the purpose of this research, we shall be limiting the scope to public blockchain networks. By using a public blockchain, the data is not stored in a centralised location, instead, it is distributed between all participants in the network. By using this design, data is accessible to any participant in the network and secure at the same time, which means that the participants do not need to trust each other, including the owner of the data. This is because every participant in the network holds a ledger which contains every transaction taken place since the genesis block, and each participant can contribute to the creation of new blocks. Adding new blocks to a blockchain is an irreversible operation, meaning that once a block has been added to the chain in a validated state, this block or the data contained within the block can never be removed, even by the original author of the data. This is done by the way the blockchain structure is built, as every block contains two hash values, one for the previous block and one for itself. Any attempts to tamper with a block would invalidate the entire chain due to mismatches in hash values. As [8] stated, a consensus algorithm is used to achieve mutual trust between every participant in the network, as the creation of new blocks on the chain has to follow a strict protocol. At the time of writing, the two most popular consensus algorithms are: proof-of-work, and proof-of-stake. Bitcoin and Ethereum currently both operate with proof-of-work, however proof-of-stake is being considered by Ethereum as it is more cost effective and wastes a lot less energy. The number of research publications is greatly increasing and spreading around the globe [14], a few notable researches will be reviewed next.

2.3 Blockchain in Education

[10] stated that an application for blockchain in education would be to store records of achievement and credit, which would be added by the awarding institutions and be later accessed by the students. Having certificates published on the blockchain provides solutions to the issues regarding paper certificates, by providing public information regarding whether a certificate has been truly awarded to a certificate holder. However, as [10] mentioned, the blockchain does not verify the honesty of either party. Hashing techniques can also be used on the document such that rather than publishing private information, a digest of the document could be uploaded to act as a signature of the document while preserving the privacy of the document itself [5].

Various solutions to store certificates on the blockchain have been applied to several educational institutions, and the majority are built on the Bitcoin blockchain [11]. In Malta, the Blockcerts platform which was developed by MITs Media Lab and Learning Machine has been launched and will be used to issue and verify credentials using the Bitcoin network [7], and is currently the only open standard for issuing and verifying records using the blockchain [5]. In [11] the researchers have presented a proof of concept prototype, implemented on

the open-source ARK blockchain platform, which grants academic credits to students, according to the European Credit Transfer and Accumulation System (ECTS), after they have successfully completed a course. This prototype is built on a consortium based distributed ledger, which will allow 3^{rd} parties to easily validate a student's credits, after being granted access permission.

A solution for mitigating falsification of certificate documents has been presented by [13], in which the prototype is built by using a central server acting as a database, such that institutions publishing certificates communicate with this server to obtain a QR code, and 3^{rd} parties communicate with this server to validate a certificate by simply scanning the QR code found on the document. The authors also mention that after validation testing, problems with treating user credentials were identified and later rectified from the system to avoid other major issues. A method for confirming ownership of an address has been presented by [11], in which the untrusted entity is given a randomly generated number, via private channels, representing a value amount, such that if the number was 1234, the value amount would be 0.001234 ETH. Having received the randomly generated number, the untrusted entity has to issue a transaction to the known party with the correct value amount. The known party then checks the transaction, and if the transaction amount is equal to the randomly generated number, the entity is proven to be the true owner of the address. Even though this method works, it does not stop entities working together by sharing the randomly generated number to validate their addresses for each other. [9] explains the importance of having participants protect, store, and backup their private keys not only digitally, but also in the physical world due to identity fraud. One solution to this issue would be to implement digital private keys into physical keys, such as magnetic stripe cards, devices with embedded ROM chips, and smart cards. This would allow the participant to use the application without having to remember the secret key, and if compromised, the adversary would not be able to retrieve the private key. In the event of losing the private key, the owner could personally contact the awarding body and transfer the awards to a new blockchain address, provided his/her identity is successfully proven [11]. Splitting the private key into two halves, and storing each half in a different medium is a solution to this, should one get lost, the private key is not compromised and can be easily changed [9]. Another solution would be to implement multi-signature wallets where a group blockchain addresses could be combined into one. Therefore, if one of the addresses was lost and unrecoverable, a new address could be generated as a replacement by using signatures from other addresses. This also improves security in the event of having a compromised secret key, this is because if the adversary attempts to impersonate the original owner, he/she would require official signatures from the other addresses.

2.4 Image Processing and Hard-Copy Documents

A method for evaluating the quality of certificate and bill images, such that images with poor quality are filtered out, keeping only the high-quality ones was proposed by [6]. However, this research does not consider optical character

recognition (OCR) accuracy, which is an important requirement for bills and certificate recognition systems. [1] and [3] have proposed similar solutions, which mitigate forgery in hard copy documents by means of OCR, cryptography and 2D bar-codes (QR codes). The proposed solutions are very similar and follow three important steps when creating secure physical documents, which are: 1) Retrieve textual data from the document; 2) Generate a QR code based on the data to be validated; 3) Affix QR code on original certificate for validation process. However, when generating the QR codes [3] uses the selected text to be validated, while [1] uses the digest of the specified region of interest, thus having less data to be put into a QR code. [1] stated that the main challenge in the proposed system was the accuracy of the OCR, and thus experiments regarding the accuracy in terms of error occurrence were performed. The researchers outline two major factors which affect the accuracy of OCR, which are: 1) The font used; 2) Character weight. Their results indicate that the font "Times New Roman" performed the best by showing minimum error and using bold characters in the specified region of interest gives maximum performance. Tesseract OCR was used by [3] and concluded that overall recognition with case insensitivity was considerably better than case sensitivity.

3 Research Methodology

This research focuses on the verification of physical certificates, which information has been published on a blockchain network by the rewarding academic institution via a smart contract to the certificate holder on the respective public blockchain address. This research has been staged into three phases: 1) Institution registration and setup; 2) Issue of certificates; 3) Automatic verification and validation of certificates. From previous research [2], it was noted that the structure of the system should be implemented in a way that would allow academic organisations to publish certificates on their own smart contracts rather than one centralised smart contract. With this design academic organisations have several benefits, such as: 1) Complete control over smart contract containing certificates; 2) Complete freedom in choosing which information to be published as validation material.

Every academic organisation deploys a smart contract, with which all information about academic achievements found on physical certificates are published as a transaction between the academic organisation, certificate holder and actual smart contract, the equivalent of the physical certificate. This research mostly focuses on the third stage, more specifically: 1) Extraction of textual information from the scanned certificate image; 2) Creation of data structure from textual information; 3) Validation and verification from academic certificates and certificate holder (via the corresponding public blockchain address on which the certificate is registered). It is thus the aim of this research to determine whether the proposed solution, will improve the verification and validation process needed by academic institutions and/or employers. To implement the prototype several questions had to be answered beforehand: 1) What machine learning techniques

can allow the extraction of textual information from scanned certificates? 2) How will the system handle different document layouts? 3) How will the blockchain be used to verify and validate academic certificates and the award holders?

The first step is extraction of textual information from the scanned certificate image, and for this we have opted to use Tesseract OCR. This is because, several research, [1] and [3], have shown that OCR, more specifically Tesseract OCR, is able to extract textual data contained within images of documents at a good accuracy level of 84% with Times New Roman font. One major problem encountered was that whenever logos were present in the academic certificates, the Tesseract was producing very inaccurate results. Some academic certificates prove to be problematic, due to having objects being unrecognised as letters or symbols, or due to having very small lettering. To address this problem, we have opted to use the OpenCV library to perform multi-scale template matching in order to identify the locations of the logos such that they can be removed by setting the logo area to white.

For the solution to be scaleable, a repository holding a large number of logos is required, such that when any party starts the certificate verification process, the application performs a sequential pattern matching and logo removal process. We have given the OCR different academic certificates having different text, image quality, noise levels, designs and also used different devices for scanning, being mobile devices and dedicated scanners, such that we could qualitatively analyse and identify limitations when opting for such a system. After having obtained a clean version of the academic certificate and having performed OCR on the academic certificate, the next task was to extract key information and organise it into a data structure.

In order to extract important information from an OCR result, we have opted for regular expression patterns (RegEx). Every institution uploads their own template files, which must match the certificates they will be uploading. Since there will be no standard certificate layout, the system will not know which information is important and which information is redundant on its own, therefore tags were used. The tags are customisable by the institution, however special non-customisable tags exist such as institution_name, award_holder_name, day, month, and year. These tags are then converted into named group capture RegEx pattern strings, such that useful information can be extracted with these tags. During extraction, a value representing the similarity between the template and the text is measured by using the difflib python library, which makes use of the Gestalt pattern matching technique, such that, the output with highest similarity is selected to be the correct output. In order to test this feature we have created multiple templates files for different awarding organisations in order to analyse how rigid pattern matching is when combined with OCR.

After having a complete dictionary, the next task is to verify and validate the certificate holders public address and the information found within the scanned certificate. This is needed so that an individual A, does not impersonate an individual B and claim the latter's certificates as one's own, especially since we are dealing with a public blockchain network. The certificate holders public address

has to be verified using the CertificateChain smart contract while an academic certificate requires the academic organisations smart contract for verification. In order to conduct this research, we have chosen to develop and deploy multiple smart contracts, using the Truffle framework and Ganache. The CertificateChain smart contract acts as the main smart contract for this solution, as awarding organisations and third parties both make extensive use of this smart contract for registration, verification and validation. Several other smart contracts have been developed to act as smart contracts deployed, by fictitious academic organisations, with the purpose of storing academic certificate information from the respective academic organisation.

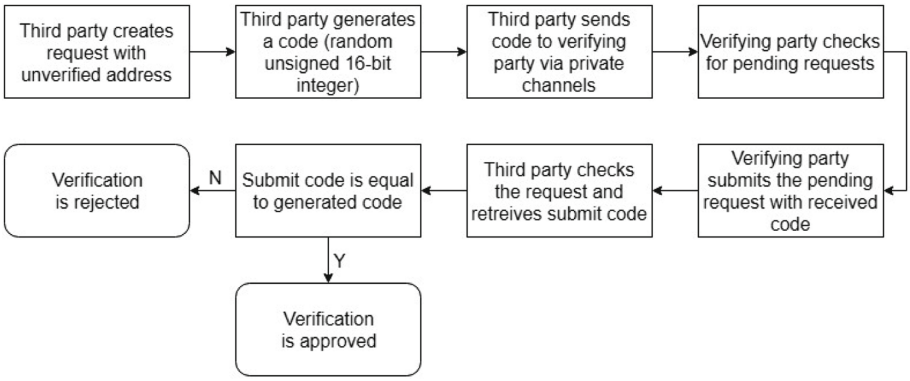

Fig. 1. Address verification process

In order to verify that the certificate holder truly owns a given address, we have adapted a solution similar to [3], which pipeline can be seen in Fig. 1, and involves the following steps: 1) 3^{rd} party create a confirmation request on the CertificateChain smart contract; 2) 3^{rd} party generates a random number between 1 and 65,535 and communicates this privately with the certificate holder; 3) the certificate holder logs-into the platform with his public key and proceeds to validate his/her pending request with the given randomly generated number; 4) 3^{rd} party checks the status and code of the confirmation response and if the codes match, the certificate holder is trusted to be the true owner of the given address. The next step is to validate the certificate information from the previously created dictionary, however, some issues had to be evaluated beforehand: 1) The current version of Solidity, version 0.5.7, does not allow functions to return an array of structures; 2) Validation depends on the template keys chosen by the academic organisation. Since the template keys are different for every institution, every smart contract is required to expose a pure function which returns an array of strings representing the list of template keys, and also, since we are not able to get the list of certificates belonging to an address with one call, the smart contract is also required to expose a function which returns the number

of certificates belonging to a specific address. The idea is that after validating the identity of the address received, the application gets the keys from the smart contract and gets the number of certificates owned by the address. Afterwards, the application starts to get the certificates owned by the address in a sequential manner and match the information stored on the blockchain with the information obtained by the OCR, giving every certificate a similarity score, very similar to how we choose a template. The certificate with the highest similarity, given that the highest similarity score is higher than a pre-determined threshold level, is shown on screen such that the third party can make some final checks before making their decision.

4 Results

4.1 Certificate Image Pre-processing and OCR

The removal of logos found within academic certificate images uses a multi-scale template matching approach. This is because OpenCV template matching requires the template to be very similar to a section within an image, therefore, if the size of the template does not match the size of the logo found within the image, the template matching operation could fail to operate as intended. Also, another limitation with template matching is that the operation will return a region in the image with highest similarity to the template, however, this does not always mean that the region is correct, which means that performing the operation with a template which is not found within the image still returns a region. The orientation of the certificate image must match the orientation of the logo for template matching and must be upright for the OCR to produce a valid result. Using different scanning devices, being dedicated scanners and mobile, made little difference in the result, as long as the certificate image result is clear for the OCR to process. Most certificates having a hand signature generated an invalid result when performing OCR due to two reasons being: 1) The signature overlapped some of the characters, thus the OCR could not identify properly the character; 2) The signature was being misread as a character to process.

When analysing the increase in time taken (in seconds), for the OCR to process certificates with different word counts, and having 300 dots per inch (dpi), a strong linear relationship can be observed, having correlation value of 0.95, thus, an increase in word count causes an increase in the time taken. Scalability is an issue with the proposed solution because each logo removal will take on average 4.32 s, thus, if each logo has to be stored in a repository, approximately every 830 logos stored will increase the time taken to automatically remove logos by one hour. A linear relationship is also present between the pixel count of the template and the time taken to perform logo removal, having correlation of 0.85. Possible solutions to such limitations shall be addressed in the final section.

4.2 Pattern Matching

To extract the useful information from the OCR result, we firstly need to identify the institution the certificate belongs to, which is done by performing a linear

search for all registered institutions in the OCR result, thus scalability is also a problem in the event of having a large number of academic institutions registered. One major limitation with the proposed solution is that this stage depends on an API and database to retrieve the institutions and respective template files, therefore in the events of having the service go down, the validation process is halted until services are back online. The RegEx patterns in the template file had to be an almost perfect match, which means that in the event of having the OCR read an extra white space or spell something incorrectly, some of the information to be extracted could not be extracted. This problem can partially be solved by creating more elaborate RegEx patterns which caters for extra or missing white spaces.

4.3 Smart Contracts

To validate the certificates on the blockchain, the CertificateChain smart contract and our institution smart contracts were deployed which require money to pay for gas used by the Ethereum virtual machine (EVM). This solution has been developed solely on Ethereum not because there are technical limitations not met on other networks, but purely as a proof of concept, which can easily be migrated to other. The gas usage for CertificateChain is of 2,309,099 whilst for two local Universities renamed as A and B is of 1,853,264 and 1,658,387 respectively based on their respective certificate data. The costs are found in Table 1, this shows the low initial cost for deploying the smart contract on the blockchain.

Table 1. Costs in ETH and EUR for deploying prototype smart contracts

	ETH costs		
Gas price (Gwei)	CertificateChain	University A	University B
15	0.034636 (€4.95)	0.027799 (€3.97)	0.024876 (€3.56)
10	0.023091 (€3.30)	0.018533 (€2.65)	0.016584 (€2.37)
4	0.009236 (€1.32)	0.007413 (€1.06)	0.006634 (€0.95)
2	0.004618 (€0.66)	0.003707 (€0.53)	0.003317 (€0.47)

EVM gas consumption uses Gwei, in which 1 ETH is equal to 1,000,000,000 Gwei, and the exchange rates are dated 2^{nd} of April 2019 16:00, in which 1 ETH is €142.93. Increasing the gas price will increase the speed of confirmation for the transaction, however, not all transactions should be created with a high gas price. In this case a lower gas price is ideal as both deployment of smart contract and publishing of academic certificates do not need to be done at instant speeds (Table 2).

When analysing gas fees for publishing 128-byte certificates according to number of graduates at University of Malta for the 2017–2018 scholastic year [12],

Table 2. Costs in ETH and EUR for publishing certificates with different sizes

	ETH costs		
Gas price (Gwei)	32 bytes/130,041 Gas	64 bytes/132,089 Gas	128 bytes/216,864 Gas
15	0.001951 (€0.28)	0.001981 (€0.28)	0.003253 (€0.46)
10	0.001300 (€0.19)	0.001321 (€0.19)	0.002169 (€0.31)
4	0.000520 (€0.07)	0.000528 (€0.08)	0.000867 (€0.12)
2	0.000260 (€0.04)	0.000264 (€0.04)	0.000434 (€0.06)

it was observed that the difference in price between 2 Gwei and 15 Gwei is €1,420.01. Gas fees for publishing 128-byte certificates is shown in Table 3 according to the number of tertiary education graduates per EU country in 2016 [4]. As can be seen in the results, the gas price makes a critical difference in the total costs. Case in point is the €2,297,168.14 difference between 2 Gwei and 15 Gwei. Having a high gas price is unnecessary for publication of certificates as these do not need to be available within seconds of graduation, thus, if a large number of certificates needs to be published, it is best to plan ahead of time and stage publishing with lower gas prices in order to avoid unnecessary costs.

4.4 Verification and Validation of Certificates and Addresses

To verify the owner of an address, a transaction must be made by both parties, therefore gas consumption must be paid to verify securely on the blockchain. The proposed solution uses 156,278 gas to create a request, which is 0.00067 ETH (€0.09), while 48,884 gas is used to confirm the request, which is 0.00021 ETH (€0.03) assuming the gas price is 4.3 Gwei and 1 ETH is €142.93. The speed of this process depends on the gas price of the transactions, if the verification party requires faster confirmation, a higher gas price such as 20 Gwei can be set and in return, higher transaction fees. The proposed solution has been adapted from [11], however, instead of using an inbuilt token, such as the EduCTX token, we are using a smart contract transaction and only pay for gas consumption. As the address verification process has very low gas consumption, the gas consumption fee is not refunded. To verify and validate an academic certificate, no transactions need to be created, this process is free and performed instantly without confirmation time. This process makes use of multiple function calls to the academic organisations smart contract due to limitations in the Solidity language. The academic certificates belonging to a certificate holder need to be fetched individually because the language does not support functions to return an array of structures, therefore, this process takes $O(n)$ time where n is the number of academic certificates belonging to a certificate holder stored in the academic organisations smart contract. Both verification and validation processes are not immune to malicious activity. In the case of owner verification, multiple users can work together by sharing the private key such that they would verify the address for each other. The proposed solution for academic certificate verification and validation crucially depends on the certificate data being published by the

Table 3. Gas fees for tertiary education certificates per EU country in 2016

		Gas price (Gwei) for 128 byte certificates			
	Total	2	**4**	10	15
EU-28	4,695.980	€291,116.68	€582,233.36	€1,455,583.41	€2,183,375.11
Belgium	119.141	€7,385.88	€14,771.75	€36,929.39	€55,394.08
Bulgaria	60.383	€3,743.31	€7,486.62	€18,716.54	€28,074.81
Czech Republic	90.725	€5,624.29	€11,248.58	€28,121.46	€42,182.19
Denmark	85.290	€5,287.36	€10,574.72	€26,436.81	€39,655.21
Germany	556.800	€34,517.56	€69,035.12	€172,587.80	€258,881.69
Estonia	10.262	€636.17	€1,272.34	€3,180.85	€4,771.27
Ireland	65.362	€4,051.97	€8,103.94	€20,259.85	€30,389.77
Greece	69.929	€4,335.09	€8,670.18	€21,675.45	€32,513.18
Spain	438.661	€27,193.80	€54,387.60	€135,968.99	€203,953.49
France	772.779	€47,906.69	€95,813.38	€239,533.45	€359,300.17
Croatia	34.028	€2,109.49	€4,218.98	€10,547.45	€15,821.17
Italy	373.775	€23,171.34	€46,342.68	€115,856.69	€173,785.03
Cyprus	8.420	€521.98	€1,043.96	€2,609.89	€3,914.84
Latvia	15.796	€979.24	€1,958.47	€4,896.19	€7,344.28
Lithuania	29.683	€1,840.13	€3,680.26	€9,200.65	€13,800.98
Luxembourg	1.682	€104.27	€208.54	€521.36	€782.04
Hungary	68.110	€4,222.33	€8,444.65	€21,111.63	€31,667.44
Malta	4.576	€283.68	€567.36	€1,418.39	€2,127.59
Netherlands	148.942	€9,233.32	€18,466.65	€46,166.62	€69,249.92
Austria	83.396	€5,169.95	€10,339.89	€25,849.73	€38,774.60
Poland	487.640	€30,230.14	€60,460.28	€151,150.71	€226,726.06
Portugal	73.086	€4,530.80	€9,061.60	€22,654.01	€33,981.01
Romania	121.788	€7,549.97	€15,099.94	€37,749.86	€56,624.79
Slovenia	30.967	€1,919.73	€3,839.46	€9,598.65	€14,397.97
Slovakia	56.280	€3,488.95	€6,977.90	€17,444.76	€26,167.14
Finland	56.066	€3,475.69	€6,951.37	€17,378.43	€26,067.64
Sweden	78.112	€4,842.38	€9,684.75	€24,211.89	€36,317.83
United Kingdom	754.301	€46,761.19	€93,522.38	€233,805.94	€350,708.91
Iceland	4.564	€282.93	€565.87	€1,414.67	€2,122.01
Liechtenstein	0.191	€11.84	€23.68	€59.20	€88.80
Norway	49.010	€3,038.26	€6,076.53	€15,191.32	€22,786.98
Switzerland	87.479	€5,423.06	€10,846.13	€27,115.32	€40,672.97
Macedonia	10.465	€648.75	€1,297.51	€3,243.77	€4,865.66
Serbia	50.326	€3,119.85	€6,239.69	€15,599.23	€23,398.85
Turkey	802.822	€49,769.14	€99,538.28	€248,845.69	€373,268.53
	5,700.837	**€353,410.52**	**€706,821.05**	**€1,767,052.62**	**€2,650,578.92**

academic organisations and cannot prevent such organisations from publishing false certificates in the first place, thus, if the academic organisation publishes false certificates, the proposed solution will identify it as a valid certificate.

5 Conclusion

We concluded that the proposed solution will improve the verification and validation process needed by academic institutions and/or employers. Using the proposed blockchain solution, academic institutions have the ability of having their academic certificates published on the blockchain network, having full freedom on the implementation of their smart contract. Institutions are required to pay gas fees in order to deploy their smart smart contracts, this fee depends on the size of the smart contract, however for our prototype the average cost for deploying smart contracts for the institutions was found to be €0.50 with gas price set to 2 Gwei. Gas fees also have to be paid when publishing new certificates onto the blockchain, and it was found to be €353,410.52 and €2,650,578.92 when publishing certificates to all 2016 tertiary education graduates in Europe when selecting 2 Gwei and 15 Gwei as gas prices respectively. In this case, transaction speed is not important, thus, low gas prices can be selected in order to cut a lot of extra costs. Since the proposed solution makes use of the blockchain network as the primary storage medium, the platform is immune to corruption and unauthorised alterations due to advanced cryptographic techniques. From these findings we have concluded that the proposed solution is very cost effective, when selecting lower gas prices, for the security benefits offered.

However, several scalability issues, discussed in the previous section, are found with the proposed solution, which need to be improved upon before implementing such a platform. The logo removal process takes approximately 4.32 s per logo, thus, approximately every 830 logos added to the logos repository will increase the time taken by one hour. From this finding we have concluded that template matching is not ideal as this requires us to store every logo as a template. One possible solution for this would be to train a neural network to identify the position of logos from the given image. The logo removal process also depends on the pixel count of the template image, such that the correlation between the linear relationship of the pixel count and time taken is 0.85. One possible solution for this issue would be to downscale the template and certificate images, such that the time taken is reduced.

References

1. Ambadiyil, S., Vibhath, V.B., Pillai, V.P.M.: On Paper digital signature (OPDS). In: Thampi, S., Bandyopadhyay, S., Krishnan, S., Li, K.C., Mosin, S., Ma, M. (eds.) Advances in Signal Processing and Intelligent Recognition Systems, pp. 547–558. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-28658-7_46
2. Curmi, A., Inguanez, F.: BlockChain based certificate verification platform. In: Abramowicz, W., Paschke, A. (eds.) BIS 2018. LNBIP, vol. 339, pp. 211–216. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-04849-5_18

3. Dlamini, N., Mthethwa, S., Barbour, G.: Mitigating the challenge of hardcopy document forgery. In: 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), pp. 1–6. IEEE (2018)
4. Eurostat: Tertiary education statistics (2018)
5. Grech, A., Camilleri, A.F.: Blockchain in education (2017)
6. Jiang, F., Zhang, L.J., Chen, H.: Automated image quality assessment for certificates and bills. In: 2017 IEEE International Conference on Cognitive Computing (ICCC), pp. 1–8. IEEE (2017). https://doi.org/10.1109/IEEE.ICCC.2017.8
7. Learning Machine: The Republic of Malta: A better pathway (2017). https://www.learningmachine.com/customer-story-malta/
8. Puthal, D., Malik, N., Mohanty, S.P., Kougianos, E., Yang, C.: The blockchain as a decentralized security framework [future directions]. IEEE Consumer Electron. Magazine **7**(2), 18–21 (2018). https://doi.org/10.1109/MCE.2017.2776459
9. Schneier, B.: Applied Cryptography: Protocols, Algorithms and Source Code in C, 20 anniversary edn. Wiley, New York (2015)
10. Sharples, M., Domingue, J.: The blockchain and kudos: a distributed system for educational record reputation and reward. In: Verbert, K., Sharples, M., Klobucar, T. (eds.) European Conference on Technology Enhanced Learning, pp. 490–496. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45153-4_48
11. Turkanović, M., Hölbl, M., Košič, K., Heričko, M., Kamišalić, A.: EduCTX: a blockchain-based higher education credit platform. IEEE Access **6**, 5112–5127 (2018)
12. University of Malta: Number of students who completed awards (2017–2018) (2018)
13. Yahya, Z., et al.: A new academic certificate authentication using leading edge technology. In: Proceedings of the 2017 International Conference on E-commerce, E-Business and E-Government, pp. 82–85. ACM (2017)
14. Yli-Huumo, J., Ko, D., Choi, S., Park, S., Smolander, K.: Where is current research on blockchain technology? a systematic review. PLoS ONE **11**(10), e0163477 (2016)