An Innovative Approach of Verification Mechanism for both Electronic and Printed Documents

Md. Majharul Haque[1], Abu Sadat Mohammad Yasin[5], Muhammad Shakil Pervez[6]
Bangladesh Bank, Dhaka, Bangladesh

Mohammod Akbar Kabir[3]
Department of Economics, University of Dhaka, Dhaka, Bangladesh

Md. Nasim Adnan[2]
Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore-7408, Bangladesh

Mohammad Rifat Ahmmad Rashid[4]
University of Liberal Arts Bangladesh

***Abstract*—Documents are inevitably relevant in our day-to-day life. Forgery of document could have severe repercussions including financial losses, misjudgments, damage of respect, goodwill, etc. Hence, documents need to be secured from threats such as counterfeiting, falsification, tempering etc., and there should be an easy way of verification about the originality of documents. There are several existing methods for ensuring authenticity and integrity with modern technologies like the block chain, Digital Signature, etc. Most of the methods are not appropriate for public usage instantly due to their intricacy, excessive costing, and implementation problem for which the easy approach of verification is yet not available for mass people. In this situation, a method of document verification has been proposed in this paper which intends to provide (i) authenticity, (ii) integrity, (iii) availability, and (iv) non-repudiation. The proposed method will serve the purpose of mass people as it has no licensing fee, easily implementable and effortlessly useable for both electronic and printed documents. It is worth to mention that the proposed method will provide a mechanism to confirm the originality of the document using only a Smartphone in no time.**

***Keywords*—Document verification; integrity; non-repudiation; blockchain; printed documents**

I. Introduction

The electronic document is much more flexible to use and communicate; however, till date we cannot omit the necessity of printed document. A wide variety of legal documents that includes national identity cards, social security cards, birth certificates, driving licenses, passports, insurance documents, educational documents, wills, power of attorney, and land titles are still needed to be in the printed form for day to day life [1]. Modern technologies opened a window for mass people to exchange information in a simpler and faster way without any hassle. The extensive growth of Information and Communication Technology (ICT) has facilitated the generation, distribution, and reproduction of huge amount of electronic or printed documents effortlessly. With the help of cheap hardware and easy-to-use digital imaging software, any person can counterfeit a legal document at his/her convenience. In this way, the benefits of ICT for the electronic and printed documents come together with problems and threats associated with ensuring digital copyright protection, preventing digital counterfeiting, proof of authenticity, and content-originality verification [2].

There are a lot of organizations around the world providing several types of approval letters, certificates or other kinds of legal documents. In most of the cases, those documents are needed to be verified by various types of agency. For example, students are applying for new degree, scholarship or job to any native or foreign institute. After the application, certificates, transcripts, etc. are ubiquitously needed to be verified. Current verification processes, either automated or manual, comes with tremendous hassle, and spending of lots of money and precious time. Even, some of the verification processes are not always reliable. An important point to be mentioned that electronic documents are generated after approval of corresponding authority in a software system in a production server. Usually, a production server is kept highly secured; whereas development and test servers may not be as secured. In development or test servers, there can have provision to generate the same document which will be a simple bypass of the legal process. Printed documents can more easily be counterfeited with a scanning machine and using Photoshop or other popular software. We are aware that some doctors and teachers have been caught red handed with fake certificate after decades of their service. Hence, in this tough world, accurate verification of any electronic or printed document is extremely required as an instantaneous service for the people of all classes.

Till date, a number of techniques have been proposed for seamless document verification [1], [2]. Among them, digital watermarking techniques intend determine the ownership and integrity for digital content [3]. Watermarking techniques intend to detect noise introduced in a document while printing.

Amidror [4] presented an intrinsic security model which integrates print-based authentication and a security strategy for the protection of valuable documents. This technique is based on more intensity profiles and mathematical transformations of the microstructure of a document to generate an attractive, dynamic visual effect [3]. The author claimed that the counterfeiting of authenticity is difficult but the way of

verification from the level of mass people is almost impossible [4]. Garain & Halder (2008) proposed a framework for automatic authenticity verification. However, only a particular area of a document can be protected by using the method which can be usable for some specialized documents [5], [6]. Zhang et al. [5] proposed a method for protecting only pictures in documents by generating photo signatures based on the cosine transformation generated from the documents to validate their authenticity.

Eldefrawy et al. [7] introduced an approach to detect the differences between various document features produced using different printing technologies. Another approach was published by Mikkilineniet al. [8] to trace documents by comparing the printing characteristics of different printing devices. These methods fail to detect any forgery if a similar printer is used to reproduce the document [7], [8]. The most interesting aspect of the approach presented by Mikkilineniet al. [8] is that the security features introduced into a document to verify their authenticities are not themselves secure [3].

According to Nia et al. [9] and Gopal & Prakash [10], digital signature and blockchain can be a solution to ensure the security of electronic documents. More recently, two techniques [11], [12] have been proposed based on Quick Response (QR) codes. However, both of the techniques use digital signature to ensure the security of electronic documents. However, a soft copy of the document is needed at the user end for digital signature or blockchain verification in addition to a third-party involvement. Hence, verification through digital signature or blockchain is still at large for the mass people [13], [14].

A number of techniques on document protection can be found in literature based on digital watermarking, steganography, digital signature and blockchain, etc. It is important to mention that most verification procedures of those techniques are tedious and time-consuming as they often involve many legal entities across the globe [15]. Hence, it is evident that very little attention has been given on how to validate the authenticity and integrity of confidential documents effortlessly for the mass people.

Under these circumstances, in this paper we propose a simple yet sophisticated method to provide a verification mechanism for both electronic and printed documents. It is expected that the proposed method will be easily accessible for the mass people and will eliminate many hindrances of the document verification process.

The rest of the paper is organized as follows: Section II illustrates the proposed method in detail. Security potentials and benefits of the proposed method are discussed in Sections III and IV respectively. Finally, we present our concluding remarks in Section V.

II. Methodology

The proposed method intends to cover the authenticity, integrity, non-repudiation, and availability of a digital and printed document. The architecture is built as a client-server model where a client connects to the server and request for document generation. The entire method is described in the following five steps.

A. Requesting for Document Generation

Present your perspective on the issues, controversies, problems, etc. as they relate to theme and arguments supporting your position. Compare and contrast with what has been or is currently being done as it relates to the article's specific topic and the main theme of the journal.

1) There is a central server that has some necessary templates (with unique template' id) based on which document will be created. In this server, a client can be registered by an authorized agreement for a definite period through the Internet. After registration, the client will get a unique Id and password.

2) In the client-side database, there is a request queue table with the following fields:

a) Request Id (Auto-incremental unique Id).

b) Template Id (for the template to be selected in the server).

c) Application Id. (The specific application for which the client needs the approval letter).

d) Application Type Id. (There can be multiple types of applications).

e) JSON Data (Necessary information and template of the application that will be placed in the server).

f) Response from the server.

g) No_of_Try (The default value of this field is 0, which means it is yet to try for generating the document. The value is increased by 1 based on the number of trying. The value is set to -1 after 3 times trying and then no more trying).

h) Status (The value can be set as "success", "fail" or "rejected").

i) Response from the server (The entire response from the server is saved in JSON format)

3) The registered client requests for connection (step 1 of Fig. 1) to the server via an API (Application Interface) request with the Client Id and password. If the credential is found valid, the server responses with a unique token number (step 2B of Fig. 1). Then, a scheduler in client fetch one by one request from the queue where the "No_of_Try" value is not equal to "-1" and the "status" is not equal to "success" (success means the document is already generated successfully) and send to the server with a token number (step 3 of Fig. 1).

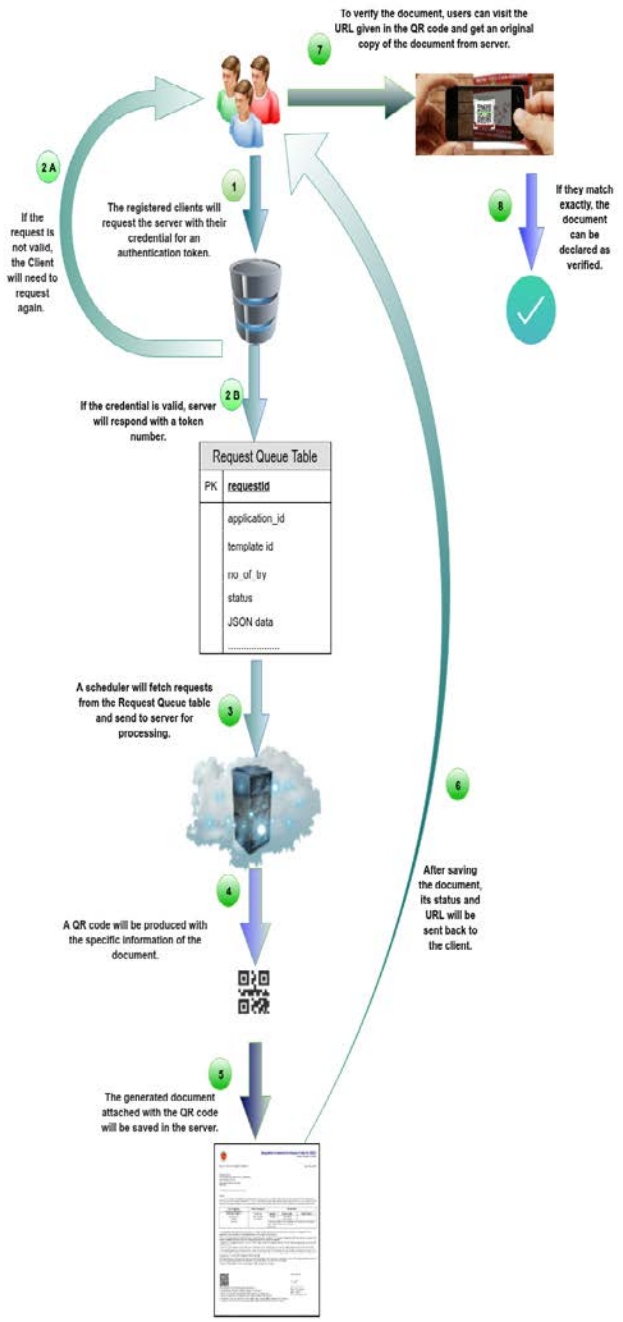

Fig. 1. Process Flow of the Proposed Method.

B. Creating Document

Present your perspective on the issues, controversies, problems, etc. as they relate to theme and arguments supporting your position. Compare and contrast with what has been or is currently being done as it relates to the article's specific topic and the main theme of the journal.

1) If the request is found valid from the Client Side, the corresponding template is selected in the server according to the Template Id. Then, the JSON formatted information (received by the server as an API request) is parsed and put in different predefined places of the selected template. Generally, there can be three types of server for any Document Management system as follows: (i) Development, (ii) User Acceptance Test (UAT), and (iii) Live. In the Development and UAT server, the generated document will have a watermark as "Development" and "UAT", respectively. In the Live server, the generated document will have no such kind of watermark.

2) The generated document is saved in the central server with a unique file name where the file name is set using a defined naming convention as lientID_TemplateID_DD_MM_YYYY_N.pdf. Here DD means two (02) digits of a day, MM means two (02) digits of a month and YYYY means four (04) digits of a year, N means an incremental number (Starting from 1 for each day), ".pdf" means the document is generated in Portable Document Format.

3) In the pdf file (generated document), a QR Code is attached which contains the full URL of the document. This URL is the same as the path of the file where this file will be saved in the server (step 4 and 5 of Fig. 1). If anybody hits the aforementioned URL, he/she will get the corresponding pdf file.

4) After generating the document, it is saved in a specific folder as APPROVAL_LETTER/CLIENT_ID/YYYY/MM where the APPROVAL_LETTER is a folder which contains several sub-folders named CLIENT_ID, then Year and then Month.

5) If we want to archive some files, we can easily decide that the files of which Client, Year and Month needs to be archived.

C. Providing Response from the Server:

After generating the document and saving in the central server, a response is given to the corresponding client with the following information in JSON format (step 6 of Fig. 1): (i) status (it can be either "success" or "fail". The status can also be rejected, if there is any mismatch in the token number), (ii) the URL of the document (it is given if the status is success otherwise it is left blank), (iii) the duration between a request received and the response provided.

D. Receiving Response by the Client from the Server:

1) The client parses the JSON formatted response from the server and saves the value of "status" field, URL of the document and the entire response in request queue table (mentioned in the second point of sub-section A) in the database (client-end).

2) Client updates the value of "No_of_Try" field in request queue table as "-1" if the status is "fail" or "rejected" after trying 3 times so that will not be tried again.

E. Verifying Document:

After generating the document and saving in the central server, a response is given to the corresponding client with the following information in JSON format (step 6 of Fig. 1): (i) status (it can be either "success" or "fail". The status can also be rejected, if there is any mismatch in the token number),

(ii) the URL of the document (it is given if the status is success otherwise it is left blank), (iii) the duration between a request received and the response provided.

1) For the verification purpose, one needs to collect the document in either printed copy or soft copy (pdf). If it is printed copy, it is needed to confirm that the QR code has been printed well.

2) Almost all modern people have a smartphone in this age. By using those devices, the QR code can be scanned from the document where a URL will be found (step 7 of Fig. 1). If we hit this URL, a document will be downloaded from the server if the URL is valid otherwise it will give "not found" error. Now, the verification is very simple as follows:

a) if there is any document downloaded by visiting the URL, we need to match the downloaded document and the document we have in hand. If both are 100% same, we can declare that the document is valid (step 8 of Fig. 1)

b) if there is any mismatch or there is "not found" error, the document can be regarded as a fake one.

III. Ensuring Security

Though we are emphasizing on document verification, some security features have also been covered with the proposed method. Four significant points have been covered here as (i) authenticity, (ii) integrity, (iii) non-repudiation, and (iv) availability. Authenticity is the assurance that a message, transaction, or other exchange of information is from the source it claims to be from [18]. Integrity refers to the document has not been distorted by anyone. Non-repudiation confirms that the authority cannot disown the document. Last of all, availability means the document can be retrievable at any time. By using the proposed method, the user can claim that the approval letter has been given from a specific authority/organization. Now, if the same document has been found in the given URL in the QR code (explained in the third point of sub-section B under section III) at any time, it ensures the four points for security issue. (i) Authenticity is confirmed in such a way that the source of document is verified by downloading the document from the URL of specific organization, (ii) If the document has been distorted by anybody then the newly downloaded document will not be same as the old document and hence the loss of integrity can be detected, (iii) The authority cannot disown the document because this can be downloaded from the URL of the authority's system that is ensuring non-repudiation, (iv) Since the document is obtainable at any time from the URL, availability is also ensured.

IV. Comparison with Latest Existing Methods

In Table I, we present a brief comparison between the proposed method and some existing methods.

TABLE I. Comparison between the Proposed Method and Some Existing Methods

Criteria	**Proposed Method**	AL-Gawda et al. [3]	Warasart & Kuacharoen [15]
Security	Very High	Very High	Very High
Counterfeit Detection Accuracy	Very High	Very High	Very High
Applicability	Both Digital and Printed Documents	Both Digital and Printed Documents	Both Digital and Printed Documents
Use of Digital Signature	No	Yes	Yes
External Dependency	Service Provider Server (expected to be online for 24×7	Digital Signature Certifying Authority	Digital Signature Certifying Authority
Licensing Fees	No	Yes	Yes
Cost	No Cost (No Overhead for the Mass People)	High Cost (Maintenance of Digital Signature for all Concerned Actors)	High Cost (Maintenance of Digital Signature for all Concerned Actors)
Verification Process	Easy (Low Computation Overhead)	Complicated (High Computation Overhead)	Complicated (High Computation Overhead)
Applicability for Mass People	Yes	No	No

V. Benefit of the Proposed Method

We believe that the proposed system has the following benefits by comparing with other systems like a digital signature or any central authority for document verification:

1) The proposed system can be developed by any organization where license from any other third party is not required.

2) User can easily get another copy of the document if a downloaded copy is lost or distorted.

3) If a user wants to give a copy of the document to another person anywhere, he/she just has to give the URL of the document.

4) Since the central server is API-based, it can be connected to multiple clients without any hassle.

5) The process of verification has almost no cost, with no process-oriented harassment and requires a very short time assist only requires visiting a URL.

VI. Conclusion

It is well-known that counterfeiting of any necessary official or unofficial document is a criminal activity which adversely affects the country, society, and even personal reputation. For eradicating this problem, a method of document creation and verification has been proposed here to ensure authenticity, integrity, non-repudiation, and availability. The entire mechanism of providing security and performing verification has been discussed here in detail including the benefits of the proposed method compared to others. The key strengths of this method are to design an efficient, secure, low-cost, and easily accessible solution for the people of almost all levels without compromising quality. We believe that this method will prevent the falsification of documents. In the arena of software development, it will also be tremendously useful to issue any certificate or approval letter. Detection of falsified documents can also be augmented using several data mining algorithms [16], [17], [18], [19].

References

[1] Van Renesse, R. L. (1997). Paper based document security-a review.In European Conf. on Security and Detection-ECOS97.

[2] Tayan, O., Kabir, M. N., & Alginahi, Y. M. (2014). A Hybrid Digital-Signature and Zero-Watermarking Approach for Authentication and Protection of Sensitive Electronic Documents. The Scientific World Journal, 2014, 1-14.

[3] AL-Gawda, M., Beiji, Z. & Nurudeen, M. (2015). Printed Document Authentication Using Two-Dimensional (2D) Barcodes and Image Processing Techniques.International Journal of Security and Its Applications, 9(8), 347-366.

[4] Amidror, I. (2002). New print-based security strategy for the protection of valuable documents and products using moiré intensity profiles. Electronic Imaging 2002 International Society for Optics and Photonics, 3677, 89-100.

[5] Garain, U.,&Halder, B. (2008). On automatic authenticity verification of printed security documents., 2008. ICVGIP'08. In Proceedings of the 2008Sixth Indian Conference on Computer Vision, Graphics & Image Processing(pp. 706-713).Washington DC, United States: IEEE Computer Society.

[6] Li, J., Zhang,X.,Liu,S.,& Ren,X. (2009). Adaptive watermarking scheme using a gray-level computer generated hologram.Applied optics, 48(26),4858-4865. generated hologram.Applied optics, 48(26),4858-4865.

[7] Eldefrawy, M. H., Alghathbar,K.& Khan,M. K. (2012). Hardcopy Document Authentication Based on PublicKey Encryption and 2D Barcodes.In Proceedings of the 2012 International Symposium on Biometrics and Security Technologies (pp. 77-81). Washington DC, United States: IEEE Computer Society.

[8] Mikkilineni, A. K., Chiang, P.J.,Ali, G. N.,Chiu,G. T.,Allebach, J. P.,& Delp, E. J., (2005). Printer identification based on graylevel co-occurrence features for security and forensic applications. In Proceedeings ofthe conference on Security, Steganography, and Watermarking of Multimedia Contents VII (Vol.5681, pp. 430-440). California, USA: International Society for Optical Engineering.

[9] Nia, M.A., Sajedi, A., & Jamshidpey, A. (2011). An Introduction to Digital Signature Schemes.In Proceedeings of National Cpnference on Information Retrieval. Retrieved from https://arxiv.org/abs/1404.2820.

[10] Gopal, N., & Prakash, V.V. (2018). Survey on Blockchain Based Digital Certificate System. International Research Journal of Engineering and Technology,5(11), 1244-1248.

[11] Schultz, Michelle Kelly. "A case study on the appropriateness of using quick response (QR) codes in libraries and museums." Library & Information Science Research 35, no. 3 (2013): 207-215.

[12] André, P.S. and Ferreira, R.A.S., 2014. Colour multiplexing of quick-response (QR) codes. Electronics Letters, 50(24), pp.1828-1830.

[13] Avram, M. G.: Advantages and Challenges of Adopting Cloud Computing from Enterprise Perspective. Procedia Technology, Vol. 12(1), pp. 529–534 (2014).

[14] Li, J., Zhang,X.,Liu,S.,& Ren,X. (2009). Adaptive watermarking scheme using a gray-level computer generated hologram.Applied optics, 48(26),4858-4865. generated hologram.Applied optics, 48(26),4858-4865.

[15] Warasart, M., Kuacharoen, P. (2012). Paper-Based Document Authentication Using Digital Signature and QR Code.In Proceedings of the 4th International Conference on Computer Engineering and Technology(pp. 94–98). Phuket, Thiland: ASME Press.

[16] Adnan M. N.: On Dynamic Selection of Subspace for Random Forest. The 10th International Conference on Advanced Data Mining and Applications (ADMA 2014), pp. 19 – 21 (2014).

[17] Adnan M. N., Islam M. Z.: ComboSplit: Combining Various Splitting Criteria for Building a Single Decision Tree. The International Conference on Artificial Intelligence and Pattern Recognition (AIPR), pp. 17 – 19 (2014).

[18] Adnan, M. N., Islam, M. Z.: ForEx++: A new framework for knowledge discovery from decision forests. Australasian Journal of Information Systems, Vol. 21, pp. 1–20 (2017).

[19] Adnan, M. N., Islam, M. Z.: Forest PA: Constructing a decision forest by penalizing attributes used in previous trees. Expert Systems with Applications, Vol. 89, pp. 389–403 (2017).