2020 IEEE Region 10 Symposium (TENSYMP), 5-7 June 2020, Dhaka, Bangladesh

Secure and Transparent KYC for Banking System Using IPFS and Blockchain Technology

Abdullah Al Mamun*, Sheikh Riad Hasan†, Md Salahuddin Bhuiyan‡, M. Shamim Kaiser§ and Mohammad Abu Yousuf ¶
Institute of Information Technology, Jahangirnagar University
Savar, Dhaka-1342, Bangladesh
Email: *abdullah.iiuceee@gmail.com, †riadzor@gmail.com, ‡mdeee15@gmail.com, §mskaiser@juniv.edu, ¶yousuf@juniv.edu

***Abstract*—The know your customer or know your client (KYC) is a guideline for the banking system to validate a customer using identity, appropriateness, risk assessment in establishing a banking relationship. With the growing concern of security, the KYC process is complex and involves a high cost for completing for a single customer. In this work, we propose an economical, swift, secure, and transparent platform for KYC document verification for the Banking system through InterPlanetary File System (IPFS) and blockchain technology. The proposed system allows a customer to open an account at one Bank, complete the KYC process there, and generate a hash value using the IPFS network and share it using the blockchain technique. Upon receiving the private key, any Bank/financial organization can retrieve, store customer data (i.e., KYC) securely using IPFS network if the customer wishes to open another account in that Bank/financial organization. The proposed system can save time, money, and repetitive work during the KYC process when someone tries to open an account at multiple Banks.**

***Index Terms*—Blockchain, Smart Contact, KYC, IPFS, Gpg4win, Decentralization**

I. Introduction

The Know your customer (KYC) is a very common term in the banking and financial sector. At this moment, the manual KYC process is outdated and has become a necessity to automate the KYC verification process. Studies around the world have made several attempts to make a better verification process for KYC. Many academics tried to propose a Blockchain-based solution. Blockchain technology recently draws the attention of the public, as a dispute that leads to the foundation that the trust-free economical transaction is possible with its distinctive method [1], [2].

The blockchain permits unnamed and secure transactions of virtual currencies (such as Bitcoin, Litecoin, etc) and saves the metadata regarding the transaction details in a database. The database is secured and impede the alteration in the transaction history by cryptography techniques. The legitimate user can write to the file using the private key. In banking, blockchain is safe and can reduce processing/transaction costs considerably. The banks or other financial organizations such as insurance industries maintain diverse policies and require multi-steps processing between parties. Besides, these require a secure transaction, short processing/settlement time. To facilitate these concerns, the researcher has proposed various distributed platforms. Raikwar et al. [3] proposed a blockchain based distributed platform for financial transaction processing in insurance industry. Puthal et al. [4] introduced a decentralized framework using blockchain which allows sharing and integration of all distributed actors. This will help industry to analyze the spread and plan further development [5].

Ever since Satoshi Nakamoto exited the scene and handed over Bitcoin development to other core developers, the digital ledger technology has evolved resulting in new applications that make up the blockchain History

Nakamoto [6] proposed a e-transaction system of coin produced using digital signatures. The system is able to track the transaction history, and it can prevent double spending problem. Since then, researchers are trying to find the potential sectors to apply the Blockchain. Nevertheless, sharing transaction information over bitcoin in costly. Currently, miners are charging around $7 per 100 KB of data [7].

The uses' KYC documents cannot be uploaded to the Blockchain network as it will be expensive. Thus, as an alternative solution, KYC documents sharing using the InterPlanetary File System (IPFS) is proposed in this paper, and then documents are shared over the Blockchain network. The IPFS is a shared dispersed document framework that looks to associate all registering computing devices with a similar system of files [8]. User can store their transaction history and hash to the IPFS network, and then stare it to the Blockchain network when required. This process will reduce the blockchain data size significantly [9].

The rest of the article is written as follow– section II depicted the literature review, section III discussed technologies used to formulate the framework; Section IV proposed a framework. Section V and section VI talked about the result section and conclusion respectively.

II. Literature Review

The KYC verification process is an integral part of regulation for the financial industry [10]. The KYC process is started when a client wants a financial transaction with a financial institution to begin [11]. Arasa et al. [12] direct an investigation of the expense of KYC dependent on the complexity level of the compliance required for the instance

978-1-7281-7366-5/20/$31.00 ©2020 IEEE

of business banks in Kenya, building up to four variables that clarify 78.3% of the consistence necessities. The data around us is increasing day by day, including the KYC documents. Soni and Duggal [13], proposed a solution using big data analytic techniques to solve the big data problem of KYC focused on Indian banks.

Y. Lootsma et al. [14] proposed to implement the Regtech (regulatory technology) like Blockchain in the banking sector to reduce the burden of the KYC process for a financial institution as well as the regulatory institution. Using the approach tax reporting can also be done. However, they did not show the full implementation of Blockchain and the cost involved with the process.

When a client wants to do the financial transaction through a payment provider, they will check the customer identity by his name from Bank if the provided information is correct through Blockchain smart contract [15]. The author provided an assumption on using the blockchain to make the identity and financial transaction through blockchain, though they did not provide any use case for document sharing like KYC docs.

A typical KYC framework could be that a client goes to a bank, the bank performs KYC, stores KYC in the Blockchain, give a customer a token and then customer give access to another bank to check the KYC information. The other bank then crosschecks the information from Blockchain [16]. Because of a range of configuration parameters, the blockchain is somewhat uncontrollable. For example, testnets like Rinkby, Ethereum cannot be adjusted easily because of their parameters like Gas limit, Mining difficulty and so on. Authors suggested using Grid'5000, as they found it highly controllable and reconfigurable testbed. Again, the authors did not provide a practical use case scenario with cost calculation.

J. Parra Moyano et al. [12] has shown the design of centralized and decentralized Blockchain KYC solution with the division of processing cost among different banks. To minimize the cost of core KYC verification and improve the customer experience, they proposed a new scheme based on distributed ledger technology (DLT). They Focused on four main points. The first is proportionality: the cost will be shared proportionally by all the institutions involved with a particular KYC verification process. They focused on Irrelevance secondly. The one who avoids the KYC process will not get any incentive. The third point of focus was Privacy. The KYC verification process has to be secured so that user privacy is not violated. Finally, they focused on No-minting. As the process is online-based, they need to focus that no false can be made during KYC verification. Whenever someone tries to edit any portion of KYC data, that editing process will automatically be void from the authoritative side. Their proposal was much effective except a few problems, which are as follows:

- The block data size will increase over time, and hence the cost involving it.
- If a customer open account only at a single bank, then the whole cost has to bear by that bank.

III. Background Technologies

A. Blockchain

Blockchain is a purely peer-to-peer version online transaction, where a customer directly sends money to others without the help of a financial organization. All transactions will be hashed to an ongoing proof-of-work chain. Each of this called a block; the running block contains the hash of previous all blocks. Therefore, the whole process is tempered proof, as a single peer cannot add a new block without the proof-of-work [6]. Bitcoin was the first fully decentralized cryptocurrency. While the central purpose of DLT was to create digital money and sending and receiving via the Internet, the technology can also be used to authenticate online document sharing using smart contracts. The smart-contract aims to involve them in properties, which are expensive and controlled in a digital manner [17].

The European Security and Markets Authority [18], discussed the possible benefits, challenges to those benefits and limitations of DLT applied to the security market. While they only focused on the security market, they provide a guideline to apply the DLT in another financial sector like Bank.

B. IPFS

The InterPlanetary File System (IPFS) is a peer-to-peer(P2P) file-sharing protocol connecting computing devices for sharing/storing files/data. The content is uniquely recognized in the global namespace using the hash code of the file. If the hash code is altered, the data can not be verified which will be identified by IPFS. Besides, IPFS identifies duplication if files with the same content are stored.

Among many, AFS [19] has succeeded widely and still used by many today. Most especially, Napster, LimeWire, Gnutella, KaZaA, and BitTorrent are unstructured distributed P2P file/content-sharing protocol used by more than 100 million 100 million concurrent users.

In other way, IPFS follows client-server model which pointed out how do we access the web [20].

C. Gpg4win

Pretty Good Privacy is a cryptography (especially privacy and authentication) approach that ensures the security of file directory, text, email, and the whole disk utilizing the digital signature. It permits the integrity-non-repudiation-authenticity of files and email.

GNU Privacy Guard for Windows, in short Gpg4win, is an open-source encryption package for email/file in Microsoft windows environment which utilizes GnuGP cryptography [1].

IV. Proposed Framework

A. Proposed Workflow

A proposed workflow for sharing KYC docs using IPFS shown in Fig. 1. The workflow uses the example of a customer approaching two different banks. In the first stage, the customer went to Bank A to provide the Bank with his KYC

[1] https://www.gpg4win.org/

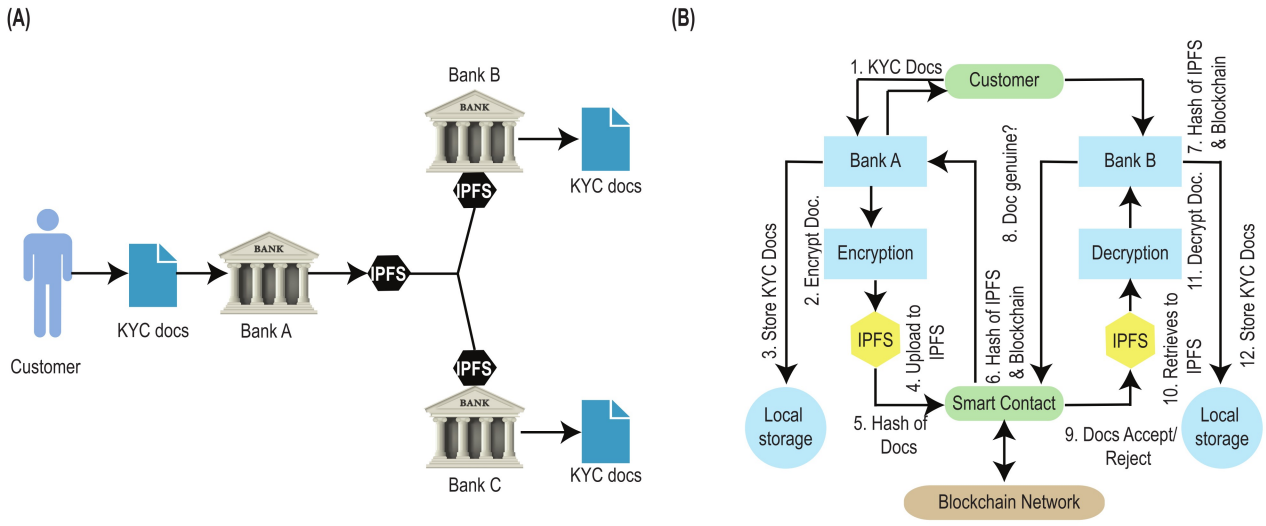

Fig. 1. (A)The workflow of KYC docs sharing using IPFS. (B) Block Diagram of Proposed KYC Solution

document for verification. Bank A goes through our proposed system design and provides the customer with a hash value and a personal decryption key. The customer will then go to Bank B and Bank C with these two keys, and both banks will testify the KYC doc in a blink. We used the IPFS network to upload and retrive KYC doc at banks end. Nevertheless, before sharing out KYC docs into the IPFS network, we considered encrypting the file for extra security and reducing file size. Since anyone can access the KYC docs from the IPFS network by just knowing their hash values. We used popular encryption software gpg4win in the Kleopatra platform so that people will have an encrypted docs of KYC.

B. Proposed Block Diagram

The proposed design for the KYC solution is shown in Fig. 2 using the example of a customer visiting two traditional banks. We considered a scenario in which, in the first phase, a customer went to Bank A to open an account. The customer submitted the account information along with KYC docs to the Bank. The bank then observed the whole information, which, if found correct, will be encrypted using the system's application (a popular encryption tool, gpg4win, and IPFS in our case) which will be available to all banks to share documents with other bank and store a copy to a local database. Afterward, the encrypted file will be stored in the private IPFS network by bank A. Later, the bank will upload the hash value from IPFS, a very small in-memory size, to the Blockchain network. Bank A also keeps a copy of the customer's KYC docs to the local database of it. Finally, Bank A will share the hash value of Blockchain and IPFS to the customer. Later on, the customer can give access to the KYC doc package by just sharing the hash value to any other institution he intended to work with.

However, now the customer can go to another bank to open another account. The customer will share his hash value from IPFS, and Blockchain to Bank B. Since Bank B will be granted access to the hash value of the document package by the customer, the Bank will get access to the Blockchain network for the required hash value. Subsequently, the bank will download the encrypted KYC docs from the IPFS network using the hash value retrieved from Blockchain. Lastly, with the help of the private key of the customer, the Bank will retrieve the KYC docs and keeps a copy of KYC docs to the bank's local database. The regulatory bank defined in the proposed solution in the central bank.

V. Results and Discussion

We tried to implement the KYC docs verification and making available to a financial institution with the help of IPFS and Blockchain network. We showed the scenario where a customer went to a bank to open an account. The bank encrypts the verified KYC docs using Gpg4win and then upload the docs in the IPFS network shown in fig. 2.

The same customer then went to another bank to make a financial transaction where he supposed to go through the same process over again. Yet fortunately, with the hash keys he received from the first bank, in our case, he retrieves the hash from the Blockchain network and retrieves the KYC docs from the IPFS network shown in fig. 3.

However, regardless of the approach, we have chosen to work with be it private or public Blockchain, our findings suggest various opportunities of increasing the efficiency of the existing financial system. More importantly, a notable reduction in cost during KYC document verification for participating institutions and less hassle for the customer could ensured by such a platform. Furthermore, the proposed system will ensure money and time saving, and efficient financing for financial organizations.

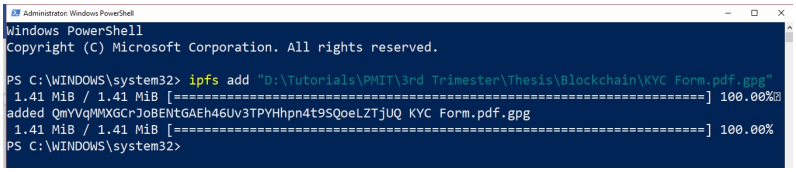

Fig. 2. Uploading the KYC Docs in IPFS Network.

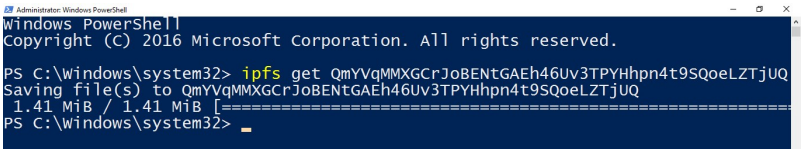

Fig. 3. Retrieving the KYC Docs from IPFS Network.

VI. Conclusion

The paper tried to implement a platform for easy KYC document verification through a document-sharing platform called IPFS. We used two different operating systems within two PCs to test our works. Both PCs were running on Windows 10 64-bit operating systems. We found that gpg4win with Kleopatra platform and IFPS for windows were installed in both PCs. The key generation and encryption processes were very smooth. We easily uploaded the encrypted file using the desktop app of IPFS using the command line interface of Windows Power Shell, and successfully uploaded and retrieved at PC2. Our research focused on a real scenario of a customer going to work with two financial institutions. The paper also showed the way of sharing the KYC docs without many difficulties between financial organizations through the wish of the customer.

This work can be extending in future by analyzing the various testing performances such as latency test, load test, stress test, etc.

References

[1] F. Glaser, "Pervasive Decentralisation of Digital Infrastructures: A Framework for Blockchain Enabled System and Use Case Analysis," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3052165, Jan. 2017.

[2] A. Rahman, S. Roy, M. S. Kaiser, and M. S. Islam, "A lightweight multi-tier s-mqtt framework to secure communication between low-end iot nodes," in *2018 5th International Conference on Networking, Systems and Security (NSysS)*, 2018, pp. 1–6.

[3] M. Raikwar, S. Mazumdar, S. Ruj, S. Sengupta, A. Chattopadhyay, and K.-Y. Lam, "A Blockchain Framework for Insurance Processes," *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, 2018.

[4] D. Puthal, N. Malik, S. Mohanty, E. Kougianos, and C. Yang, "The Blockchain as a Decentralized Security Framework [Future Directions]," *IEEE Consumer Electronics Magazine*, vol. 7, pp. 18–21, 2018.

[5] M. Mahmud, M. S. Kaiser, M. M. Rahman, M. A. Rahman, A. Shabut, S. Al-Mamun, and A. Hussain, "A Brain-Inspired Trust Management Model to Assure Security in a Cloud Based IoT Framework for Neuroscience Applications," *Cognitive Computation*, vol. 10, no. 5, pp. 864–873, Oct. 2018. [Online]. Available: https://doi.org/10.1007/s12559-018-9543-3

[6] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2010, library Catalog: bitcoin.org. [Online]. Available: https://bitcoin.org/en/bitcoin-paper

[7] D. Tonin, "Money Button CEO: How to upload large files to Bitcoin SV blockchain," Mar. 2019, library Catalog: coingeek.com Section: Tech. [Online]. Available: https://coingeek.com/money-button-ceo-how-to-upload-large-files-to-bitcoin-sv-blockchain/

[8] J. Benet, "Ipfs - content addressed, versioned, p2p file system," *arXiv:CSNI*, vol. 1407.3561, 2014.

[9] Q. Zheng, Y. Li, P. Chen, and X. Dong, "An innovative ipfs-based storage model for blockchain," in *2018 IEEE/WIC/ACM ICWI*, 2018, pp. 704–708.

[10] M. Irvine and D. King, "The Money Laundering Control Act of 1986: Tainted Money and the Criminal Defense Lawyer," *McGeorge Law Review*, vol. 19, no. 1, pp. 171–192, Jan. 1987. [Online]. Available: https://scholarlycommons.pacific.edu/mlr/vol19/iss1/9

[11] J. Parra Moyano and O. Ross, "KYC Optimization Using Distributed Ledger Technology," *Business & Information Systems Engineering*, vol. 59, no. 6, pp. 411–423, Dec. 2017. [Online]. Available: https://doi.org/10.1007/s12599-017-0504-2

[12] J. Parra-Moyano and O. Ross, "KYC Optimization Using Distributed Ledger Technology," *Business & Information Systems Engineering*, vol. 59, 2017.

[13] A. Soni and R. Duggal, "Reducing Risk in KYC (Know Your Customer) for large Indian banks using Big Data Analytics," *International Journal of Computer Applications*, vol. 97, pp. 49–53, Jul. 2014. [Online]. Available: http://adsabs.harvard.edu/abs/2014IJCA...97i..49S

[14] "Initio — Blockchain as the Newest Regtech Application— the Opportunity to Reduce the Burden of KYC for Financial Institutions," library Catalog: www.initio.eu. [Online]. Available: https://bit.ly/2YhYhMr

[15] M. Mainelli and M. Smith, "Sharing Ledgers for Sharing Economies: An Exploration of Mutual Distributed Ledgers (Aka Blockchain Technology)," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3083963, Nov. 2015. [Online]. Available: https://papers.ssrn.com/abstract=3083963

[16] W. M. Shbair, M. Steichen, J. François, and R. State, "Blockchain orchestration and experimentation framework: A case study of kyc," in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, 2018, pp. 1–6.

[17] N. Szabo, "Formalizing and Securing Relationships on Public Networks," *First Monday*, vol. 2, no. 9, Sep. 1997. [Online]. Available: https://firstmonday.org/ojs/index.php/fm/article/view/548

[18] "Blockchain & Distributed Ledger Technology (DLT)," library Catalog: www.worldbank.org. [Online]. Available: https://www.worldbank.org/en/topic/financialsector/brief/blockchain-dlt

[19] J. Howard, M. Kazar, S. Menees, D. Nichols, M. Satyanarayanan, R. N. Sidebotham, and M. West, "Scale and performance in a distributed file system," *SIGOPS Oper. Syst. Rev.*, vol. 21, no. 5, p. 1–2, Nov. 1987. [Online]. Available: https://doi.org/10.1145/37499.37500

[20] "What is IPFS? Interplanetary File System: Complete Beginner's Guide," last Accessed 31 Jul 2019. [Online]. Available: https://blockonomi.com/interplanetary-file-system/