Utilisation of Blockchain Technology for KYC process for banks in India using Aadhar Number

MSc Research Project
FinTech

Roshan Ramchandran
Student ID: X18120245

School of Computing
National College of Ireland

Supervisor: Noel Cosgrave
Student Name: Roshan Ramchandran

Student ID: X18120245

Programme: MSc in FinTech
Year: 2019

Module: Research Project

Supervisor: Noel Cosgrave

Submission Due Date: 16/09/2019

Project Title: Utilisation of Blockchain Technology for KYC process for banks in India using Aadhar Number

Word Count: 5104
Page Count: 15

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author’s written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ..........................................................
Date: ..................................................................

 PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST 

Attach a completed copy of this sheet to each project (including multiple copies) □

Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies). □

You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. □

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only
Signature: ..........................................................
Date: ..................................................................
Penalty Applied (if applicable): ..........................................................
Utilisation of Blockchain Technology for KYC process for banks in India using Aadhar Number

Roshan Ramchandran
X18120245

Abstract
As the economy of India is growing day-by-day, more people have started using its financial services i.e. it is becoming more financially inclusive. As a result, the Know Your Customer (KYC) process has to be both time and cost efficient. It is a mandatory process that has to be followed by every bank for every customer that wants to avail their services. However, the customer has to follow the same process again if they intend to work with another bank. The main objective of this research is to propose a smart contract based blockchain model which will reduce the time and cost involved in the KYC process. The proof of identity that has been considered for this research is a 12-digit unique Aadhar number which contains the customer’s personal information including their biometrics. The smart contract would allow the KYC details of the customer to be shared with other banks so that they do not have to do the KYC again. The whole system would be under the supervision of the national regulator of India i.e. Reserve Bank of India.

1 Introduction
As there is a huge increase in the rate of criminal activities related to the finance industry such as money laundering, providing funds for terrorist activities and other financial scams and frauds, KYC (Know Your Customer) process has become very crucial for banks and other financial institutions. The personal information of a customer such as their identity, home address, phone number, email address, etc. is collected as a part of KYC verification process (Soni and Duggal, 2014). According to an act passed in 2002 in India, known as the Prevention of Money-Laundering Act (PMLA), KYC is a mandatory process that every bank and financial institution has to follow to prevent any potential financial fraud or funding for criminal activities.¹

One of the biggest challenges faced by the banking industry is incurring of the ever-increasing cost of the KYC verification process. According to a survey conducted by Thomson Reuters in 2016, the cost spent by the banks and other financial institutions to meet the specific requirements of KYC are estimated to be around USD 60 million per year.² Furthermore, if these financial institutions fail to keep up with the KYC and anti-money

laundering (AML) rules and regulations, they would be subject to more charges or fines by the regulator.

For any customer to open an account with a bank, it is necessary to provide a document for the proof of identity like a passport or a unique 12-digit Aadhar number in India so that the bank can do the KYC in a legitimate way. Along with the proof of identity, the customer has to submit the data related to their finances such as financial statements, investments, and other information related to their financial data such as loans or any pending bills. Hence, as the name suggests “Know Your Customer” is the process where banks get to know their customer in a better way. There are two purposes of collecting these documents from the customer: 1) Collecting the proof of identity to identify that the customer is genuine and not using someone else’s identity and 2) the financial data that will help the banks in knowing about the creditworthiness of a customer and if they are eligible to do business with the bank (Arner et al., 2019).

The main purpose of this research is to solve the issues related to KYC by designing a decentralised model to perform KYC for financial institutions in India. A smart contract based blockchain technology has been designed to improve the process of KYC verification with banks which will save them a lot of money and give customers a better experience with the banks. The decentralised nature of the blockchain technology enables other banks to collect the customer’s data from the blockchain network and process their application within no time. Also, there will be no duplicate records on the network and it is completely decentralised. The duplication can be further reduced by only considering a unique 12-digit Aadhar Number allocated to all the citizens of India which can be used as a proof of identity.

The different sections of this paper are organised as follows: Section 2 critically analyses the related work in the domain of blockchain mostly in a financial sector. Section 3 gives an idea about the Research Methodology adopted for this research which explains the existing KYC process, a brief explanation about blockchain and its concepts. Section 4 is the Design Specification and Implementation which gives an idea about the tools used the research and how it was implemented. Section 5 discusses the evaluation of results and shows the output. Finally, Section 6 gives the conclusion and future work for the research.

2 Related Work

Since the introduction of the blockchain technology by Satoshi Nakamoto in 2008, when it was used as a permission-less public network for Bitcoin, there has been a tremendous advancement in the technology which were specifically designed to meet the ever-increasing requirements in different areas. Pop et al. (2018) analysed the Bucharest Stock Exchange (BVB) and concluded that it had some drawbacks. The most important drawback is that the stock exchange is managed by a central authority which is not transparent, can be vulnerable to cyber-attacks and it manages the transaction fees which is usually high. As a result of this, they proposed a decentralised system on the blockchain system for the entire stock exchange that will be transparent with no central authority and highly secure. The proposed system uses smart contracts to ensure the proper execution and settlement of all the orders. Dowling et al. (2018) have filed a patent for their work on blockchain based letter of credit (BLC). In the
supply chain industry, a letter of credit (LC) is a document issued by a bank used in international trade transactions that allows the seller to get the payment from the buyer as long as specific conditions are fulfilled. However, as the LC is paper-based, it is less secure, less transparent and cannot be tracked in real-time. The method proposed by Dowling et al. (2018) to solve this issue is blockchain based letter of credit (BLC). BLC stores the list of supply chain flow events and the documentary of the trade transaction, which when followed correctly and in a timely manner will trigger the payment for the contract and it will be transferred to the seller.

Mengelkamp et al. (2018) focused on their research on the case study of The Brooklyn Microgrid and proposed that a blockchain network can be used to trade energy between the residents of the neighbourhood. The need for renewable energy is increasing day-by-day and its demand is more than ever as more and more people are going for the renewable energy sources to protect the environment. Due to increasing demand on renewable energy, the houses with surplus energy can transfer the energy to the houses that need them and in return get paid for the energy that has been transferred. A smart contract can be implemented and can trigger payment whenever the energy is transferred. The research conducted by Pachaiyappan and Kasturi (2018) suggests that KYC can be conducted on blockchain using Smart Contracts. This will reduce the cost involved in the overall process and along with reducing the time taken. At the Bangalore Tech Summit in November 2017, State Bank of India, a bank from the public sector in India, announced that they are planning to implement the blockchain enable KYC process.3 Smart Contracts will be deployed to increase the efficiency, transparency and security during the KYC process in the banking system. As the blockchain is immutable, the data would be protected and tamper proof. It was on a beta test mode and there has been no update since then.

Brown et al. (2016) developed a platform – Corda in collaboration with a blockchain enterprise company - R3. It is a platform designed on the distributed ledger technology to record and process the financial agreements between the participating firms by automating the entire process. The smart contracts will manage the financial agreements between two or more participating companies in such that it follows the current rules and regulations and will be compatible with any changes in future regulations without the need of a central controller. Wang, Guo and Cheng (2019) proposed a new financial loan management system in China known as loan on blockchain (LoC). This system is mainly designed for the poverty alleviation loan, a special loan product of Fujian Rural Credit Union (FRCU) designed for people a certain poverty line in rural areas in China. The proposed loan management system is based on smart contracts over permissioned blockchain – Hyperledger Fabric. The main issues addressed by this system include decentralisation of the management system, creation of transparent system where the data updating process can be traced and will be tamper proof and increasing the security of the system.

In their paper, Hyvarinen, Risius and Friis (2017) mentioned the loopholes in the tax refund system whenever an individual or a company invests in foreign companies. The participating companies have reached an agreement the person or the company making an

3 https://inc42.com/buzz/bankchain-sbi-blockchain-kyc/
investment can claim a refund on the dividend payment so that they do not have to pay the tax twice. However, there have been cases where documents were forged to claim the tax refund causing the damage which was estimated to be around USD 1.8 billion in Denmark itself. They proposed a blockchain based system to prevent such tax frauds and increasing transparency with respect to the flow of dividends. In their paper, Zhong et al. (2018) have addressed the problems related to e-learning systems such as lack of user interaction for asking and answering questions and the interoperability in the e-learning systems. The proposed model uses blockchain technology and smart contracts to address these issues. The research conducted by Ahmad et al. (2018) concluded that the audit logs that store crucial data and can be used to store, audit and track any changes in the data are very susceptible to the attacks by the hackers or cyber criminals and it can be tampered very easily. To overcome this issue, they design a BlockAudit system which will store audit logs over a permissioned blockchain Hyperledger system.

Miraz and Donald (2018) are of opinion that the digital information in the Securities Exchanges are vulnerable to cyber-attacks. As a result, they have analysed the use of Blockchain in ensuring the security of stock exchange transactions along with a focus on its technological and legal aspects. Their research proposes a hybrid blockchain model which can be designed specifically for different stock exchanges. Also, during the design process of such a model, it should follow the regulations of the respective country. Srivastava et al. (2018) proposed a blockchain based system where the credits of a student are stored on the blockchain network instead of a college repository. This will address the issues related to centralised storage such as the security issues, higher access times, credit transfer between two universities or between the university and the ministry of education for ranking purposes. Their solution will also mean that all the data will be tamper proof. In their paper, Yu et al. (2019) analyse and summarise the traditional audit process and have addressed the issues related to it including security issue, privacy issue, data tampering. To overcome these issues, they have designed a blockchain technology that uses smart contracts to solve these problems.

3 Research Methodology

3.1 The Existing KYC Process

The main reasons due to which the Prevention of Money-Laundering Act of 2002 was introduced in India are the previous recorded events of financial frauds and scams, terrorist funding and criminal activities and money-laundering. So, the banks and financial institutions are obliged to follow a proper KYC onboarding process as ordered by the regulating authority to avoid doing business with the customers involved in the above-mentioned activities. “The KYC process consists of an exchange of documents between the customer and the financial institution that intend to work together” (Moyano and Ross, 2017). In this process, the customer has to submit an identification proof such as an Aadhar Number or a Passport which has information such as the name, date of birth, address which is very crucial
in finance industry. This also means that the banks have to monitor the financial data of the customer such as their bank statement, their source of income, investments and their patterns of transactions and the risk involved while onboarding the customer about their creditworthiness. Due to this, the banks have to pay huge costs to complete the process and if this process is not followed according to the regulations, they would have to pay hefty fines to the regulatory authority (Moyano and Ross, 2017).

Whenever a customer wishes to work with the bank for opening an account or taking a loan from them, a KYC process has to be initiated. Firstly, the customer agrees with the terms and conditions of its relationship with the bank which depends on the service. The next step is for the customer to give all the documents required by the bank to perform the KYC process including the identity of proof and the financial data. Lastly, the bank performs deep analysis of the documents submitted by customer to check their creditworthiness and create an internal document that would certify that the customer has been accepted or rejected and that the KYC verification was done according the regulations. The whole process has to be repeated whenever the customer wanted to open an account with a new bank. The current regulation says that when the customer intends to work with another bank, the cost of the new KYC process has to be borne by the new bank itself. An illustrative example of the current KYC process in Figure 1 where the customer works with three different banks. In this example, it is clearly shown that the customer has to submit the documents and perform the KYC process thrice with three different banks and even the cost involved is three times of what was required for a single KYC process. However, it is crucial to know the difference between the “core KYC verification” process and the processes that are set by a particular bank. The process which has to satisfy the minimum requirements as specified by the regulating authority to perform the KYC process is the Core KYC verification process whereas the one which requires the submission of some additional documents depending on the banks is a bank-specific KYC process. The solution proposed in this research is purely based on the core KYC process and it will be shared by the banks within the same jurisdiction (Moyano and Ross, 2017).

![Fig. 1. The KYC process followed currently](image_url)
3.2 Aadhar Number

The Indian Government created the Unique Identification Authority of India (UIDAI) in January 2009 which was given the task of assigning unique identification numbers to all the citizens of India. This 12-digit unique identification number is known as Aadhar Number which stores the personal details of a person including name, sex, home address, phone number, email address, bank account details, etc. (Mudliar, Parekh and Bhavathankar, 2018). It also stores the demographic details such as fingerprint and iris along with other details such as passport number, permanent account number (equivalent of PPS number in Ireland). All these details are stored in a smart card in the form of a QR code which the citizens can use to access different government services (Prakash, Muniyal and Acharya, 2019).

3.3 Blockchain Network

Blockchain was first introduced to the world when it was launched in 2008 as the base technology for Bitcoin. All the transactions which is transparent, removes corruption and does not require any central management or authority. This creates an immutable record that can be altered as there is a time-stamp on each entry with a unique identification code which is also linked to the last entry (Nakamoto, 2008). All the transactions are stored in the form of a digital register which is called a ledger and it is not stored by any single person or an organisation but on a decentralised network. This ledger is accessible to everyone on the network and once a transaction has been recorded on the ledger, no user on the network could be able to make changes on it. As the data is stored on all the nodes on the network and not stored centrally at a single location or a computer, there is great amount of transparency and it removes the need for a single entity to manage the entire network. The IT experts agree that the blockchain is the new trending technology that will revolutionise every aspect of global business (Duy et al., 2018).

Basically, blockchain technology uses peer-to-peer (P2P) networks without any need of a central server. The user on the network who wants to start to transact on the network would have to broadcast their information on the entire P2P network. Once the information is broadcasted on the network, other nodes can view the transaction and approve its authenticity using different consensus mechanisms or algorithms such as Proof-of-Work, Proof-of-Concept, Practical Byzantine Fault Tolerance (PBFT), Proof-of-Stake, etc. The authenticated data is then stored in the new block which is ready to be added to the blockchain network. And finally, the consensus algorithm, with the permission and confirmation of all the nodes on the blockchain network, adds the new block to the existing blockchain network. Hence, if there is even a slight change in any of the information on the block, a new hash of the data will be created with the latest time stamp and thus eventually all the blocks in the blockchain will be affected as each block has the information from the previous block. This makes it practically impossible to hack the information on the blockchain network. Also, it removes the need of third-party vendors or intermediaries like banks and brokers and thus, saving a lot of money for the user (Duy et al., 2018).
From the security standpoint, the technologies such as Hashing, Cryptography, Digital Certificates and Signatures are used in the blockchain network to make sure that the transactions are valid, secure and their integrity is maintained without the need of any third-party security service provider. The structure of a blockchain has been illustrated in figure 2. As seen from the Figure 2, each block contains data of its own along with the time stamp and hash of the previous block. This entire information is stored in the form of a Hash which can be considered as the digital identity of the block (Duy, et al., 2018).

Fig. 2. A typical structure of a blockchain (Duy, et al., 2018)

3.4 Smart Contracts

The main aim of the blockchain technology was to mainly to create new currencies or rather cryptocurrencies such as Bitcoin, Ethereum, Litecoin and many more. However, it can also be used to manage the operation of decentralised systems using smart contracts. Smart Contract is a computer code which can be triggered automatically and implemented to take specific actions as defined by the user when certain conditions are fulfilled (Moyano and Ross, 2017). It is very much similar to a normal contract on paper which is agreed between two parties with specific clauses that can be triggered when the condition is fulfilled. Figure 3 and figure 4 illustrates the difference between a traditional contract and a smart contract. Szabo (1997) suggests that the smart contracts should be used and utilised in a way that it can be implemented in most of the aspects of businesses and can be controlled by the digital means. Since its introduction in Satoshi Nakamoto’s paper for Bitcoin (Nakamoto, 2008), blockchain technology has been used in many new innovative projects in different industries for a better service or product. One such platform known as “Ethereum” uses this technology to run applications on a decentralised network.

The utilisation of smart contracts for the distributed ledger technology was first suggested by Szabo (1994) which are generally called “blockchain contracts” or “self-executing contracts” or “digital contracts”. In this system, traditional contracts can be replaced by a series of programming codes and can be stored and retrieved on the system which is regulated by a series of computers on the blockchain network.
4 Design Specification and Implementation

The proposed solution for the KYC verification is a smart contract based blockchain network where the customer details will be stored once the KYC is performed by the first bank. The solution in this research was proposed on following three assumptions:

1. All the financial institutions in India are supposed to follow the same rules and regulations for the KYC process and would eventually have the same agreement with the customers before the KYC verification process.
2. The mean cost of perform the KYC would be mutually decided and agreed by the participating financial institutions in the entire blockchain network. The costs would however, vary according to the service opted for by the customer such as purpose of work, business of the client, the documents submitted, etc.
3. The regulating authority of India i.e. Reserve Bank of India (RBI) would have be central managing authority on the network and will make sure that all the financial institutions are following the procedures for a smooth administration.

The proposed solution uses Ethereum smart contracts to store information on the blockchain network. Table 1 shows the software requirements for the implementation of the proposed solution. The smart contracts were written on Solidity language on Remix IDE which can be run any web browser like Google Chrome, Mozilla Firefox, IE Edge. Ganache hosted the blockchain network on the local server which had 10 Ethereum accounts for free that can be used for the purpose of development. Metamask is Google Chrome Extension that has to installed while implementing the smart contract. It is a wallet that connects your smart contract to different blockchain networks. For client-side implementation, the web3 package was used that would interact with the Ethereum smart contract with the help of JavaScript. In Solidity, the input was taken using strings and was to be stored in an array. After some trials, it was found that the bytes input was more suitable for the smart contracts. However, the user would not understand the bytes array input; so, a function was added in the client-side html file to convert the string input to bytes input for smart contract.

The smart contract was deployed on the blockchain test network and not on the main network as it would require real ethers for every single transaction. The test networks have test accounts and fake ethers in the account so that it can be used for checking the functionality of the smart contract before its actual implementation on the main Ethereum network. As Ganache was used in this research, an RPC server was used with the port number – 7545 as a testing network. Instead of Proof of Work, the test network uses Proof of Authority where the predefined nodes verify the transactions. Once it has been approved, it is then added to the blockchain network.

The actual implementation of this system can be explained with the help of following steps: 1) Customer will open the portal, 2) Go to Customer Login => New Customer, 3) Fill in the required fields and submit => The customer will receive a unique index key, 4) The data will be collected by the bank, 5) Bank will do KYC of the customer, 6) Once verified, the bank will update the portal that the KYC of the customer has been successfully completed, 7) When the customer wishes to work with another bank, they can give them their index key to the bank and the bank will look for the customer’s data on the blockchain network and verify the information online.

<table>
<thead>
<tr>
<th>Software</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solidity</td>
<td>v0.4.20</td>
</tr>
<tr>
<td>Metamask</td>
<td>Google Chrome Extension</td>
</tr>
<tr>
<td>Ganache</td>
<td>v2.1.0</td>
</tr>
<tr>
<td>web3</td>
<td>v0.20.7</td>
</tr>
<tr>
<td>Notepad++</td>
<td>v7.5.8</td>
</tr>
</tbody>
</table>

Table 1

5 Evaluation

The Ethereum smart contract was successfully compiled and deployed using Remix IDE on the RPC port number 7545 which acts as a test network for blockchain. It can be observed that ethers were deducted from the fake account on test network to carry out the transaction successfully. The following code was written on the smart contract that was deployed on the blockchain network.

```
pragma solidity ^0.4.24;

contract KYC{

    struct Customer {
        bytes name;
        bytes dob;
        bytes aadhar;
    }
```
bytes phone;
bytes email;
bytes bank_name;
bytes customerOk;
}
constructor() public{
}

Customer[] allCustomers;

function getDataSize() public view returns(uint) {
    uint total = 0;
    for(uint i=0;i<allCustomers.length;i++) {
        total = total + allCustomers[i].name.length + allCustomers[i].dob.length +
                          allCustomers[i].aadhar.length + allCustomers[i].phone.length + allCustomers[i].email.length +
                          allCustomers[i].bank_name.length + allCustomers[i].customerOk.length;
    }
    return total;
}

function getDataHash(uint index) public view returns(bytes32) {
    return blockhash(index);
}

function addCustomer(bytes name, bytes dob, bytes aadhar, bytes phone, bytes email,
bytes bank_name, bytes customerOk) public {
    allCustomers.length ++;
    allCustomers[allCustomers.length-1] =
Customer(name,dob,aadhar,phone,email,bank_name,customerOk);
}

function getCustomersCount() public view returns (uint){
    return allCustomers.length;
}

function getCustomerData(uint index) public view returns (bytes name, bytes dob, bytes
aadhar, bytes phone, bytes email, bytes bank_name, bytes customerOk) {
    Customer storage c = allCustomers[index];
    return (c.name, c.dob, c.aadhar, c.phone, c.email, c.bank_name, c.customerOk);
}

function stringToBytes32(string memory source)public view returns (bytes32 result) {
    bytes memory tempEmptyStringTest = bytes(source);
if (tempEmptyStringTest.length == 0) {
    return 0x0;
}

assembly {
    result := mload(add(source, 32))
}

The details of the log file for which contains the information of each and every transaction can be shown in Figure 5. It contains the contract address, the hash code of the transaction, block number, etc.

![Fig. 5](image)

The inputs and outputs of the deployed smart contract as viewed on Remix IDE are shown below in Figure 6 and Figure 7 respectively:

![Fig. 6](image) – Input that can be given to the smart contract
The smart contract that was created for this research fetches the input from the user, stores it on the blockchain and can be retrieved whenever required. This framework assures the core KYC verification process has to be done just once by the first bank and other banks can utilise the information from the blockchain itself. After the implementation of this model, Figure 9 will illustrate the improved version of the KYC process that was shown in Figure 1. Figure 8 shows the front page of the client-side application.
6 Conclusion and Future Work

The main objective of this research was to design an improvised KYC system for India that can reduce the cost and time involved in the process. Ethereum Smart Contracts based Blockchain model was used to design this system. Aadhar Number is key input for the successful completion of the KYC verification, as it is the most important document in India. The customers would be able to work with any bank of their choice that are on the blockchain network without going through the KYC process again. Smart contracts deployed on the blockchain network ensured that the customer is KYC compliant and is ready to work with the bank. As the blockchain network was hosted on a test network, the Proof of Authority was done for the designed blockchain system instead of Proof of Work. However, the client-side application on web3 still needs to be enhanced as it does not support some of the functionality of Ethereum.

The future work for this research would be privatise the whole blockchain system with the help of Hyperledger Fabric and create a database to store the customer’s details such as login ID, password, their documents to further speed up the process.

References


Mudliar, K., Parekh, H. and Bhavathankar, P., 2018, February. A comprehensive integration of national identity with blockchain technology. In 2018 International Conference on Communication information and Computing Technology (ICCICT) (pp. 1-6). IEEE.


