FRAMEWORK FOR DATA MANAGEMENT IN PUBLIC SERVICE DELIVERY APPLICATIONS IN SRI LANKA USING BLOCKCHAIN TECHNOLOGY

Jeewani Goonathilaake, Nuwanthika Deshapriya, Ruvini Jayakody and Methviru Dharanidu

School of Computing, University of Colombo, 35, Reid Avenue, Colombo 07, Sri Lanka

ABSTRACT: The research paper discusses about how data management issues found in the current Sri Lankan public service delivery systems namely, difficulties in data accessibility, data manipulation possibility, data loss and privacy preservation can be mitigated through the application of Blockchain technology. Three distinct public service delivery systems were analyzed in order to derive a common framework which can be used to design a Blockchain based solution for the public sector processes. Systems were redesigned using Blockchain technology and two prototypes were developed using two different existing platforms and qualitatively evaluated for the four criteria focused by the research. Based on the characteristics extracted from the selected systems, a Generic Guideline was designed comprising the six areas that need to be considered when selecting a Blockchain for Data Management in a Public Service Delivery application and five steps that should be followed in adopting a Blockchain for the specific needs of the government institution.

KEYWORDS: Blockchain, Public Service Delivery Systems

INTRODUCTION

In Sri Lankan government institutions, public data is stored and managed on behalf of the public. Certificate issuing agencies such as Land registry department, District Secretariat, Department of registration of persons store critical public data which is used in verification purposes. So the integrity of the data in such certificates is crucial. Furthermore there are public data which needs specific privacy measures regardless of being open to view. For example medical records of a patient will need more privacy than a verification document data like a death certificate. For the data management in public entities, computerization of records is initialized under egovernment transformation in Sri Lanka and under that computerization of Birth, Marriage and Death certificates (BMD project), land title information (eLand registry project) and computerization of health data of government hospitals (eHR project) [1] were some of the most successful initiatives. In those systems, storage and data management were done using database options such as Centralized databases and Distributed databases.

However these public data management processes still reflect prominent problems such as accessibility difficulties when receiving a service, possibility for fraud and error, privacy preservation and data loss possibility. Nevertheless the database solutions incorporated in the current system are unable to provide a single option to address the above mentioned problems collectively. A centralized database may accomplish the challenge of accessibility, but there is a data loss possibility unless a backup is kept. Since a central authority is there, a malicious database admin could alter data in the source. A decentralized database may promise the geographical independence, however if a node is down with failure, data in that node will be unavailable unless replication is done. If full replication is done, data availability can be promised, but maintenance is costly. Privacy of records is beyond the authorship of the owner of data in all database solutions. All of the above mentioned database solutions need a security layer implemented separately as it is not available as an inbuilt feature.
For that an expert’s ideas and effort is needed which is both costly and time consuming. When designing a DBMS, threat is not modeled and audit trails are needed to monitor user actions. However, audit trails are not inbuilt, which also require an extra effort for implement and audit trail is not a preventive mechanism [2] mechanism for fraudulent activities. Among potential solutions, Blockchain [3] is a trending concept that shows its capability to cater for a better storage medium, a distributed ledger. Therefore the applicability of Blockchain technology to serve public data management in Sri Lanka is worth to be explored.

Research Problem

The main research problem is —Applicability of Blockchain Technology to mitigate data management issues in public data management in Sri Lanka. The research problem is further segregated into 4 research questions as,

Q1: Reducing the difficulties of data accessibility in public data management systems using Blockchain technology
Q2: Reducing the possibilities of data management frauds and errors using Blockchain technology
Q3: Reducing the data loss possibilities using Blockchain technology
Q4: Privacy Preservation using Blockchain technology

LITERATURE SURVEY

In Sri Lanka like in other countries, Public Data Management systems are computerizing under egovernment initiatives [1]. Civil registration systems, Land title management systems and Health records systems are computerized with usage of centralized, decentralized and cloud database solutions in many countries [4]. Though computerization of records is done under egovernment reforms, all database solutions need to implement security as a separate layer. According to recommendations —Report of Committee on Computerization of Land Records [5] of Department of India, the project needs to implement security measures to be reliable. Provisioning backup for SQL database, backup storage managing under an official, implementing both physical and cyber security, implementing policy for confidentiality, accountability, access control and data access over network, and scheduling security audits should be carried out. Database solutions also subject to data loss vulnerability and data manipulation if not well maintained. In 2016 there were 269 data breaches occurred in government institutions and from all data breaches occurred from 2013, 11.46% were directed to government institutions [6]. If a centralized database is used, there is a centralized authority in charge of the database that possesses power to manipulate the database and forge records or delete records. In a government organization people with power could misuse this property which makes the solution non reliable. If a distributed database is used, when one node is down, data stored there would be in lost unless a replication exists. If a cloud database is chosen, there is also a third party to be trusted for the integrity of data and unauthorized data modification is possible via a cyber-attack. Therefore, every database solution possesses a significant vulnerability in the management of public data. No single solution will possess power to provide distributed access, data availability and data integrity.

Apart from data security, in the areas such as public health record management preserving privacy has become the main issue. In the international context several researches have been done in order to address the privacy concerns in large scale health related data management systems. In the research conducted by B. Sangeetha et al. [7], they have explored the capabilities of protecting privacy in a
cloud-based environment. In the process they have developed a novel framework where the patients can encrypt their medical data when uploading to the cloud. Ability of using multi-key approach in the cloud-based environment is also examined by Poonam Patel, Amar Buchade [8]. This research also highlights that existing techniques are not capable of providing complete privacy.

**Blockchain Technology**

The Blockchain concept established with the arrival of crypto currency, Bitcoin [9]. A Blockchain can be simply defined as a distributed ledger, where all the transactions or digital events that have been executed are shared among participants through a peer-to-peer network.

Block is the basic organizational unit of data in the Blockchain. The very first block in the Blockchain is called the Genesis Block [10] and it is always hard-coded to establish the chain. A block consists of two parts namely, block header and body. Block header contains the hash value of the previous block, block size, a random Nonce value, Block height; no of precedent blocks, Merkle root hashes, timestamp of creation time of block and the number of transactions included in the block. The body contains non-empty transactions or records. Blocks are linked by the hash values of previous block. If the hashes are not matching, the Blockchain breaks from that point and it is used to identify non-genuine transactions [11]. Merkle root is the component that vouches for the integrity of individual records or transactions included in the block. The transactions are first individually hashed. Then they are combined and hashed in a tree-like order. The top level is named as the Merkle root which is created using the systematic hashing combination of all transactions [12].

The process of adding a new block to the existing Blockchain is called —mining. Miners engage in a process of solving a complex mathematical problem and the first miner who is capable of identifying the solution for the problem will be the one who add the block to the existing Blockchain. Blockchain constantly grows as miners add new blocks to the chain to record the most recent transactions. The blocks are added to the Blockchain in a linear, chronological order and transmitted all over the nodes of the network. Once the data is entered into the Blockchain, those data cannot be altered or deleted in the original place. That is because of the —append only feature in Blockchain. Modified data will be added to a separate block in the chain with the timestamp. So the auditability of the valid transactions is ensured automatically.

According to the conducted researches, Blockchain has several areas of application such as crypto currencies in financial sector [13][14], smart contracts enabling [15], e-voting application building [15], managing digital rights and intellectual property rights [13][16], IoT applications [17][18] and enabling Public Services [19].

**Blockchain for Public Data Management**

Several researches have made attempts to justify the usage of Blockchains in the public service delivery applications. Some researchers have been focused on introducing theoretical frameworks [19], [20] for adopting Blockchains in public sector services while some focusing on introducing implementable Blockchain solutions in different public services like voting [13], banking, public funds administration [19], land registry systems [15]. Even Though there is a wide range of public services, most of those researches were focused on addressing some common concerns existing within public service delivery applications. Those common concerns may include improving the data management efficiency, authentication of public data records and preserving the privacy of data owners. In the attempts of improving the data management efficiency some researchers have directly adopted Blockchains as a decentralized database [28]. However due to the scalability issues; some researches suggesting to use
Off-Blockchain data repositories to store data; while Blockchains maintaining only the pointers to that data [21],[22] specially when handling large scale data repositories. In the area of public data authentication researches were mainly focused academic certificate authentication [23] and ownership verification of digital assets [24], [25] or non-digital assets [26]. However in the privacy preserving scenarios, Blockchains were mainly used to either protect the identities of system users [27] or to provide the control of data to its original owners [28]. Implementations have been done using Blockchain for data management in academic certificates and professional certificates as a storing and a verification repository [23][29][30], health data repository [21], Land Title Management like Bitland [26] which is a smart contract based land title tokenizing system conducted in Ghana and Kenya.

METHODOLOGY

The research process was structured as a six step process in order to ensure the completeness.

1. Studying current Public Data Management Systems and identifying the existing challenges in delivering value to general public. The identified main concerns in public service are namely, data accessibility, data loss, concerns related to frauds and errors, and privacy related issues.

2. Investigation of the features of Blockchain technology to address challenges in public data management systems, which cannot be addressed through current database solutions. A comparison between Blockchains and other different data management solutions such as centralized databases, decentralized databases and Virtual private databases were done based on four criteria. Through the comparison, it is justified that the characteristics of Blockchain technology possess the ability to achieve all four criteria.

3. Identifying three public data management systems which would utilize the properties of Blockchain technology to create value by resolving issues in their current systems. From the available Data Management Systems, Birth Marriage Death certificate issuing process [1], Land Title Management, and Health data records management systems were chosen to design systems which incorporate Blockchain technology. Reasons for the selection were based on the egovernment initiatives deployed worldwide. In most of the countries, the first preferred system for egovernment transformation is the Identity management and Birth Marriage Death certificate management, land registry management and health data management. Another main reason behind the selection is that all these processes are accessed by the majority of the Sri Lankans and these were some of the main processes which were focused by the egovernment initiatives launched by the Information and Communication Technology Agency Sri Lanka (ICTA) [1]. A major portion of the data relevant to these processes has been converted into digital form and the rest is converting in ongoing projects.

Therefore, with consideration to the feasibility of the proposed solution we decided on the above mentioned three processes.

4. Designing of three selected data management systems architecture using Blockchain technology. The objective of the designing was to use Blockchain characteristics can be used to address the existing concerns of those particular systems.

5. Designing an abstract guideline/ framework by selecting properties from the selected data management systems. The selected systems are redesign with Blockchain to address the concerns which are specific to the particular system. Since the data management processes of different systems are not similar, there are some significant differences between those three designs. At this
phase our attempt focused on generalizing designs of the selected systems to come up with a common data management framework that could be used in government data management systems.

6. Evaluating Abstract Prototypes. The extracted designs evaluated qualitatively based on four main criteria,

- Possibility of deleting or modifying a record in the Blockchain
- Tracking history of records
- Accessibility providing ability
- Preserving privacy of records

DESIGNS

The main focus of the research is to identify the suitability and applicability of Blockchain technology for the Sri Lankan public data management systems. As for the design phase, an analysis was done regarding the three public service delivery systems that were chosen, to identify the existing issues and to design a suitable approach incorporating Blockchain to overcome the identified drawbacks with respect to different centralized or decentralized data management solutions which have been incorporated into the public sector organizations.

Birth Marriage Death System Design

The Birth Marriage Death certificates management process (BMD project) [1] is one of the Sri Lankan e-government initiatives which replaced the manual process of re-issuing the copies of Birth, Marriage and Death certificates by a single computerized system. Physical certificates that are manually issued at the point of initiation will be collected by the relevant district secretariat office and convert them to digital form by scanning the documents and store in a database with a set of user critical data like the document type, district secretariat, division, name, date field mentioned in the certificate and serial number. The certificates belongs to particular district secretariat office will be reside in their own database situated at their premise and will be available for retrieval by the citizens who are in need of obtaining another copy of their certificate or in case where the certificate they own is damaged or lost and whose certificates.

However isolated databases at each divisional secretariat office restrict the access of information in the current system due to the fact that individuals must visit the divisional secretariat which owns that certificate according to the location of physical initiation of that certificate in the first place to obtain a copy, irrespective of their present residence. This also adds a risk of data loss if a corruption in the database occurs or in case of a natural disaster, since the certificate information belongs to a particular divisional secretariat is only stored in the local premise. Recovery in case of a failure will also be based on a backup that resides at the local premise which might also subject to the same failure or a natural disaster if any case. Then the recovery might be based on the physical certificates maintained in the premises which are also facing the possibility of getting easily damaged.

Also, there is a difficulty to identify any change for the documents in the database due to the possibility that a person at the local government institute with access is able to replace any image with a fraudulent document of a modified certificate and use that as the original certificate without any alarm.
Furthermore there is a concern with the mobility of the document which expresses the necessity of carrying the physical certificate for the purpose of verifying details about a particular individual. These certificates might get misplaced, damaged or most importantly this will open up the risk of being able to use a fake certificate as the original document. At present third party institutes which need to verify certificates for different purposes like Banking, Immigration, and Insurance rely on the physical documents which are being given to them by a certain individual. Therefore it’s important that these documents are in a form which is accessible and convenient for the public to carry around and to use in any case of verification of details, and a mechanism to check against the system whether the document is original.

**Proposed Solution**

In the proposed solution we thought of suggesting a mechanism which will use Blockchain as an enhanced mechanism to the existing system which will address and mitigate the concerns arising through the issues. The proposed design incorporates the existing system and its structures as well, along with a Blockchain and a wallet based approach to individual citizens. A mobile wallet in the sense is actually a mobile application which includes the scanned images of Birth, Marriage and Death certificates as requested by an individual at the point of registering the mobile wallet as proposed in the design. Like in the current system, the solution is common to all the three different types of certificates that are being managed under the BMD project, since we have used the existing system as the foundation of the proposed design. In the context of value added service, a Blockchain will be implemented to store the hash values of the digital documents and act as a mechanism which allows the third parties to verify the certificates without needing physical documents as proof. This also has been used in the design to eliminate unauthorized modifications to the scanned images residing in the database and to identify when such instance occur. The Blockchain concept in the design supports the mobile wallet concept to create value addition to the users with relevant to the certificate obtaining procedure, by acting as a security enhancement which leads to reliability and the effectiveness of the wallet functionality and the purpose in the design. This will ensure the mobility of the documents as well as the identification of any changes to documents in the database without any authorization from the user, as well as acting like a backup for the personal digital documents in case if there is a data loss occurred in the local government body. The proposed design is further explained in the below sections with regards to the processes associated with the system.
1. **Image scanning process:** This phase of the process is already exists in the currently implemented BMD project, which is being carried out by the data operating officer at each divisional secretariat office. The main reason for not changing the existing process is due to the well-established nature of the current mechanism of converting the physical documents to the digital form well as the nature of the certificate, certificate creation and the time period when a specific individual is starting use these certificates for verification purposes. We did not considered to have a digital version of the certificate solely with the user because for validation purposes third party organizations are relying on a centralized authority, which is the government play the validating authority for legal purposes. Therefore instead, the proposed solution is designed in a way that is using these scanned images of the certificates in the database, for the user’s mobile devices as well which are being transferred to the mobile device at the point of registration of the mobile wallet where the users can decide whether they want to obtain the Blockchain based mobile wallet mechanism as a value addition for their interaction with these public service providing authorities.

2. **Downloading and setting up the mobile wallet:** Users will be able to download the mobile wallet/application from the app store, which is suggested to be developed for the BMD project according to our design. The user then has to setup the mobile application by providing a PIN number and several onetime answers for a specific set of security questions which are then used to create the wallet ID. Wallet ID will be the unique value or the number which is used to differentiate each mobile wallet owned by different users of the system. Here the wallet ID will be stored within the device. The reason for using a wallet ID instead of a random number is to make the recovery of the wallet easy in case of losing the mobile device containing the mobile wallet. Wallet ID will be able to retrieve from the value stored in the Blockchain by following the same procedure of registration of the mobile wallet an in case where the users forgot the answers for the security questions and the pin number the
existing record of hash values with relevant to the previous wallet ID will be flagged and a new ID can be obtained as per request.

3. Physical Authentication and the creation of Hash: As an additional step, physical authentication was introduced to the process due to the need of correctly verifying the owner of certificates which are about to get transferred to a mobile device of an individual. An operator at the divisional secretariat will compare the physical verification details like National identity cards, provided to him by the citizen and the search result obtained using the currently existing system by filtering out the critical data. It’s also important that the operator check whether the record is already being associated with a wallet or not before continuing to identify an attempt to fraudulently obtain someone else's information.

Setting up the content After the user get physically authenticated he/she can send the mobile wallet ID to the operators web service to carry out the transferring of the digital certificate details from the local government authorities database to the users mobile device. Hash 1 will be the hash value of the digital certificate document and the Hash 2 is the hash value generated using combining the Hash 1 value and the Wallet ID the user sent to the operator's web service. Here these hashes will generated for each certificate that the user is requesting and will be recorded in the Blockchain as a single transaction where Hash 2 value is the transaction ID and the value of Hash 1 is the transaction value. The process recording the Blockchain transaction involve a proof of work based mining protocol by a mining network consists of miners from each of the district secretariat office around Sri Lanka. The two hash values will be sent to the user's mobile device along with the digital certificate transferring. The Blockchain which is proposed to implement in the design is a private Blockchain but it is opened to access and view the transaction by authorized third party organizations to carry out the verification of digital scanned images if certificates presented to them by the citizens. Web portals can be created for institutes like banks or other government bodies to access the Blockchain, by searching through the transaction ID provided by the citizen through the wallet which is the Hash 2 value. Verification can be done by comparing the Hash 1 stored in the wallet and the Hash 1 value in the particular search result of Blockchain transaction.

In the 1st phase of designing Blockchain was used as the sole database for storing the digital version of the certificates rather than using a separate normal database. However we faced some limitations with the storage capacity of a particular block [11] which aroused concerns with mining effort, cost and time taken to complete the mining process of a particular block. In the next design phase we incorporated Blockchain to the current system as an enhancement that address the issues related to verifiability, security and the mobility of the digital Certificates stored in the currently deployed system by incorporating the mobile wallet concept as a supporting concept. Through the design the inconveniences occurring to the users because of the geographical barriers, carrying the physical documents as proof for verification purposes will be mitigated. The proposed design also contribute to the security aspect of the process by eliminating the possibilities of fraudulent modifications to the digital data stored in the databases which are either carried out with or without the acknowledgement of the owner of the digital certificates. The data residing in mobile devices can be used as a backup copy of data in case of a need. Since through the design we focused on providing a web portal third party organizations are also having the convenience of correctly verifying the individual that they are about to provide the service and eliminating the risk of involving or facing a fraudulent action and also by not involving in huge documentation procedure incurring a huge amount of time. One of the limitations of the current design is that since the mechanism is introduced on top of the existing system of the BMD project as an optional enhancement mechanism, whether the users will be engaged with the mechanism.
Land Title Management System Design

Land Registry Department provides a particular set of services as registering a particular land under the name of owner, once the deed is registered and issuing copies of stored certificates at request [31].

In Sri Lanka when the ownership of a land is transferred to a new user a new deed document is written in front of a notary in the presence of witnesses. When such registered deed is brought into a land registrar department office, the registration of that land for the new owner is done. Request for a registration is done using the deed and an application issued by the Land Registry Department.

Folio records are the documents that hold the ownership details for a particular land. Folio records are created with the details of a deed and some registration process related data. The process of writing a folio record is considered as the registration of a land title. Registration clerk is the person who does registration and there is a specific clerk for a particular court division. The unique identifier for a particular registration is the Daybook No. The Daybook No is generated in the order of registrations starting from the January 1st for a particular year. A folio record will contain details about a land including Division of the land location, Folio No(All registrations for one land block goes under one folio No), Name of Land, Plan No and Date, Name of Surveyor, Land Lot No (Specific number that identifies the land), DayBook No, Grantors, Grantees, District, Province, Boundaries of land, No and Date of Deed, Name of Notary and Signature of Registrar who authorized the registration.

There are different roles in the Land Registry such as Counter which enters the data in the application and deed to proceed with registration and forward documents to Assistant Registrar, Assistant Registrar who divide deed documents to the relevant Registration Clerk of the court division where the land located, Registration Clerk who writes a folio record to register a land under the owner’s name and Registrar who authorize the registration In the existing eLand Registry, there is no facility to view folio or deed documents in real time. To request copies, requester should visit to a branch. Folio records and deeds are stored in separate tables in the centralized database.

The eLand Registry is most affected by the data manipulation because if the data get modified and remains undetected, integrity of land titles is lost. Furthermore using a centralized database leads to data loss can if proper backups are not kept.

Proposed Solution

Using Blockchain properties the identified issues can be suppressed. When incorporating Blockchain in to the Land Title Management, only a segment of folio data is entered into Blockchain according to the proposed solution. The data that holds the ownership for a land block is entered into Blockchain while rest of data inserted into folio table in database. A temporary table is used to carry that data up to the registration and once that data is entered to the Blockchain; they get flushed from temporary table. Once a registration is complete, critical data that holds ownership details will be with Blockchain and rest of data will be in the database. The Blockchain transaction ID which generates when the data entered to Blockchain is used as the common key between the database table and Blockchain. Validation of data that is entered into the Blockchain is done in the registration process itself and the validation completes with the Registrar’s approval.

Private Permissive Blockchain is used because the Land Registrar should have the authority of registrations. Customized block structure containing ownership details for a land is used for the Blockchain and dedicated mining is done using a pool of miners representing the Land Registry offices.
The Registration Process When the Deed is handed over to the Counter along with the Application for Registration of Deed,

a. Data is entered to the local SQL Database
b. Issue a RRRN No(Day Book No) for the registration
c. Deeds to be registered are forwarded to Assistant Registrar
d. The Assistant Registrar divides the Deeds to the relevant Registration Clerk
e. The Registration Clerk checks folio records in the Blockchain for the history of Registration of Land Block
f. Registration clerk registers the Deed
g. Forward the Registered Deed for Registrar Approval
h. On receipt of approval, Registration Clerk enters critical Folio Data to the Blockchain

Blockchain token will comprise of, Day book No (RRN_No) – unique identifier, Folio, Deed No, Lot No, Grantors and Grantees, Hash value of Deed and Signature of Registration clerk. After adding a particular folio record to the Blockchain, a unique Blockchain transaction ID is generated. It is used as the link between database and Blockchain.

i. When the data entered into Blockchain, the pool of miners which represent the Department of Registrar General, mine and add blocks to the Blockchain

When the block is added, rest of folio data is submitted into database along with the scanned copy of deed and Blockchain transaction ID. Once the Registrar authorizes the submitted transaction, deed is returned to citizen along with the Blockchain transaction ID which can be used as a reference number to query.

Document Retrieval When a citizen request a copy of a Deed or Folio stored in Land Registry Office,

a. Citizen submits an application with required details to locate Deed or Folio.
b. Operator
c. Query the Database + Blockchain using provided details and provide a copy

Viewing Documents

a. For verification purposes, external organizations,
b. Get access through a portal and query using Blockchain transaction ID
c. View required document
When evaluating the design against the identified issues, the main concern of frauds can be addressed. Since the data that verify the ownership for a particular land is in the Blockchain, they cannot be deleted from the original place based on the —append only‖ feature on Blockchain. Even if the scanned copy of deed is replaced with a forged one, since the hash value of the original certificate is in the Blockchain, frauds can be detected. Digital signature of the Registration Clerk is added with Blockchain token to maintain the accountability.

Data loss is also addressed because even if the database crashed and become unavailable, the Blockchain transactions will be there. Since every node in Land registry has a copy of Blockchain, even if 99% of nodes are down, system can be still recovered with the use of the single node which was not down.

Since Blockchain is decentralized and every node has a copy of Blockchain real-time accessibility for verification also can be promised because of availability is greater.

Even though a portal is there, it can be queried using the Blockchain transaction ID. Therefore only the people who have that ID can access. Blockchain transaction ID is not assumable because it is a complex number. So privacy of the folio records are ensured with the owner.

Through the portal, real-time verification of folio records can be done by searching using the Blockchain transaction ID received with the registered deed.

Merkel tree is there in the block structure. So if one token in a block is forged, that particular token can be identified by going through hash values of Merkel tree.

**Electronic Health Records System Design**

Along with the Sri Lankan e-government initiative, a separate project has been undertaken to automate the health record management processes as well. Even Though the project is not completed the existing open-source medical records management system is operational in several hospitals covering end to end processes of a hospital [32]. The existing e-health record management system in Sri Lanka operating with standalone databases which maintains the data relevant only to that particular hospital. Hence the utilization of those medical records are only limited to that particular hospital. However it should be highlighted that ICTA is operating with plans to create a national patient database for the
country [33]. Hence at the latter stages all those standalone systems will be integrated with each other [34].

When analyzing the existing system, several issues were identified. However it is clear that the privacy of the patients is one of the prominent issues. In the electronic health record management system we view privacy as the system’s ability to hide the patients’ identities while interacting with the system. Patients should be given the opportunity to obtain medicine without exposing their personal identities. According to electronic health record management system scenario; we cannot actually restrict the doctors accessing medical records since it will restrict the doctor’s primary objective of using a such kind of a system. However when it comes to personal details, [ eg: Name, Address, DOB ] still efforts can be made to restrict accessing those personal information without prior consent of the patients.

It is true that hiding patients’ identities may not be applicable in all clinical situations. For an example, in the normal clinical situations hiding patient’s identities may not be suitable. But there are some clinics where hiding patients’ identities are necessary.

Furthermore, in the existing system the patients have no control over their own health records. All the records are entered and maintained in the system by the hospital staff. Hence a particular patient does not have any idea about who accessed medical records, who updated medical records, for what purposes they are used. Furthermore in the Sri Lankan public health management system context; the data is owned mostly by the hospitals. Apart from the doctors other parties such as data entry operators also engage in managing the patient data in a hospital. Hence those parties can also access the personal records of the patients and can view the medical status of the patients without patients’ consent.

Hence the proposed Blockchain based solution is mainly focused on protecting the identity of the patients while interacting with the system and to allow the patients to control their identity related personal data. If the identity is controlled by the patient; patient can decide to whom he can allow to see his/her personal details. Hence doctors can only see those personal information if the patient is allowing him to do so. On the other hand in our proposed Blockchain solution; any person who have access to the medical records may not be aware about to whom a particular medical record belongs to. Hence without interacting with the patient, no one can isolate the medical history of a particular patient.

Proposed Solution

The proposed solution we attempt to separate the medical information from the identity related information of the patients. In the process of separating those medical information and identity related information, it is important to note that it is not feasible to go for a complete Blockchain based solution due to the highly volatile nature of the medical records and specifically due to the storage concerns [35]. Hence we do not propose to store all the data in a Blockchain. The proposed system is equipped with three main components.

1. **The medical database** - contains all the medical information of patients
2. **The Blockchain** - contains all the identity related data of patients
3. **Patient App/Card** - Contains the unique symmetric key of the patient

The Blockchain will only contain the identity related details of a patient along with the Patient ID [PID] which is the primary key of the existing database. All the personal details (such as name, address) and the PID of that particular patient will be encrypted separately using the unique symmetric key of that particular patient and will be stored in the Blockchain as a single transaction. For each patient there will be a separate SID created in the initial data input. Whenever the personal
details need to be altered it will be recorded as a separate transaction with the same SID. It is important to highlight that PID of a particular patient is also stored as an encrypted value. Every patient will be issued a unique symmetric key. That symmetric key will be held by the relevant patient in a mobile app or in a patient card. This key will be only kept by the patient and no one will have the access to that particular key. Even to access the medical records of a particular patient a doctor will have to obtain the PID of that patient through the Blockchain. In order to do that the patient will have to provide his symmetric key to the doctor. Since the patient's details will not be disclosed to any party without the consent of the patient. Any person who has the full access to the Blockchain will see a list of encrypted transactions along with SIDs. They cannot directly identify the PID of a particular patient by just looking at the Blockchain.

**Table 1 Key values and data stored in the main components**

<table>
<thead>
<tr>
<th>Data</th>
<th>Blockchain</th>
<th>Medical DB</th>
<th>Patient App/Card</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Details</td>
<td>Personnal Details (encrypted)</td>
<td>Medical Data</td>
<td>SID</td>
</tr>
<tr>
<td>PID (encrypted)</td>
<td>PID</td>
<td>PID</td>
<td>Unique Symmetric Key of the patient</td>
</tr>
<tr>
<td>SID</td>
<td>SID</td>
<td>SID</td>
<td></td>
</tr>
</tbody>
</table>

The medical database will contain all the medical data of the patients. The existing medical database will be used with one modification to the system. In the existing system all the personal details of patients are stored in the —Patient table. In our approach we will only add the PID to the Patient table in the relational database. All the other details will be stored in the Blockchain. Hence the existing —Patient table will be altered and restricted only to store PID. However it should be highlighted that the relations in the existing electronic health record management will not be affected. Due to the availability of the key field —PID! all the relations in the RDBMS is not compromised. Only difference would be that there will be no personal details stored in the —Patient table. Hence anyone who has the full access to the —Patient table cannot see any personal details such as Name, address, etc. There will be only a single column to store the PID values. However the relationships in the database will remain unharmed since the primary key —PID! is not removed. Due to the availability of PIDs the medical practitioners can continue the use of existing electronic health record management system as it is. However it should be highlighted that even a person has the full access to the medical database; he can only see a list of PIDs and relevant medical records. They can't specifically identify a particular person since the personal details are not stored in the database.
Patient Creation Process Once a data entry operator submits the personal details of a patient to the system, the system will issue a new PID and a unique symmetric key for the patient. The symmetric key will be a randomly generated number which contains twenty alphanumeric characters. All the personal details (such as name, address) and the PID of that particular patient will be encrypted separately using the unique symmetric key and will be stored in the Blockchain as a single transaction. Furthermore exposing the PID to outside world is not acceptable. If anyone can see the PID of a particular patient printed in a patient card; anyone can query and see the medical history even though his identity is not known. Hence in order to identify different transactions which are unique for a particular patient in Blockchain we have introduced a separate key value as SID. SID is a a randomly generated number which contains ten alphanumeric characters. It will act as the key value for the records in the Blockchain.

Patient Data Retrieval Process Whenever a doctor wants to access the medical records of a patient, doctor should obtain the PID of that patient. The doctor can make a request to provide the PID. If the patient agrees to give away his/her PID and personal details he/she can provide the symmetric key. Patient will have to give the permission to access the PID by giving away the barcode card or by providing the permission through the mobile app. If the patient refuse to provide his/her key then no one can obtain the patient's PID or the personal information. Once the symmetric key is provided the doctor can obtain the relevant PID from the Blockchain and use it to query data from the medical database directly.

Derived Data Management Framework

When selecting a specific Blockchain platform to build a newer Blockchain project or to implement it from the ground level, for public sector data management, there are different aspects that need to be considered such as the nature of the Blockchain and extent of usage. Based on existing literature and the three designed systems, we designed a generic guideline that can be used to select a suitable Blockchain approach for public sector.
The Generic Framework/ Guideline we designed for Public Data Management of Sri Lanka comprises of six components and five main steps that should be considered when implementing a Blockchain solution for a public data management system. The six components are namely,

- Nature of Blockchain
- Extent of Blockchain usage
- Mining
- Node involvement
- Infrastructure
- Issues and Concerns

**Figure 4 Derived Data Management Framework**

**Nature of Blockchain**

There are four main categories of Blockchain as, Public Blockchains, Private Blockchains, Public Permissioned Blockchains, and Private Permissioned Blockchain

In a **Public** Blockchain, any node can read and write the data stored on the Blockchain. Any person can become a member of the Blockchain network to store, send and receive data after downloading the required software on his device. A Public Blockchain is completely decentralized where the permissions to read and write data onto the Blockchain are shared equally by all the connected users, who come to a consensus before any data is stored on the database [36]. In a **Private** Blockchain, the permissions to write data onto the Blockchain are controlled by one organization which is highly trusted by the other users. This organization may/may not allow users to have access to read the data, as public readability might not be necessary in most cases.

Limited/restricted read permissions also provide a greater level of privacy to the users, a feature not available in Public Blockchains. The organization in control has the power to change the rules of a Private Blockchain and may also decline transactions based on their established rules and regulations. In a Private Blockchain, the transactions are quicker as they can be verified through consensus by a
small number of devices. There number of people verifying the transaction is fewer than in a Public Blockchain [37]. A Permissioned Blockchain is a hybrid of a public Blockchain and a private Blockchain. It often referred as a Consortium Blockchain, where instead of allowing any person to participate in the verification of the transaction process, a few selected nodes are used for mining and verification [36].

The Blockchain solution which is ideal for data management in government organization as we suggest is a private permissioned Blockchain based on several reasons. One of the main reasons behind the suggestion is the extent of control, of Blockchain that should be given to a particular organization/government institute. Since the information stored on the Blockchain is sensitive and mostly personal information with regards to citizen of a country when allowing the permission to access the Blockchain, a form of access control is a necessity. And when it comes to the Mining functionality it needs to be under the authorization of the particular group of mining representatives of each public service delivery organization which will make a network of private miners making the Blockchain a private Blockchain. Since the designs does not involve a compensatory mining mechanism as well as the requirements of customized block structures with respect to each and every public service delivery system this is a necessity.

Extent of Blockchain usage Blockchain can be used for Data Management in several methods as,

- As a single Database option
- As an Access Control mechanism to limit access to a separate Database
- As a privacy preservation option
- As a verification repository

If the Blockchain is used as a single database solution, all data should be stored in the Blockchain itself. There the mining will be costly because of the increase of block size and the time it takes to conduct the mining function, since the mining effort required will increase with the block size or in other words the size of the data in a particular block [37].

If the Blockchain is used as an Access Control Mechanism, the critical data that needed to be protected against viewing/modifying by every user in system can be separated from a traditional database and that specific data can be stored in the Blockchain with a pointer to the original database, creating limited access to the original database.

If the Blockchain is used as a privacy preservation option, the sensitive data that need non-disclosure can be removed from the original database and store them in the Blockchain.

If there are documents or files that need the verification for their originality, the hash values of the documents or files can be stored in the Blockchain while storing the documents in the original database. If scanned images of certificates or documents are also included in the public data, rather than storing the particular image + rest of data as a record in Blockchain, it is more efficient to store the image and non-critical data in a separate database and storing the hash value of image with other critical data in Blockchain. The reason is that as mentioned in the above section, the increase of block size result in the increase of mining effort [38]. If the cost that needs to be incurred because of the requirement of high mining effort is bearable, images can be stored along with rest of data in Blockchain.
Mining

From the existing mining techniques, the ideal mining for the public data management is the customized proof of work to mine digital asset data. That is because the content of a block is depending on the particular data management process. Therefore the mining should be customized according to a particular public service delivery system process and based on the information that a particular service need to include in a block and its block structure.

Node Involvement

Any computing device that is connected to the Blockchain network is called a node. For a Blockchain network, there are several possible types of nodes as,

- Mining Nodes
- Nodes able to Query only
- Nodes with data adding functionality

For Data Management of Public Service organizations, Nodes with data adding functionality is offered to the data entry points at a particular Public Data Management Institute. The external entities that need to query the Blockchain can be offered with the Only Query functionality and the respective records can be searched according to the values that will be assigned as transaction ID in each design with respect to different public service management systems. The mining nodes can be offered to the pool of permissioned miners representing different branches of same government institution such as Divisional Secretariat offices and Hospitals.

Infrastructure

In Order to deploy a Blockchain network different types of infrastructure needed. The mining nodes are needed to be consist of hardware with high capacity CPU and Graphic cards. Special mining software are needed to be developed with the mining protocols to be in the verification. Client nodes can run on standard Desktop specifications. If the volume of data to be stored in the Blockchain is huge, high capacity storage devices are needed for every node which holds a copy of Blockchain.

Issues and Concerns

There are concerns and issues specific to the type of Blockchain application. Even though the data inside Blockchain are secure, physical access to the devices holding the application should be limited only to the authorized parties simply because nature of data that is being used in public service like the personal and private data of a specific individual or a party. And simply because we need to be aware of the secureness of the data entry stage because it determines whether the information feed into the Blockchain database is accurate or not and that has a huge effect on the reliability of the information and the verification capability provided by the Blockchain. This concern leads to why we need a permissioned Blockchain for the Public service delivery systems, because of the fact that we need a point of data entry or access with an individual who will be held responsible for the correct data entry at the most initial stage so that the benefits of the Blockchain characteristics will be able to incorporate into the process thereafter. After selecting a Blockchain based on the generic guideline it can be implemented and customized for a particular Data Management process. For different data management applications separate Blockchains can be used or they can be stored in a single Blockchain under different streams [39]. The five key steps to be followed is mentioned below.
Step 01: Identify Nature of Public Service

Different public services have various concerns on decentralization of data, privacy of data, data loss and possibility of frauds and manipulation in different magnitudes. And also in each and every public service delivery system the data that should be stored in a Blockchain will vary depending on the data size, criticality, and sensitivity of data and how frequently this data is being accessed or used. Therefore when designing a particular Blockchain based design for a specific process a thorough analysis need to be conducted to identify what aspects need to enhanced and what aspects should be compromised. Systems like Land Title Management are concerned much on the possibility of fraud and manipulation than the privacy preservation. Systems like Electronic Health Record Management System value the privacy preservation than the possibility of fraud and manipulation of data. Furthermore as the users of such public services may have different expectations for instance, to use the Land Title Management system as a way to prove land ownerships this creates the necessity of having data publicly available. On the other hand the users of an Electronic Health Record Management System would want their data hidden from other parties. Therefore, the specific nature of the Public Service should be clearly identified to select the characteristics of Blockchain that need to address the domain specific challenges.

Step 02: Identify Nature of Data

Public service delivery systems associates many categories of data based on their importance for the public service delivery process. There could be a data segment, which can be considered as critical. If that data is exposed to manipulation, the integrity of the service may fall into question or if that data segment is lost, the operation of service gets interrupted. Furthermore there may be some data that are sensitive or privacy critical where if they are exposed it could cause serious damage on the owners of such data. Frequency of refreshing of data and volume of data that gets added in a specific time period also should be identified as the nature of data. The component _'Extent of Blockchain usage'_ can be decided after identifying the nature of data. The specific nature of the data will decide the characteristics of Blockchain that need to be utilized to support data. A system like Land Title Management holds data that defines the ownership details for a particular land. Therefore such data record can be inserted into the Blockchain as a token or an asset. For data that gets update in very short intervals of time are not suitable for a Blockchain due to the mining difficulty. If the volume of data is huge, it takes a huge mining power and such data also not ideal for a Blockchain.

Step 03: Modeling of Data

This phase is focused on mapping the existing systems data structure with the new required data structure. If there is a Database System already available for the particular Public Service, completely replicating data into a Blockchain may be not needed. The data should be decomposed in such situation based on the identified nature of data. The critical data should be entered into the Blockchain while the rest of data can reside in the tables in the Database System. For example if there are scanned images of documents are in the existing system, rather than moving them to a Blockchain, hash values of the documents can be entered into the Blockchain for the mining convenience and used as a verification repository for such system. Frequently updating data suits to be stored in a database rather than adding to Blockchain, if they are not critical for the system. To establish the link between data records in Blockchain and the Database System, a common key should be used. The link between the Blockchain and the database should be defined according to the scenario of the public service. Based on the data that go into the Blockchain, the block structure should be designed. Block header will be almost common for all systems. It should contain the previous block hash, Nonce value, Block height, Block size, Version, Number of records and Merkle root. The maximum size of a record that should be entered into Blockchain and the block size should decide the number of records that could be inside of a single block. Block size can be determined based on the available mining capability.
**Step 04: Determination of additional security requirements**

Issues and Concerns component that is specific for the considering public service decide the security requirements. Usually when data is transformed to a Blockchain; all the data is publicly accessible and viewed by any node in the Blockchain. However depending on the requirements or nature of the public service exposing all the data without any control would be problematic. If the data entered into the Blockchain need to be hidden, for such situations options such as symmetric key encryption can be utilized. Even for the Blockchains different access levels can be provided based on the requirement.

**Step 05: Infrastructure modeling**

Based on the nature of Blockchain, the Infrastructure modeling should be done. For public data management, the ideal Blockchain would be a Private permissive Blockchain. For such Blockchain, node involvement should be decided based on the nature of involved users. As explained under the Node involvement component different nodes are needed for a Blockchain with different access levels. Number of different nodes needed for a particular branch of an Institute should be decided by the volume of data added per day and the number of system users. For mining, a pool of miners representing different branches of same Institute can be taken. Since almost all government public service processes are confirmed after the authorization of a responsible government worker who is assigned for the task, mining should schedule to be triggered after the authorization of such person. From examples considered, Land title blocks are mined when the number of Registrar Approved records fulfills the number of records that goes in one block. If there is a database functioning along with the Blockchain, keeping regular backups of database is need to be done to ensure the completeness of data. Since every government institute has branches representing different geographic locations, replications of databases also can be kept at other branches. However replicating the full database in several number of locations would be very costly. Hence the designers should identify the optimal number of replications that they need to maintain based on the industry standards and based on the nature of Blockchain. In case of an error, the backups or replications should be used to recover the data segment that is stored in the database.

**IMPLEMENTATION**

Based on the designs for Land Title Management and Electronic Health Record Management System, prototypes were developed to evaluate and verify the ability of Blockchain technology to overcome the considered four challenges. Platforms were used because building a Blockchain from scratch cannot be done within the research time period. Therefore when transforming a data management design into a prototype exact mapping was not achievable. To develop the prototypes two different platforms were used namely Multichain [40] and Hyperledger Composer [41]. The reason for the selection based on the preference of Blockchain application builders who use Blockchain to store digital assets. Most of them preferred Multichain and Hyperledger which are open source platforms. Land Title Management was developed using Hyperledger Composer and eHealth record system was developed using Multichain platform [40]. Hyperledger Composer is a Blockchain solution which can be easily adopted to store digital assets. Therefore, for Land Title Management System it was used. Multichain [40] is ideal for storing chunks of data and flexible to customize as it has a set of APIs that can be connected. Therefore, it is used for eHealth record System.
RESULTS AND FINDINGS

Developed prototypes were qualitatively evaluated for data accessibility concerns, data loss concerns, concerns related to frauds and errors and privacy related concerns. Land Title Management Prototype was evaluated for Data Manipulation and data loss avoidance possibilities and those tests got passed. Due to the limitations in the platform used to build it privacy preservation and accessibility providing couldn’t be tested. The electronic health record management system prototype which was developed using Multichain platform; was also qualitatively evaluated to verify whether it is capable to address privacy issue and the other three research questions. When addressing the privacy issue; the control of identity related personal data set of a particular patient is transferred to the respective patient. Hence the privacy issue is successfully addressed by the proposed solution using symmetric key capabilities along with Blockchains. The data manipulation issues were mainly addressed in the prototype by utilizing the append only nature of Blockchains and by decomposing patient data. Even if an intruder obtains full control over the Blockchain or medical database; intruder is not capable to uniquely identify and manipulate the records of a particular patient. In addressing the data loss possibilities and the data accessibility issues the decentralized nature of the Blockchains is utilized. All the identity related personal details are replicated among all the nodes of the prototype Blockchain nodes. Hence the data loss possibility is low and the accessibility is high. However it should be highlighted that medical records are not replicated among all the nodes.

CONCLUSIONS AND FUTURE WORK

Conclusion

Computerization of public data records managed by government is done under e-government initiatives of many countries including Sri Lanka. Up To now particular projects came up with different database solutions such as centralized, decentralized, virtual private and cloud databases for storage purposes. However these solutions have weaknesses where the data loss is still possible and data manipulation is also possible with correct attack. Therefore there is no single database solution which could address the difficulties in accessibility, data loss and possibility for fraud and error. Blockchain technology which emerged with Bitcoin showcased properties that could be used to address these issues that are currently being in question. However it’s essential to explore ways on how to actually incorporate Blockchain based mechanisms to the existing public service delivery systems mitigating the issues while considering on the limitations arise with the use of Blockchains. In this research several different ways of using Blockchain technology is being analyzed with relevant to the Sri Lankan context which proved to be effective to use as a database solution to create a data management framework for public data which ensure distribution, data availability and data security. From the initial information gathering, these four challenges were identified and they were taken as the objectives to be achieved in the research. Three public data management systems were chosen by considering the e-government initiatives that are already taken. Those selected systems were affected by the four considered challenges mostly. In research approach selected systems were designed using characteristics of Blockchain technology to address the challenges. Designs were theoretically evaluated for the ability to address the selected challenges. Abstract prototypes were developed based on the designs and they were qualitatively evaluated for the considered four criteria of challenges.

Qualitative evaluation passed the all four criteria considered although few limitations were recognized. However, evaluations of the prototypes were not precise due to the limitations in the platforms used to build prototypes. Since Blockchain is a complex concept, building a Blockchain from scratch is not possible within the research time period. Therefore existing two platforms were used which matched
the requirement. Number of existing open source platforms are less and their customization ability was also a bit inflexible

Future Works

The Blockchain solutions can be further scale into a single Blockchain for the Sri Lankan government where the critical information relevant to different public service delivery systems will be resides within a unique stream in Blockchain [39]. And also the wallet concept that we are presented in the Birth Marriage Death Certificate Management solution can be incorporated with the Land Title Management System where a single wallet may able to hold all real estate ownership for a particular person. If such wallet can be incorporated in Land Title Management, real estate transactions can be enabled and land Title token can be converted into a tradable object, like a Bitcoin. Furthermore for the designs that we have presented in the research incorporate proof-of-work mining concept, where there are many alternatives for different mining algorithms with their own pros and cons. There is a future work opportunity in identifying or developing a specific mining algorithm that will suit more to the requirements and the infrastructure capabilities of the government sector. And also one of the limitations we faced was the size of a particular Block in the Blockchain and the limitations of it. Therefore research opportunities are there to introduce a flexible Block size Blockchain or a Blockchain architecture where the Block size can be varies according to the transactions that are being recorded.

REFERENCES

[26] [K. Biswas and V. Muthukumarasamy, —Securing Smart Cities Using Blockchain Technology,l 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2016.].


[34] K. Wickramasuriya, "Hospital Health Information Management System (HHIMS) project status in Sri Lanka", ICTA - Colombo 05, 2017.


