УДК 004.031.43
DOI: 10.20535/SRIT.2308-8893.2017.3.01

BLOCKCHAIN AS A SERVICE FOR MEDICAL RECORDS

A. PETRENKO, R. KYSLYI, I. PYSMENNYI

Abstract. Storing and sharing healthcare data is challenging. Despite different data types that can be used on different platforms, also there is a question of security of gathered data. Storing data in a traditional way may cause data leak and it unavailability in critical moments. In this work we present decentralized way of storing patient data that can be used to avoid security and unavailability problems. A blockchain is a distributed and decentralized ledger that contains connected blocks of transactions. Unlike other ledger approaches, blockchain guarantees tamper proof storage of approved transactions.

Keywords: cloud computing, medical records, distributed databases, computer science, online care helping, sensor data storage.

INTRODUCTION

Patient data, such as medical history, tests results or other sensitive information should be stored in such a way, that patient himself, any doctor or clinical institution in which patient appeared should have access to it and possibility to use it. This will lead to significant increase of effectiveness of medical help, and research. Despite evidence that the value of healthcare data exchange is large, these issues, described below, remain significant barriers. Main idea is to provide access to patient data on demand — as needed (in case of emergency, planning examination etc.) to give doctors an instrument for making decisions based not only on current symptoms, but also on historical data. And at the same time provide patient himself with more relevant information about his health and possible issues. For this purposes clinical data have to be standardized according to international standards (e.g. ISO 12967 [2]).

Failing to secure the patient record has financial and legal consequences, as well as the potential to impact patient care. Securing patient data is a challenging task. Data privacy involves ensuring only authorized parties may access the medical record, especially considering that using blockchain technology means that full copy of all data are stored on different nodes across the network . This impacts any healthcare system, as patient privacy is not only an ethical responsibility, but a legal mandate. Patient data is also an asset to the institution, and unauthorized access could compromise competitive advantages or reveal proprietary practices.

The main goal of this work is to describe an approach to effectively and securely share healthcare information within all interested parties. We believe that a patient's record should be consistent and easy to access, and the terms of its access strictly dictated by the patient. As a secondary goal, this data should not only be shared, but shared in such a way that all interested parties should have an ability to use anonymized data for research purposes.

To meet these requirements we assume that user data has to be encrypted by any trusted cryptographic algorithm and accessed using digital signature of the patient.

BLOCKCHAIN

A blockchain is a distributed transaction ledger. The blockchain itself is composed of blocks, with each block representing a set of transactions. As a data structure, a blockchain has several interesting properties. First, blocks are provably immutable. This is possible because each block contains a hash, or numeric digest of its content, that can be used to verify the integrity of the containing transactions. Next, the hash of a block is dependent on the hash of the block before it. This effectively makes the entire blockchain history immutable, as changing the hash of any block $(n - i)$ would also change the hash of block n. The blockchain itself does not depend on a central, trusted authority. Rather, it is distributed to all nodes participating in the network. Because no centralized authority may verify the validity of the blockchain, a mechanism for reaching network consensus must be employed.

In one of the most widely used implementations of blockchain technology — Bitcoin, a Proof of Work function is used to ensure network consensus. This strategy requires that any node wishing to add a block to the blockchain must complete a computationally expensive (but easily verifiable) puzzle first. At a high level, this ensures consensus of the network because there is an opportunity cost (the computation time) to building a block. There are several other techniques used, such Proof of Stake and Proof of Activity, but all are designed to drive the network to consensus on blockchain validity. Miners are nodes that assemble the blocks and add them to the blockchain. It is through the miners that the consensus strategy is enacted, usually via some incentivisation protocol. In Bitcoin, for example, miners are incentivized by collecting transaction fees and also by a reward for adding the block to the blockchain. In general, however, there should exist an incentive for them to only build on top of valid blocks, which in turn drives the entire network to consensus.

A transaction has the following characteristics:

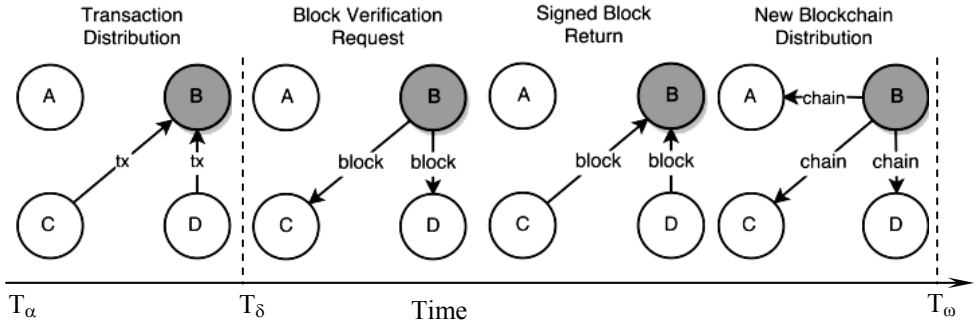
- hash: The SHA256 hash of the resource payload. Although the actual resource itself is not entered into the blockchain, its content can be verified using the transaction hash upon retrieval;
- contributor Signature: The digital signature of the originating node;
- record data itself;
- secure Index: An encrypted index allowing for data discovery without leaking information about the record.

The hashes of all transactions in a block contribute to the hash of the Merkle Root, or Block Header. The Block Header contains the following metadata used to validate the new block:

- hash: The SHA256 hash of the block. Assume the Merkle Root has two children c0 and c1, with a previous block bn−1. Let the hash of bn equal the hash of the hashed concatenation of the of the bn−1, c0, and c1 hashes;
- previous Block Hash: The hash of the previous block, for validation purposes;
- contributor Signatures: For each node that contributed to the block, a digital signature is required. This is to ensure that the block remains valid after it was assembled by the miner;
- miner Elections: Each node that contributed to the block is required to provide a random number encrypted with the node's private key. This will be used to seed the election of the next miner, which is discussed below.

ADDING DATA TO THE BLOCKCHAIN

For purposes of such system can be used different implementations of blockchain, currently we suggest to use hyperledger (as it is supported by key players in blockchain development, such as IBM, Huawei and Intel), figure.



Algorithm: Creating a new block and adding it to the blockchain

Hyperledger uses a NoSQL DB in the backend for storage of data (transactional info). Every transaction is identified by a unique txn id and the relevant info is encrypted, then mapped to this id and stored in the DB. Now if we want to query all txn in the same ledger, we must keep the txn id with us which we will pass to the NoSQL DB as keys. The DB will return all the relevant information about the transaction to you. On the high-level algorithm can be presented as follows (can be used with different implementations of blockchain):

input : A set N of Nodes participating in the network;

input : A blockchain, B representing a sequence of {b0...bn} where bn is the current (last) block on the chain;

input : T, the end of the Message Distribution phase;

1) $a \leftarrow$ ElectMinerNode(bn, N);

2) $P \leftarrow \{\}$; //Begin with an empty set of pending transactions.

3) while CurrentTime() < T do

foreach $n \in N - \{\alpha\}$ do

$P \leftarrow P \bigcup$ GetTransactionsFromNode (n);

4) $b \leftarrow$ AssembleBlock (P);

5) $N0 \leftarrow \{ n \in N \mid (\exists t)\ [t \in P \wedge$ IsOrigin(n,t)]}; // $N0$ is all nodes with ≥ 1 transaction;

6) foreach $n \in N0$ do

SignBlock (B,n);

7) B0 $\leftarrow$ AddBlock(B,b)

8) foreach $n \in N$ do

DistributeBlockchain($B0,n$).

VOTING

To save consistency of data, we need to make voting schema different from classic (where block accepts when it is accepted by more than 50% nodes of the network). Only trusted nodes should have an ability to vote. In this implementation, trusted nodes can be identified by their public keys.

The vote object contains the details of the vote being made as well as the signature and identifying information of the node passing the vote:

1) node_pubkey;
2) signature;
3) previous_block;
4) voting_for_block;
5) is_block_valid;
6) invalid_reason;
7) timestamp.

SECURITY

As said above, data security is fundamental priority for the system. A multi-faceted approach to security for our proposed network includes:

Blockchain Encryption. Nothing in the blockchain should be stored in plain text. Public information, or information intended for all nodes in the network, is expected to be encrypted by a network-shared key, while sensitive information should be encrypted by the originating node.

Smart Contracts. Patients may authorize access to their record only under certain conditions or for a specific reason. This notion of the codification of usage agreements is called smart contracts. There is precedent for their use on a blockchain, and given the complexities involved with our healthcare use case, smart contracts will play an important role. The intent is to ensure that patient authorization is codified and executable — for example, a patient may want their data shared only for research of a certain type, or for a given time range. These smart contracts can be placed directly on the blockchain as transactions, providing not only assurances of validity but an audit mechanism as well.

Value of the system

Patient has control over the data.

Patient has no problems with reaching his data and sanding it to required specialist or institution.

More data can be used by specialist which may lead to more accurate diagnosis.

Data can be shared for research activities including clinical trials, enabling larger and more diverse patient populations.

SUMMARY

There are a lot of challenges in storying such sensitive information as medical records. This model will create a possibility to run general medical network that will allow benefiting all healthcare industry by using more accurate and fresher data — while patients will receive more accurate diagnoses, doctors will receive wider datasets for researches.

All this, will make possible to use wider different cognitive tools like IBM Watson for healthcare purposes.

REFERENCES

1. *BigchainDB*: A Scalable Blockchain Database. — Available at: https://www. bigchaindb.com/whitepaper/
2. *INTERNATIONAL* STANDARD ISO 12967-1. ISO 12967-1:2009(E). — Available at: http://hsevi.ir/RI_Standard/File/1551
3. *Marques Rodolphe.* How We Integrated With MongoDB / Rodolphe Marques. — Available at: https://blog.bigchaindb.com/how-we-integrated-with-mongodb-d6a45e776d6b
4. *Peterson Kevin.* A Blockchain-Based Approach to Health Information Exchange Networks / Kevin Peterson, Rammohan Deeduvanu, Pradip Kanjamala, Kelly Boles, Mayo Clinic. — Available at: https://www.healthit.gov/sites/default/files/12-55-blockchain-based-approach-final.pdf
5. *King Sunny.* PPCoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake / Sunny King, Scott Nadal. — Available at: https://peercoin.net/assets/paper/peercoin-paper.pdf
6. *Thomas Roy*. Fielding. Architectural Styles and the Design of Network-based Software Architectures / Roy Thomas. — Available at: https://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf
7. *Hyperledger* Whitepaper. — Available at: http://www.the-blockchain.com/docs/Hyperledger%20Whitepaper.pdf
8. *Ashish K. Jhaa* The use of health information technology in seven nations / Ashish K. Jhaa, David Doolanb, Daniel Grandt, Tim Scott, David W. Batese. — Available at: https://pdfs.semanticscholar.org/2b67/6d6013995a65cca36596d7d906233af32b39.pdf

Received 04.07.2017