Navigating Healthcare Through Challenging Times
D. Hayn et al. (Eds.)

doi:10.3233/SHTI210099

A Mechanism for Verifying the Integrity and Immutability of Tuberculosis Data Using IOTA Distributed Ledger Technology

Vinícius LIMA[a,b,1], Filipe BERNARDI[a,b], Rui RIJO[c], Jó UEYAMA[d] and Domingos ALVES[e]
[a] *Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil*
[b] *Bioengineering Postgraduate Program, São Carlos School of Engineering, University of São Paulo, São Carlos, Brazil*
[c] *School of Technology and Management, Polytechnic Institute of Leiria, Leiria, Portugal*
[d] *Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, Brazil*
[e] *Department of Social Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil*

Abstract. Background: Intensified research and innovation and rapid uptake of new tools, interventions, and strategies are crucial to fight Tuberculosis, the world's deadliest infectious disease. The sharing of health data remains a significant challenge. Data consumers must be able to verify the consistency and integrity of data. Solutions based on distributed ledger technologies may be adequate, where each member in a network holds a unique credential and stores an identical copy of the ledger and contributes to the collective process of validating and certifying digital transactions. Objectives: This work proposes a mechanism and presents a use case in Digital Health to allow the verification of integrity and immutability of TB electronic health records. Methods: IOTA was selected as a supporting tool due to its data immutability, traceability and tamper-proof characteristics. Results: A mechanism to verify the integrity of data through hash functions and the IOTA network is proposed. Then, a set of TB related information systems was integrated with the network. Conclusion: IOTA technology offers performance and flexibility to enable a reliable environment for electronic health records.

Keywords. Electronic Health Records, Health Informatics, Tuberculosis, Computer Security

1. Introduction

Measured by the number of people who die each year, tuberculosis (TB) is the world's deadliest infectious disease [1]. To achieve the milestones for reductions in cases and deaths, the World Health Organization (WHO) has defined the global strategies and targets for TB prevention, care, and control, as part of The End TB Strategy, which defines 3 pillars to reduce TB burden [2]. One of these pillars consists of incorporating

[1] Corresponding Author: Vinícius Lima, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil, E-mail: viniciuslima@usp.br

an environment supported by digital health to promote intensified research and innovation and rapid uptake of new tools, interventions, and strategies.

In Brazil, the Tuberculosis Ecosystem (SISTB) for monitoring tuberculosis treatment via the Directly Observed Treatment (DOT) strategy and patients' follow-up has been developed [3], consisting of several modules. Information about the treatment may be recorded online or offline, as well as remote self-registration of medication intake by the patient through the upload of a video for a health team (VDOT) [4]. Also, the SISTB is able to communicate and share health data with other national and international repositories through a semantic interoperability layer [5].

TB records are composed of a variety of health data, such as demographics, treatment, exams, daily registration of medication intake, and contacts tracking. These pieces of data are linked together and visualized according to the objective of the health professional. Therefore, TB health information systems are segmented in several sections, where each one presents data within their context.

Despite these advances in the attendance of TB patients, the interinstitutional sharing of health data remains a significant challenge. The ineffective process of sharing data in the health sector results, in part, from the lack of trust between providers and data interoperability between healthcare IT systems and applications [6]. Data consumers must be able to verify the consistency and integrity of data. Trust relationships often exist between providers in the network and/or healthcare organizations but are particularly difficult to maintain due to the lack of consensus that the data can generate if the parties do not use the same healthcare system with a shared provider directory [7].

To address such challenges and enable patient and health stakeholders' control and autonomy, solutions using distributed ledger technologies (DLT) may be adequate. DLT's also offer opportunities for clinical research, real-time access to individualized data and the ability to set permissions for accessing and auditing data to ensure their integrity and security [8].

A DLT is based on the principles of peer-to-peer (P2P) network and cryptographic primitives. Each member in the network holds a unique credential and stores an identical copy of the ledger and contributes to the collective process of validating and certifying digital transactions for the network [9]. Information is encrypted and digitally-signed to guarantee authenticity and accuracy. Blockchain is one example of a DLT, which became famous with the digital currency Bitcoin in 2008 [10], but there are other approaches with distinct operation.

In this sense, this work proposes a mechanism and presents a use case in Digital Health, in which the main goal is to allow the verification of integrity and immutability of TB electronic health records using a DLT-based approach.

2. Methods

This research seeks to address an open issue in security of health data. Based on the relevance of the theme, a research question was defined to drive the development of the solution, as follows: "How to deliver a mechanism to allow systems' users to verify the integrity and immutability of a given TB health information?".

Due to the sensitivity of data, one key challenge to be faced is to provide guarantees that a given record, e.g., a video or an electronic health record, was not corrupted. An approach deployed in the SISTB Ecosystem is proposed to enable the possibility of

consumer systems/users to check the integrity of a health record by using hash functions and a DLT.

Hash functions are algorithms that transform a message of random length and generate a fixed-length output that acts as a unique identifier of a message. Any minimal changes to the original message will lead to an entirely different hash value [11]. In this sense, the solution is relevant because it may increase reliability of stored data.

2.1. IOTA

An attractive alternative to traditional blockchain architectures for DLTs is the Directed Acyclic Graph (DAG). DAG ledgers can provide many gains over traditional blockchain, including performance, scalability and transaction costs. In this work, we adopted the IOTA, developed in 2015, with no trading fees, blocks, or mining. Based on the M2M (Machine to Machine) principle, IOTA was designed specifically for the Internet of Things (IoT) industry with an interconnected architecture through a tangled network, the Tangle network [12].

All the data flow management can happen in channels with the exchange of Masked Authenticated Messages (MAM). MAM protocol is one of the most remarkable features of IOTA. It allows data to be shared securely via an encrypted channel. In this way, it is possible to ensure that the recipient of a transaction receives data with integrity through a trusted source [13]. To decrypt data stored in transactions, the key must be known, so only authorized parties can access the data.

The IOTA community maintains two public networks, and each one has its own Tangle to which nodes can attach transactions: the Mainnet and the Devnet. In addition, a private node can be deployed and attached to the network.

3. Results

The main outcomes of this work consist in the following: i) the deployment of a private node in the IOTA Devnet public network; ii) a mechanism to perform verifications of integrity of data stored in the SISTB and in the VDOT App. The SISTB Ecosystem relied on IOTA technology to increase trust in data accessed by multiple actors.

The private node was deployed in a virtual server. The software, called Hornet [14], is provided by the IOTA community. Although being connected to a public network, owning a private node allows an organization to securely interact with the Tangle, with specific permissions and without depending on the availability of third-party nodes.

The IOTA node automatically exposes an Application Programming Interface (API) to enable interoperability with external applications. The API allows writing and reading operations, i.e., sending and retrieving transactions, as well as the communication through confidential channels with encrypted data (MAM).

To communicate with the node, an additional auxiliary API was developed in Node.js that directly connects to the node's API using programming libraries [15] maintained by the IOTA community. This API was able to authenticate and send/retrieve transactions to/from the Tangle and work as a bridge to the network. Finally, through specific endpoints of the interoperability layer available in the SISTB Ecosystem, authorized systems were able to connect to the auxiliary API to interact with the Tangle. Figure 1 summarizes the communication flow and the integration between components.

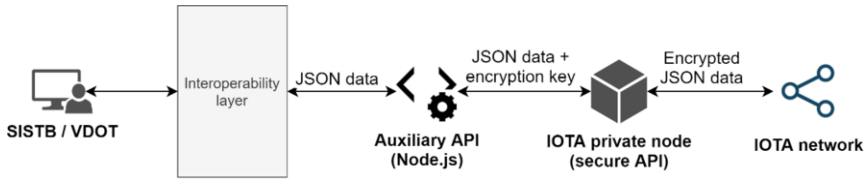

Figure 1. Communication flow

The SISTB Ecosystem needed the auxiliary API to interact with the Tangle. When an electronic health record was created or updated, or a video was submitted through the VDOT App, a data entry trigger calculated a hash for the data and called the auxiliary API, supplying the record ID, the hash and a timestamp as a JSON object. Then, a communication channel was established with the IOTA node, which effectively submitted the transaction (with data) to the Tangle. After successfully committing it, the transaction's ID was returned to the auxiliary API, which was stored it in the history of transactions for further reading operations. Figure 2 presents an example of data stored in the Tangle.

MESSAGE JSON


```
{
   "cod": "11245875",
   "hash": "E9E16315EAE1308D458BCBE079000C0632F5F2C35100DAF6B14330646E241536",
   "timestamp": "2020-12-07 12:34:05"
}
```


Figure 2. Data in JSON format stored in a transaction. This transaction can be verified at: https://explorer.iota.org/devnet/transaction/MLJOTLKQSVVXHJZBQCVMNBYSVMKCKCZBOMWYWGPEAERZZFDMQBHRCPPRXREHKIWFCQPOKANNGVIPEF999

Considering that this solution uses a shared network, it is important to keep data confidential in the Tangle. Therefore, the MAM protocol was used, which receives the message and a symmetric key known only by auxiliary API to encrypt and decrypt data. Systems from the SISTB Ecosystem do not know the key, because this API is the single point that writes and reads data from the Tangle.

Finally, a consumer system (or user) that wants to verify the integrity of a specific data can calculate the hash of the data by itself and compare it with the hash stored in the Tangle. If hashes match, it means that the data was not changed. When disclosing data for users, interfaces of the systems usually present the hash calculated in real-time, that is, when the user navigates to the page that contains the data, the transaction ID associated with that data and the hash extracted from the Tangle. If there is a mismatch, the user is alerted about a possible data corruption. Also, with the transaction ID in hands, the user can manually double-check the hash in the Tangle through an explorer tool provided by the IOTA community.

4. Discussion

IOTA technology is a free and open-source tool that enables a reliable environment for electronic health records underpinned by DLT features, such as data immutability, traceability and tamper-proof characteristics. Most existing studies on IOTA applications in health information systems have focused on conceptual designs and systems for verifying the integrity and immutability of electronic health records or on proposals with

managerial aspects of information [16]. Unlike the literature, a functional prototype was implemented to demonstrate the feasibility of the solution proposed in this work.

The main challenges of traditional blockchain-based solutions have been gradually overcome by IOTA Tangle and MAM. Despite being a constantly evolving solution, afforded by your active community, solid aspects in terms of cost, efficiency, scalability and flexibility in managing data access have been identified as advantageous in relation to their predecessors. Moreover, in health information systems, where the information is sensitive, the solution must comply with the latest regulations (for example, LGPD, a Brazilian GDPR). The immutability, traceability and tamper-proof characteristics reduce the risk of irregular processing and storage [17].

In this work, it is assumed that the party (SISTB) storing and sharing TB data can be trusted. One may say that, in this case, using a DLT tool becomes obsolete. However, the proposed solution seeks to protect a data consumer (other systems or users) from obtaining malicious or altered data in case of an unexpected security breach.

Finally, although a private node was deployed, a public network was used. The private node guarantees the connection with the Tangle with a dedicated and protected API. To increase availability, additional nodes can be configured for redundancy. Ideally, migrating for a private network is desirable, allowing a full control of nodes and participants.

5. Conclusion

This work presented the use of IOTA technology to deliver a mechanism to enable the verification of integrity of tuberculosis related data stored in the SISTB Ecosystem, enhancing trust on data in a reliable, safe and controllable way. IOTA offers performance and flexibility to enable a reliable environment for electronic health records.

Although the approach presented in this work was motivated by the TB scenario, its application is suitable in other health areas. The only requirement is to be able to calculate hashes for pieces of data in a given health information system to, then, write it into the IOTA network.

As future work, it is expected to establish a whole private network based on IOTA technology, so confidentiality and access control could be increased, which will allow the use of the network for additional purposes.

References

[1] Goldrick B. Once Dismissed, Still Rampant: Tuberculosis, the second deadliest infectious disease worldwide. AJN, Am J Nurs. 2004;104(9):68–70.

[2] WHO. A Global Action Framework for TB Research in Support of the Third Pillar of WHO´s End TB Strategy. 2015;75. Available from: https://www.who.int/tb/publications/global-framework-research/en/

[3] Crepaldi NY, Costa Lima V, Andrade Bernardi F, Albano dos Santos LR, Yamaguti VH, Pellison FC, et al. SISTB: an ecosystem for monitoring TB. Procedia Comput Sci [Internet]. 2019;164:587–94. Available from: https://doi.org/10.1016/j.procs.2019.12.224

[4] Albano dos Santos LR, Andrade Bernardi F, Prado GCS, Costa Lima V, Crepaldi NY, Marçal MA, et al. The perception of health providers about an artificial intelligence applied to Tuberculosis video-based treatment in Brazil: a protocol proposal. Procedia Comput Sci [Internet]. 2019;164:595–601. Available from: https://doi.org/10.1016/j.procs.2019.12.225

[5] Pellison FC, Rijo RPCL, Lima VC, Crepaldi NY, Bernardi FA, Galliez RM, et al. Data Integration

in the Brazilian Public Health System for Tuberculosis: Use of the Semantic Web to Establish Interoperability. JMIR Med Informatics. 2020;8(7).

[6] Cichosz SL, Stausholm MN, Kronborg T, Vestergaard P, Hejlesen O. How to Use Blockchain for Diabetes Health Care Data and Access Management: An Operational Concept. J Diabetes Sci Technol. 2019;13(2):248–53.

[7] Zhang P, Walker MA, White J, Schmidt DC, Lenz G. Metrics for assessing blockchain-based healthcare decentralized apps. 2017 IEEE 19th Int Conf e-Health Networking, Appl Serv Heal 2017. 2017;2017–Decem:1–4.

[8] Liang X, Zhao J, Shetty S, Liu J, Li D. Integrating blockchain for data sharing and collaboration in mobile healthcare applications. IEEE Int Symp Pers Indoor Mob Radio Commun PIMRC. 2018;2017–Octob:1–5.

[9] Wüst K, Gervais A. Do you need a Blockchain? IACR Cryptol ePrint Arch [Internet]. 2017;(i):375. Available from: https://eprint.iacr.org/2017/375.pdf

[10] Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008; Available from: https://bitcoin.org/bitcoin.pdf

[11] Teh J Sen, Tan K, Alawida M. A chaos-based keyed hash function based on fixed point representation. Cluster Comput [Internet]. 2019;22(2):649–60. Available from: https://doi.org/10.1007/s10586-018-2870-z

[12] Bhandary M, Parmar M, Ambawade D. A Blockchain Solution based on Directed Acyclic Graph for IoT Data Security using IoTA Tangle. Proc Fifth Int Conf Commun Electron Syst (ICCES 2020). 2020;(ICCES):827–32.

[13] Bhandary M, Parmar M, Ambawade D. Securing Logs of a System - An IoTA Tangle Use Case. Proc Int Conf Electron Sustain Commun Syst ICESC 2020. 2020;(ICESC):697–702.

[14] Community TI. HORNET - The IOTA community node [Internet]. Available from: https://github.com/gohornet/hornet

[15] Community TI. iota.js GitHub repository [Internet]. Available from: https://github.com/iotaledger/iota.js

[16] Zheng X, Sun S, Mukkamala RR, Vatrapu R. Accelerating Health Data Sharing : A Solution Based on the Internet of Things and Distributed Ledger Technologies Corresponding Author : 21.

[17] Hawig D, Zhou C, Fuhrhop S, Fialho AS, Ramachandran N. Designing a distributed ledger technology system for interoperable and general data protection regulation–compliant health data exchange: A use case in blood glucose data. J Med Internet Res. 2019;21(6):1–13.