Mutichain Enabled EHR Management System and Predictive Analytics

Chapter · January 2020

DOI: 10.1007/978-981-15-0077-0_19

7 authors, including:

Meghana Kshirsagar
University of Limerick
72 PUBLICATIONS **572** CITATIONS
SEE PROFILE

Aditya Patil
Government College Of Engineering Aurangabad
1 PUBLICATION **8** CITATIONS
SEE PROFILE

Gauri Vaidya
University of Limerick Ireland
15 PUBLICATIONS **58** CITATIONS
SEE PROFILE

Mayur Rahangdale
Government College Of Engineering Aurangabad
1 PUBLICATION **8** CITATIONS
SEE PROFILE

Mutichain enabled EHR management system and Predictive Analytics

Meghana Nagori[1], Aditya Patil[1], Saurabh Deshmukh[1], Gauri Vaidya[1], Mayur Rahangdale[1], Chinmay Kulkarni[1] and Vivek Kshirsagar[1]
Government College of Engineering, Aurangabad, Maharashtra, India-431005

adityacpatil6598@gmail.com[1], saurabh.98desh@gmail.com[1], gsvaidya2608@gmail.com[1], rahangdalemayur@gmail.com[1], chinmaypk15@gmail.com[1], kshirsagarmeghana@gmail.com[2], vkshirsagar@gmail.com[2]

Abstract. One of the challenges in biomedical research and clinical practice is that we need to consolidate tremendous efforts in order to use all kinds of medical data for improving work processes, to increase capacity while lessening costs and enhancing efficiencies. Very few medical centers in India have digitized their patient records. Because of less interoperability among themselves they end up having scattered and incomplete data. Health data is proprietary and being a personal asset of the patient, its distribution or use should be accomplished only with the patient's consent and for a specific duration. This research proposes multichain as a secure, decentralized network for storing Electronic Health Records. The architecture provides users with a holistic, transparent view of their medical history by disintermediation of trust while insuring data integrity among medical facilities. This will open up new horizons of vital trends and insights for research, innovation, and development through robust analysis. The platform focuses on an interactive dashboard containing year, month and season wise statistics of various diseases which are used to notify the users and the medical authorities on a timely basis. Prediction of epidemics using machine learning techniques will facilitate users by providing personalized care and the medical institutions for managing inventory and procuring medicines. Vital insights like patient to doctor ratio, infant mortality rates, and prior knowledge of the forthcoming epidemics will help government institutions to analyze and plan infrastructural requirements and services to be provided.

Keywords: Electronic Health Record, Predictive Analysis, Multichain, Decentralization, Data security.

1 Introduction

In 2014, according to the latest NSSO report, 44 out of each 1,000 Indians wind up getting hospitalized in a year. In 2016, the total number of hospitals in India reached 1,96,312, not all of which store data digitally [1]. Interoperability among the few or-ganizations that use the Electronic Health Record (EHR) systems is difficult to achieve due to the differences in their technological tools used.

An essential prerequisite for giving good patient care is the guarantee of the quality information in clinical health databases. Keeping the records on centralized servers makes them vulnerable. Hackers recently breached 1.5 million Singapore pa-tient records, repeatedly targeting their Prime Minister [2]. It requires vast efforts to gather all the invoices and send them to insurance suppliers, pharmacists etc. as infor-mation is not accessible on a solitary digital platform. In critical applications, absence of dependable information and sluggish interfaces proves devastating. Hospitals are wasting ample resources into duplicating the work that can be simply achieved by a refined framework and by expanding the proficiency of inadequately structured frameworks[3].

To improve the efficiency of patient care conveyance, healthcare parties must exchange the patient information among themselves, independent of their organiza-tional and technological particularities. With the headways in innovations, there is a scope to increase the efficiency of Electronic Health Record Management, making di-agnosis increasingly accurate and faster. Consolidating these advancements in health-care and information technology would stimulate prodigious change in the healthcare industry [4].

Blockchain's potential to accelerate and enhance research and development, care conveyance with fewer expenses is tremendous. The platform houses an interac-tive dashboard containing all vital pieces of insights gained from the data like infant mortality rates, major diseases in an area, epidemics etc. provided, will help govern-ment officials to plan accordingly and take measures to try and eradicate these in fu - ture. The dashboard contains year, month and season wise statistics of various dis-eases which can be used to notify the user or the medical authorities on a timely basis.

2 Related work

Security, trust, traceability, and control are the promises of the blockchain, and we intend its use for storing sensitive health data and for the operation and management of supply chains. The most dire need to have such a platform can be elaborated and justified with these factual problems faced around the globe and the work done using similar platforms to overthrow them.

2.1 Verifying the Authenticity of Returned Drugs by tracking Supply of Au-thorized drugs in Supply Chain

In 2016, the US pharmaceutical manufacturer sales were $323 billion. The 2018 forecast revenue for the top ten global pharmaceutical companies (which represent half of the sales of the top fifty companies) is $355 billion. Based on these estimates, the saleable returns market at 2–3% of total sales is 7–10 billion. Instead of obliterating these exemplary drug shipments, pharmaceutical companies instead chose to re-sell them. However before they can re-sell these returned drugs, the pharmaceutical companies have a legal obligation to verify the authenticity of the returned drugs[5].

The Drug Supply Chain Security Act (DSCSA), in the United States stipulates that all US manufacturers must implement serialization or barcoding of drugs at a package level by November 2018. Also, by the same time next year, these serial numbers must be used to verify the authenticity of the returned drugs[5]. A far better approach is to have pharmaceutical manufacturers record the serial numbers of their packages on a blockchain, which serves as a decentralized and distributed ledger. Wholesalers and customers can then verify the authenticity of a drug package by connecting to the blockchain. The research aims to have a similar existing architecture while giving the user the power to control their own data.

2.2 Transparency and traceability of consent in Clinical Trials

Patient consent involves making the patient aware of each step in the Clinical Trial process including any possible risks posed by any procedure undergone. Traceability for stakeholders and transparency for patients is expected from the data of clinical trial consent that is altered[6].

The Food and Drugs Administration reports that relatively 10% of the prelimi-naries they screen include a few issues identified with assent gathering. Frequently there are reported cases of misrepresentation, for example, issues of predating assent records[7].

Tracking the intricate information stream with various differing partners and ar-chiving it progressively through a time stamping work process is a vital step toward demonstrating information consistency.

2.3 Incentivized access to medical data

MedRec is a system that gives a transparent and accessible view of medical history by prioritizing patient agency, proposed by Asaph Azaria. It provides proof-of-con-cept that uses blockchain as a mediator to health information. MedRec incentivizes medical researchers and healthcare stakeholders for mining against access to aggre-gated and anonymized medical data, as a byproduct of continuing and anchoring the system through Proof of Work.

An attainable and anchored shared network can be gathered just by giving huge information, to enable experts while interfacing with patients and providers. MedRec was approved just for therapeutic records [13].

3 Problem Setting

In this section we define the mathematical context of the architecture, Let U be a vector of users' details with attributes as age, gender, diagnostic result. The set of streams used as wallets to store the transactional details of the user is defined as W. We refer to the wallet of a user as Wu_i for every user $u_i \in U$. Let H with location, no. of doctors, no. of. beds as attributes be the vector of Hospitals' details.

$U=\{u_1, \ldots, u_m\}$

$W=\{Wu_1, \ldots, Wu_m\}$

$H = \{h_1, \ldots, h_n\}$

For given set of users and hospitals T is the set of transactions where each transac-tion is defined as t_k.

$T = \{t_1, \ldots, t_k\}$

$t_k = (Wu_m, h_n)$

t_k is the relation between the patient (or the user) and the hospital. Each transaction has timestamp as attribute. From the transaction and hospital vector, the hospital to patient ratio, doctor to patient ratio is calculated. Using the attributes of the hospitals, users and transactions, area specific diseases are predicted. Personalized care will be provided to patients by doctors based on the corresponding subset of transactional de-tails T.

4 Proposed Architecture

The platform uses Multichain API to build and deploy blockchain for implement-ing a shared database using administered distributed system. It also provides public-private key cryptography, transaction mechanisms and fault-tolerant mechanisms that can survive any malevolent actions, forming a secure and decentralized network for storing Electronic Health Records. MultiChain blockchain provide timestamping, au-thenticity and invariability to the streams which are used as a database which compre -hensively allows append-only operations.

The medical authority issues a wallet to the patient on his/her first visit, trigger -ing the creation of a stream in the background. A MultiChain blockchain could con-tain infinite number of streams, these streams act as wallets for the patients. The data published in each stream is stored by every node thus promising its security. The nodes can access any stream and perform actions like adding transaction i.e. publish reports to streams and retrieve from stream, either as text or a file of formats like pdf, jpg, png, dcm with limit of 64MB on-chain and 1GB off-chain.

Instead of the original data, we can set the hashes of the data within the transactions to manage the data scalability issue. Files are stored off-chain using lossless compression algorithms, digital tokens or hashes of which are stored on blockchain using SHA-256. Data is anonymized, i.e care is taken to protect the private and personal information of the patient, and is analyzed using machine learning techniques to predict vital insights and provide customized analysis- descriptive, predictive as well as prescriptive. A trustful, secure environment will be created across the healthcare industry for collaborative and meaningful data exchange among the health centers. It is a multi-node platform with hospitals, insurance companies, government officials as the nodes, thus aiming at bringing revolution in the healthcare industry in the country.

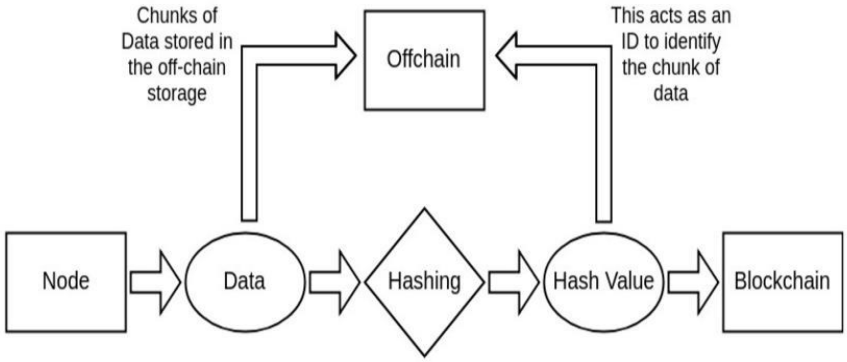

Fig. 1. Off-chain data in Multichain

4.1 Multichain

Multichain is an open source platform for the intra or inter organizational deploy-ment of private Blockchains. It is an augmented version of Bitcoin's Blockchain based on similar principles. It also lays accentuation on end-user decision enabling the user to control whether the chain is private or public. Also the user controls who can interface with the system, the objective time for blocks, the screening of individuals who can associate with the system along with extreme block size, and metadata. Inte-grated management and user permissions help in eradicating the problems of mining, privacy, and openness.

4.2 Off-chain storage

With large amounts of data generated, there raises a question of confidentiality and scalability. As the transaction copy is shared among all the nodes in the network, there would be a threat to data privacy. However, this could be solved by using en - cryption and compression techniques. The data will be encrypted using a key and then shared across the network. The nodes having the key will only be able to access the data.

While taking scalability into consideration, as multiple file formats can be uploaded on Multichain, it will be difficult to store the large amounts of data generated. The solution for this is, instead of storing the original data on chain, we can store the hashes of the data within the transactions. Thus actual data will be stored off-chain, ensuring the performance, speed and additional layer of security of Multichain.

4.3 Hashing

Hashing of data helps supplanting delicate information with a unique value that is not sensitive. This randomly generated non-sensitive value acts as a unique identifier and is the “Hash or token” for a sensitive record reducing the risk of unauthorized access. It can be used as a constant value, and used by many end users. This allows users to interact with the data directly, without having to unscramble and re-encrypt data each time they access the information in multichain platform.

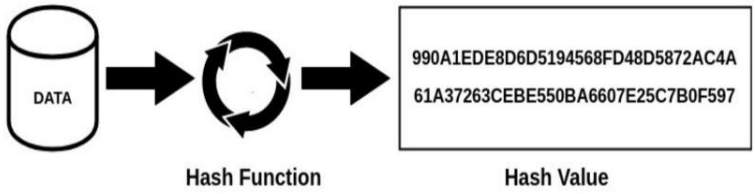

Fig 2. The hashing process

4.4 Data Analysis

Three stages for analysis on the data generated are descriptive, predictive and pre-scriptive.

1. Descriptive Analysis

This is the initial phase in data analysis which summarizes and organizes the data collected for better understanding. Such data is extremely helpful in forecasting vital facts. Data is turned into actionable insights and further studied in Predictive analysis phase.

2. Predictive Analysis

Above Descriptive analysis gives factual information that can help better adminis-tration, treatment and enhance diagnosis methodology. The information gathered is timely verified and is consumed by Machine Learning(ML) Engines for getting diver-sified predictions on the patient medical background. These predictions lead to impro-vised diagnostics and better healthcare.

3. Prescriptive Analysis

Every medical condition associated with some signs and symptoms is used as data to predict whether the features of corresponding conditions lie in a particular disease region, and if yes, then classify the disease and suggest preventive measures. Based on decision optimization technology, these capabilities enable doctors to recommend

best course of action for patients.

4.5 Machine learning

The estimation of machine learning in healthcare is its capacity to process immense datasets past the extent of human ability, and after that dependably convert ex-amination of that information into clinical bits of knowledge that guide doctors in ar - ranging and giving consideration, eventually prompting better results, bring down ex-penses of consideration, and expanded patient fulfillment.

The analytics produced will recognize and propose preventive measures for epi-demics, exact forecasts about an illness to happen, give appropriate medicare depen-dent on the medical background.

4.6 Workflow

The workflow of the system as shown in Fig. 3 can be summed up in following steps-

User's bio-metric verification, registration by authorized medical authority and is-suance of a digital wallet(i.e the health wallet) of the user.

Record storage on local servers using compression techniques to reduce storage space thus reducing related costs, every time a user visits a medical center with his/ her consent.

Anonymization of data and use of machine learning techniques on this anonymized data to predict vital statistics like spread of a disease, epidemics, etc.

The hash is stored on the chain where the transaction is verified and appended to the user's digital wallet.

User can now log in to the portal and can view his/her complete medical history along with analysis of his/her health profile.

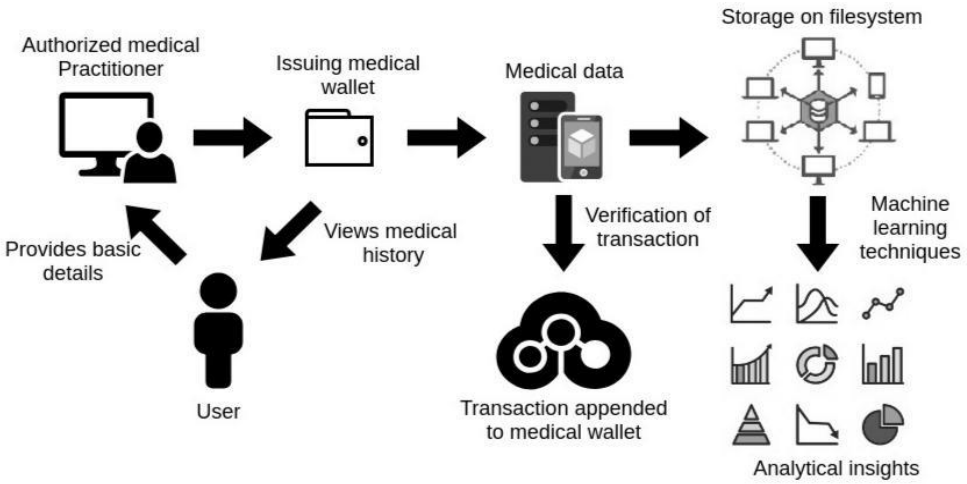

Fig. 3. The complete workflow

5 Conclusion

EHRs combined with analytics have an extraordinary potential to help clinical studies. A certain extent to reform clinical research and provide control to stakehold-ers which include patients, health systems, analysts, industry and society, can be guar-anteed by the proposed framework.

The platform is intended to be utilized and shared among assorted healthcare fa-cilities like research facilities, small clinicians and experts. It will help to collect in-formation from all professionals involved in the patient's care, in order to bring about decentralization in record management. Platform does not offer any Crypto asset for trading and is solely based on using the Open Source blockchain technology and re-lated frameworks to create a solution for storing and sharing Health records on dis-tributed network.

Machine learning is used in prediction of various diseases. Using algorithms, creating various self learning frameworks helps specialists for better and easier diag-nosis of different diseases. Choices identified with patient medical services can be made for the extent of productive results.

References

[1]. Kurian O. The Wire. [2018-03-25]: Why It's a Challenge To Make Quick Sense of India's
Health Data: https://thewire.in/health/imr-mmr-data-nss-srs

[2]. Xiao Yue, Huiju Wang, Dawei Jin, Mingqiang Li, Wei Jiang, Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control.
https://link.springer.com/article/10.1007/s10916-016-0574-6

[3]. Vikram Dhillon, David Metcalf, Max Hooper. Blockchain in Health Care: https://link.springer.com/chapter/10.1007/978-1-4842-3081-7_9

[4]. Laure A. Linn, Martha B. Koo, Blockchain For Health Data and Its Potential Use in Health IT and Healthcare Related Research:
https://www.healthit.gov/sites/default/files/11-74-ablockchainforhealthcare.pdf

[5]. Pharmaceutical Commerce May 2017 Article on US drugs 2016 Sales.
http://pharmaceuticalcommerce.com/latest-news/us-drug-single-digit-growth

[6]. Umesh Chandra Gupta, Journal on Informed consent in clinical research: Re-visiting few concepts and areas.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3601699/

[7]. Mehdi Benchoufi, Raphael Porcher, Philippe Ravaud on Blockchain proto-cols in clinical trials: Transparency and traceability of consent.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5676196/

[8]. Dr Gideon Greenspan, MultiChain Private Blockchain - White Paper
https://www.multichain.com/download/MultiChain-White-Paper.pdf

[9]. Asaph Azaria, Ariel Ekblaw, Thiago Vieira, Andrew Lippman, MedRec: Us-ing Blockchain for Medical Data Access and Permission Management. https://ieeexplore.ieee.org/abstract/document/7573685?reload=true

[10]. Four genuine use case of blockchain: https://www.multichain.com/blog/2016/05/four-genuine-blockchain-use-cases/

[11]. Off-chain storage via multichain https://www.multichain.com/blog/2018/06/scaling-blockchains-off-chain-data/

[12]. Satoshi Nakamoto on blockchain https://nakamotoinstitute.org/bitcoin/

[13]. Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System – Bit coin.org https://bitcoin.org/en/bitcoin-paper

[14]. Gupta A, Fraud and misconduct in clinical research: A concern. *Perspect Clin Res.*2013; 4(2): 144–7. 10.4103/2229-3485.111800 [PMC free article] [PubMed] [CrossRef]