

Survey on Medical Data Storage Systems



Suyog Gatkal, Vinayak Dhage, Dhanashree Kalekar, Sanket Ghadge

Abstract: Nowadays digital data storage and digital communication are widely used in the healthcare sector. Since data in the digital form significantly easier to store, retrieve, manipulate, analyses, and manage. Also, digital data eliminate the threat of data loss considerably. These advantages pushing many hospitals to store their data digitally. But, as the patients reveal their private and important information to the doctor, it is very crucial to maintain the privacy, security, and reliability of the healthcare data. In this process of handling the data securely, several technologies are being used like cloud storage, data warehousing, blockchain, etc. The main aim of this survey is to study the different models and technologies in the healthcare sector and analyses them on different parameters like security, privacy, performance, etc. This study will help the new developing healthcare systems to choose appropriate technology and approach to build a more efficient, robust, secure, and reliable system.

Keywords: medical, healthcare, communication, Data storage, privacy

I. INTRODUCTION

The healthcare sector is a very important sector and it decides the health of the citizens of the country. The availability of efficient health services is essential for maintaining and improving public health. It is important to provide quality health care to prevent, monitor, and treat diseases that reduce premature deaths. Effective healthcare can help boost the country's economy and growth by increasing worker productivity, raising the educational standards of the population, and reducing poverty. Also, a good and efficient network of health facilities enables you to fight epidemics and contagious diseases. Initiatives to improve health care across the region will help save large sums of money for people who are spending on illnesses. This sector helps the government to improve the physical and mental health of the country. The state of the healthcare sector in the country can be considered as a measure of the government's vigilance towards the citizens. Therefore, healthcare must be fast, reliable, appropriate, and streamlined which would be achieved by the digitalization of the sector. Digitalization of any organization increases the efficiency of the organization by enabling the digital tools to maintain analyse and modify data.

Manuscript received on September 09, 2021.

Revised Manuscript received on September 12, 2021.

Manuscript published on September 30, 2021.

* Correspondence Author

Suyog Gatkal, Jspm Narhe Technical Campus, Pune. (M.H.), India. E-mail suyoggatkal20@gmail.com

Vinayak Dhage*, Jspm Narhe Technical Campus, Pune. . (M.H.), India. E-mail vinayakdhage1234@gmail.com

Dhanashree Kalekar, Jspm Narhe Technical Campus, Pune. (M.H.), India. E-mail kalekar.dhanashree89@gmail.com

Sanket Ghadge, Jspm Narhe Technical Campus, Pune. (M.H.), India. sanketvg7299@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Digitalization also helps organizations to organize their services effectively and systematically which enhances the quality and the productivity of them. For example, a hospital can create and maintain a digital record of each patient which can be utilized by the hospital in successive visits of that patient. Speed and reliability of the healthcare sector are limited by the traditional data storage methods which contain the data storage in the paper document form. this method of data storage led to risks of data loss by natural calamities like floods earthquake, accidents like fire and human mistakes. Also, this data storage is methods need large manpower and physical space to maintain it. Speed of searching for particular information stored in non-digital form is one of the most time-consuming tasks if data stored is very large. the overall efficiency gets affected by this problem so digitalization is important in data storage of the healthcare system. Apart from this, the data stored in the paper form can be misused by a staff of the hospital who has access to it. Another important aspect of the healthcare management system is data sharing. Since traditionally each institution has its own patient information system, the data of the patient is distributed among the different institutes. Even if we create the centralized medical system which harvests the data from all the medical institutes, the patient needs to authorize the hospital since he is the owner of his personal information. But in a critical condition of the patient cannot authorize the hospitals which can endanger the life of the patient. In this paper, in section II we have discussed the different Survey Dimensions of the paper which consist design goal, Assumptions and Performance matrix. In section III, we did Literature survey in which we discussed some recent healthcare systems. In section IV we summarized the result and performance of these systems.

II. SURVEY DIMENSION

1.1 Design goal

Some design goals need to be met for system design. Frist is the system must be *Robust*. For the smooth working of the activities, the system must be Robust and errorless. The second goal is the system must provide *Security*. Since the hospitals have private data of patients, the system must be secure and only accessible to trustworthy peoples. The third goal is *Scalable*. Since the system may store the data of some hundred patients to few lakh patients, we need to make the system scalable. The fourth goal is *accessibility*. The system must be accessible from anywhere. Since hospital may be located anywhere even in the remote region the system must be able to provide the required information to that place. Another design goal is speed. since the system is going to be accessed by many clients at a time system need to be fast enough to serve the requested data in a reasonable amount of time.



1.2 Assumption

While designing the system we need to assume some requirements are available. A system is connected to the Internet. A system has a minimum required hardware configuration to perform the required tasks. A system needs to have a suitable Operating System installed in it along with required application software that are needed for working of the system. The system needs to have a continuous power supply for its seamless working. A system needs to have a staff equipped with the required skill set to operate it.

1.3 Performance Metrics

Access time is the important performance parameter for the data storage system since it decides the waiting time of a client system. Another performance parameter is no. of I/O request that the system can handle since this parameter decides the no. of clients a system can handle at a time.

III. LITERATURE SURVEY

In the survey, we have studied some recent systems that are proposed by the researchers in the last three years. we compared these systems on various parameters. In the paper, Erwin Halim et al. [1] proposed a system that can combine the records among the different healthcare organizations. This system uses the famous *Nonaka and Takeuchi model* of knowledge conversion for integrating the medical information about the patient, which is distributed at a different organization, into a single but complete medical record about the patient.



Fig. Nonaka Takeuchi model

In this model, first of all, a patient will convey the tacit knowledge to the doctor via interaction. This step is called *socialization*. In this step the doctor will get tacit knowledge from the patient about its health condition. In the next step, the doctor will try to make the digital record of the patient based on the current condition of the patient. In this phase, tacit knowledge is converted into explicit knowledge. This phase is called the *Externalization*. In the following step system will combine the information about the patient from the different healthcare organization and form a complete medical record of the patient. This step is called *Combination*. In the last step, the knowledge that is combined in the combination

process is converted into tacit knowledge again. This process is called as Internalization.

This system consists of several issues. Like anyone in a central authority have the access to the whole database of all patients. also, data can be easily tampered by anyone if he could get access to the central database since the data is centralized. Unlike in traditional medical system, where staff verifies the identity of the person and then disclose the medical records to the patient, in the proposed system if anyone get access to the password of the patient's account then he/she can access and tamper the medical information of a patient. In the paper, Hoai Luan Pham et al. [2] proposed a system based on the blockchain technology which uses the public ledger to store the transactions, since theoretically it is impossible to tamper the data in a blockchain in a reasonable amount of time. This system uses the Ethereum as an underlying technology to develop the system. Ethereum is a public decentralized, open-source blockchain that enables the smart contract functionality. just like any other contract, a smart contract is a contract between parties, which has all the conditions of the contract mentioned in it. The smart contract is a code written in some programming language (Mostly Solidity) that specify the conditions of the contract. This code gets executed when a certain trigger occurs. This code is stored in the public distributed ledger. This code cannot be changed by anyone but accessible to everyone due to the inherent feature of the blockchain. This feature allows the code to be non-changeable and transparent. Ethereum has its own currency to pay fees of these transactions called Ether. To complete the transaction, we need to pay "gas" as a transaction fee. If we give more gas for the transaction, the miner gets that fee as a bonus reward and hence, he will prioritize our transaction over the other transactions and hence the block will be added to the blockchain faster. This system is developed using several Smart contracts hence it is transparent and non-changeable. First of all Hospital (healthcare service provider) will create the Remote Healthcare System Smart Contract (RHSSC). Then the hospital will publish the smart contract address of that smart contract to all the doctors of that hospital and patients through any publishing medium like social media, Email, Website etc. After publishing this address Patient and doctor need to register themselves to a hospital through Registration Smart Contract. Now the hospital will authorize any doctor to monitor the patient health record. This record that the doctor is going to monitor, is generated by the sensor network on the body of the patient. These sensors will collect the data like heart rate, blood pressure etc. which can be used to predict the current health of the patient. Sensors spread across the body will send the data continuously to the smart device, mostly a smartphone. The smart device will perform the primary analysis on the data and predict whether the condition of the patient is stable or patient needs emergency health support. If the patients' health is normal, then the smart device will discard most of the data and only periodically record data into the blockchain.

Also, in a normal condition, the “gas” used for the transaction is very low, since it is not necessary to add the data into blockchain immediately which will save the transaction fee. When the smart device detects emergency for a patient according to data sensed by sensors, it will immediately try to add the data to the blockchain with high "gas" as a transaction fee. Due to high transaction fee block with emergency information will be added to the blockchain immediately. When emergency data is added to the blockchain system will immediately inform the corresponding doctor to help the patient along with body parameters of the patient and the location of the patient.

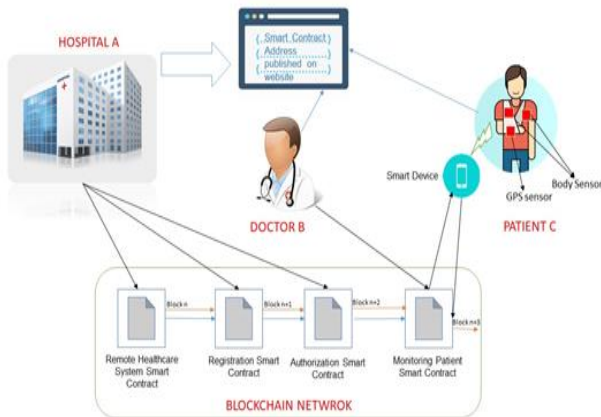


Fig. 2 Remote healthcare system using smart contract.

Issues in this system include large usage of space since data is stored in a public distributed ledger. Many copies of the same data stored over several computers spread over the world which is wastage of a large amount of space. The second issue with this system is if patients "gas" ended in an emergency condition then the transaction will not proceed and the patient's life will be in danger. The third issue with this system is, till now Ethereum uses the Proof of Work for validation of block in a blockchain. This makes the transaction slow blockchain network has a high amount of transaction in the queue. Apart from this proof of work is needs a great amount of computational power and electricity. Some countries also banned cryptocurrencies. hence this system can not be implemented in all the countries. In the paper, Yilong Yang et al. [3] proposed a system which aims at the integration, interoperability, innovation and intelligence for the sharing of patient's Electronic Health Record. In this system, information sharing is necessary for effective healthcare service and improving patient health. MedShare is a robust Electronic Healthcare System built for exchange medical information among hospitals which helps the patient who wants to switch healthcare provider due to some reasons. This is a query-based system developed on a hybrid cloud for data exchange in an emergency where we can get medication history, reports and records of patients. Hybrid cloud is a cloud system in which public and private cloud combined to take advantages of both. Private cloud of a particular hospital contains all Electronic Health Records of patients visited to that particular hospital. On the other hand, Public cloud which is monitored by the government, contains the patient's ID mapped with addresses of clouds of all the hospitals he visited before. The system consists of five entities: Patient, Hospital, Private Cloud, Public Cloud. When patient visits the Med Share enabled hospital A first time, hospital A will generate the patient ID which is universally unique for the patient ID. This ID is stored on to

the public cloud along with the address of the private cloud of the hospital A. Also in a private cloud, the Patient ID is stored along with other patient's data like Name,

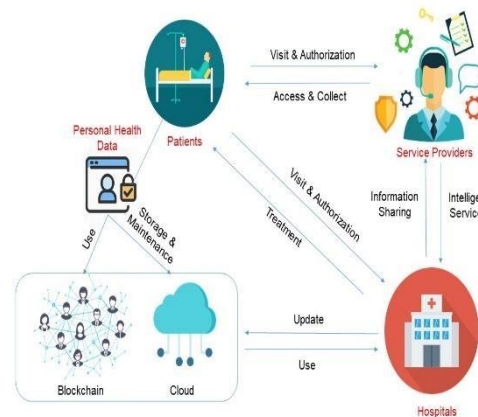


Fig. 5 Medshare

address, medical history etc. When the patient visits the same hospital A again, then record from the private cloud will be used to get the personal information of the patient and his/her medical history. But when that patient visits the hospital B where he never visited yet, He will be asked for his patient's id. Private cloud of that hospital will then request the public cloud for the patient information with that patient id. If the public cloud finds the Entry of that patient, then it will request for information of that patient to the private cloud using the address associated with that patient ID. Private cloud of Hospital A will search its repository for that patient ID and send the requested information to the public cloud. This information will be sent back to the private cloud of Hospital B which is now available for use in a Hospital B.

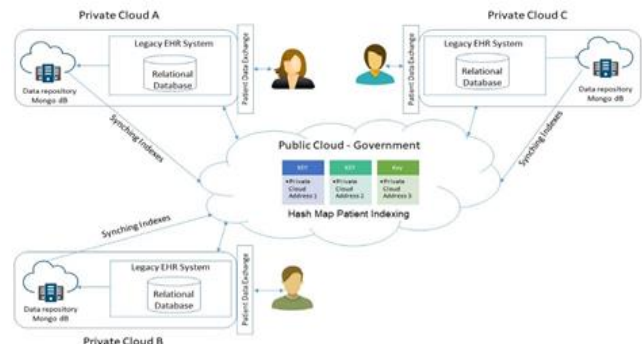


Fig 4 The architecture of medical blockchain system.

Issues with this system are system needed continues monitoring of the public cloud by the government because of the reliability of the data exchange is highly depends upon the public cloud. The second issue is the extra cost and memory required for data exchange from the Electronic Health Record format to the unified data format. The third issue with a system is many organizations will oppose sharing the data through the public cloud because of competition. In the paper, Yi Chen et al. [4] proposed the medical record management system built using blockchain to provide a distributed and decentralized way to store and manage the data. The main characteristics of this system are security, privacy, immutability, ease in sharing and ownership of patient over his data.

Survey on Medical Data Storage Systems

This system has these properties since the patient has full control over his/her health-related information.

In this system data of the patient is stored on the cloud and encrypted by the symmetric encryption key. This is the only key that can decrypt the data of the patient. The hash value of this data is stored in a blockchain along with the storage address of the patient's data and access permissions related to data. If someone tries to change the data in the cloud, its corresponding hash value gets changed and hence data get invalidated. The validity of this data can be checked by anyone, but since the data is encrypted, no-one can extract true information from that data. Also, a storage address, hash value and access permissions are accessible to everyone and immutable due to the inherent character of blockchain. When a patient goes to the hospital for his check-up, all the information generated about the patient like blood pressure, sugar, test reports etc. will be converted into digital format.

The hospital will create the digital signature of this data and then encrypt this data with the symmetric encryption key. After this encryption doctor will encrypt the symmetric encryption key with the patient's public key and send it to a patient along with encrypted data. No one in the way can decrypt the data in between since data is encrypted by symmetric encryption key and then the symmetric encryption key is in turn encrypted by the public key of the respective patient so it can only be decrypted by patient's private key. Once this data reached to the patient, he will decrypt the symmetric encryption key by using his private key and using this key he will decrypt the data sent by the hospital. Now patient will the data, using the digital signature of data which is available on blockchain created by a doctor. Once the data is validated patient will encrypt the data using symmetric encryption, create its hash value, store the data on the cloud. And finally post the block on the blockchain with information about data i.e., the storage address of data on the cloud, hash value of data, and access permissions. There are some issues with this system. One of the important issues with this system is even though the data is secure if someone gets access to the cloud through any means he can corrupt the data of patient since all the medical record is not stored on the blockchain. The second important issue with the system is during the key exchange the system is vulnerable to Man-in-The-Middle (MITM) attack since the one can easily send false data to the patient, pretending itself as a hospital/doctor.

IV.SUMMARY

1.4 Title: Indonesia Medical Knowledge Management System.

Description: This system is developed to integrate the medical history which is distributed among the different hospitals considering patient data privacy. Also, it enables the sharing of data among hospitals in an emergency with the permission of government authority.

Advantages:

1. Emergency support available.
2. Data privacy maintained.

Disadvantages:

1. Security is not up to mark.
2. Government authority can misuse the data.

1.5 Title: A Secure Remote Healthcare System for Hospital Using Blockchain Smart Contract.

Description: This system is developed to securely store the medical data collected by the sensors on the patient's body using blockchain smart contract. It also provides emergency support for the patients in a need.

Advantages:

1. Highly data tamper-proof system.
2. Very hard to attack the system.

Disadvantages:

1. High data redundancy.
2. Highly resource consuming and power-consuming System.

1.6 Title: Medshare: A Novel Hybrid Cloud for Medical Resource Sharing Among.

Description: This system is developed to share the data between the private cloud of each medical institute. The system uses the public cloud as an intermediate party to share the data between hospitals.

Advantages:

1. Data redundancy avoided
2. Reusability of hardware and software present in a different healthcare provider.

Disadvantages:

1. There is extra cost and storage required for conversion of Electronic Health Record format to the unified data model.
3. Public cloud requires continues monitoring of government.

V.CONCLUSION

In this review, we studied the medical system recently proposed by researchers in order to make medical systems more robust, secure and efficient. We tried to find out the advantages and disadvantages of each system over the other one so that one can choose a proper system according to their need.

REFERENCES

1. Nihar Ranjan, Midhun Chakkaravarthy, " A Brief Survey of Machine Learning Algorithms for Text Documents Classification on Incremental Database", Test Engineering and Management, ISSN 0193-4120, Volume 83, May-June 2020, pp 25246-25251
2. E. Halim, P. P. Halim and M. Hebrard, "Indonesia medical knowledge management system: A proposal of medical KMS," 2017 International Conference on Information Management and Technology (ICIMTech), Yogyakarta, 2017, pp. 322-327, doi: 10.1109/ICIMTech.2017.8273559.
3. H. L. Pham, T. H. Tran and Y. Nakashima, "A Secure Remote Healthcare System for Hospital Using Blockchain Smart Contract," 2018 IEEE Globecom Workshops (GC Wkshps), Abu Dhabi, United Arab Emirates, 2018, pp. 1-6, doi: 10.1109/GLOCOMW.2018.8644164.
4. Y. Yang et al., "Medshare: A Novel Hybrid Cloud for Medical Resource Sharing Among Autonomous Healthcare Providers," in IEEE Access, vol. 6, pp. 46949-46961, 2018, doi: 10.1109/ACCESS.2018.2865535.
5. Chen, Y., Ding, S., Xu, Z. et al. Blockchain-Based Medical Records Secure Storage and Medical Service Framework. J Med Syst 43, 5 (2019). <https://doi.org/10.1007/s10916-018-1121-1>
6. S. Khadka, "Privacy, security and storage issues in medical data management," 2012 Third Asian Himalayas International Conference on Internet, 2012, pp. 1-5, doi: 10.1109/AHICI.2012.6408435.

7. M. S. Saranya, M. Selvi, S. Ganapathy, S. Muthurajkumar, L. S. Ramesh and A. Kannan, "Intelligent medical data storage system using machine learning approach," 2016 Eighth International Conference on Advanced Computing (ICoAC), 2017, pp. 191-195, doi: 10.1109/ICoAC.2017.7951768.
8. Secure Storage Systems for Healthcare Records. In: Jonker W., Petković M. (eds) Secure Data Management. SDM 2007. Lecture Notes in Computer Science, vol 4721. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-75248-6_12
9. Nihar Ranjan, Midhun Chakkaravarthy, " Evolutionary and Incremental Text Document Classifier using Deep Learning", International Journal of Grid and Distributed Computing, ISSN 2005- 4262, Vol 14 No. 1, pp. 587-595
10. Nihar Ranjan, Kunal Phaltane, " A Survey on Techniques in NLP", International Journal of Computer Application, ISSN 0975- 8887, Volume 134 No 8, January 2016, pp 6-9
11. Nihar M. Ranjan, Rajesh S. Prasad, "Automatic text classification using BPLion-neural network and semantic word processing, Imaging Science Journal, ISSN 1368-2199, September 2017, pp. 1-15
12. Nihar Ranjan, Vishnu Panickar, " Text Document Classification using Convolution Neural Network", Journal of Emerging Technologies and Innovative Research, ISSN 2349-5162, Volume 7, Issue 6, June 2020, pp. 329-332

AUTHOR PROFILE



Sanket Vijay Ghade is currently in his final year of bachelor's degree in computer science. He is pursuing this degree from JSPM Narhe Technical Campus, Savitribai Phule Pune University, India. His area of interest is in Artificial Intelligence and Cloud Computing



Dhanashree Dhiraj Kalekar is currently in his final year of bachelor's degree in computer science. He is pursuing this degree from JSPM Narhe Technical Campus, Savitribai Phule Pune University, India. His area of interest is in Artificial Intelligence and Cloud Computing



Vinayak Anil Dhage is currently in his final year of bachelor's degree in computer science. He is pursuing this degree from JSPM Narhe Technical Campus, Savitribai Phule Pune University, India. His area of interest is in Artificial Intelligence and Cloud Computing



Suyog Chandrakant Gatkal is currently in his final year of bachelor's degree in computer science. He is pursuing this degree from JSPM Narhe Technical Campus, Savitribai Phule Pune University, India. His area of interest is in Artificial Intelligence and Cloud Computing