A Secure Electronic Health Record Storage System Based on Hyperledger Fabric, IPFS, and Secret Sharing Scheme

Puja Sarkar, Lopamudra Pathak, Rohan Molia, Sima Boro, and Amitava Nag

Abstract Electronic health records (EHR) systems are now becoming widely used as a system for storing and managing patients' health records. Patients should be able to access their health records as and when required. Traditional EHR systems have challenges with data security, integrity, interoperability, and management. The typical client–server system is vulnerable to single-point-of-failure since it is centralised. Medical data that is dispersed across different EHR systems is frequently difficult to access. Due to its positive features such as security, privacy, secrecy, and decentralized, blockchain technology has the potential to significantly enhance the healthcare sector, despite all of the existing problems. In this paper, we present a health data storage system based on the permissioned Hyperledger blockchain, an InterPlanetary File System (IPFS), and a multi-image secret sharing approach. Hyperledger Fabric is a peer-to-peer tamper-proof private-permissioned blockchain network that allows an organisation to simultaneously participate in multiple, separate blockchain networks via channels. And further, multi-image secret sharing is applied to provide security to sensitive medical image data during transmission against illegal access or alteration. Our system is integrated with IPFS for EHR storage. The proposed framework provides a secure, decentralised, and distributed model of EHR management system. Hyperledger Fabric is a tamper-proof private-permissioned blockchain network that allows an organisation to participate in many blockchain networks at the same time using channels. Furthermore, multi-image secret sharing is used to protect sensitive medical picture data from unauthorised access or manipulation during transmission. Using which our system is free from single-point-of-failure. For EHR storage, our system is connected with IPFS. The proposed framework provides an EHR administration system that is secure, decentralised, and distributed.

Keywords Blockchain-based EHR system · Hyperledger Fabric · InterPlanetary File System · Image secret sharing

P. Sarkar (✉) · L. Pathak · R. Molia · S. Boro · A. Nag
Central Institute of Technology, Kokrajhar, India
e-mail: p.sarkar@cit.ac.in

A. Nag
e-mail: amitava.nag@cit.ac.in

J. K. Mandal and D. De (eds.), *Frontiers of ICT in Healthcare*, Lecture Notes in Networks and Systems 519, https://doi.org/10.1007/978-981-19-5191-6_29

1 Introduction

An electronic health record is a health-related, extremely sensitive record used to diagnose and treat patients with care. The amount of patient data is increasing at a breakneck speed. The access, sharing, and sensible distribution of EHR among numerous healthcare stakeholders such as hospitals, clinicians, and patient families is one of the most essential and critical difficulties the healthcare industry is now experiencing. The electronic health record system has a variety of flaws, including the following:

Interoperability: The capacity of multiple information systems to communicate with one another is referred to as interoperability. The information should be able to be shared and put to various uses.

Information Asymmetry: According to opponents, today's biggest concern in the healthcare sector is information asymmetry, which refers to one side having better access to information than the other. This difficulty emerges in the case of EHR systems or the wider healthcare sector since doctors or hospitals have access to the patient's information, making them central. A patient must go through a lengthy and rigorous process in order to view his medical records. The data, which is centralised at a single healthcare organisation, is only accessible to hospitals or organisations.

Data Breach: Without the knowledge or authority of the system owner, information is stolen or taken from the system, requiring the need for a better platform in the healthcare industry [1, 2].

To overcome all of the aforementioned concerns, we devised a system based on blockchain technology. Blockchain is a distributed, decentralised ledger that records transactions in an ever-growing chain of unchangeable blocks linked by cryptographic hashes. Due to their transparency, blockchains can instantly detect fraudulent information, and smart contracts on the blockchain can operate as security measures [3, 4]. These smart contracts can record evaluations and provide data that can be utilised to design new and more effective treatments.

For patient privacy and confidentiality, such as health-related personal information, a private-permissioned blockchain is appropriate. It focuses on specific security and interoperability weaknesses and issues, as well as present EHR system roadblocks. With all of the advantages of blockchain technology, there are also some disadvantages.

Using a blockchain network to store massive volumes of data makes it slow and computationally expensive. So, in our system, we kept our data files separate in off-chain storage, and we employed an InterPlanetary File System (IPFS) to do so. Our objectives can be described as follows: A decentralised platform that would store a patient's medical records and allow physicians and other interested parties, such as the patient and the patient's trusted party (PTP), to access them. A patient- centric paradigm is one in which the patient determines who has access to their data and how it is used. In the event of an emergency, data can be easily shared with other organisations. Before uploading important EHRs to off-chain storage, make sure they are secure. This paper is organised as follows: Sect. 2 discusses comparable efforts by others. The preliminary knowledge of the technologies involved is covered in Sect. 3. In Sect. 4, the proposed system and workflow are discussed. Section 5 looks at the conclusion and security analysis.

2 Related Works

Shahnaz et al. [1] established a strategy for adopting blockchain technology in the healthcare sector for EHR to address issues including data leaks, information asymmetry, and scalability. The purpose of their suggested framework is to first embrace blockchain technology for EHR and then to provide secure electronic record storage by creating granular access controls for users of the proposed framework, as well as off-chain record storage, thereby addressing the scalability issue.

Tith et al. cite [2] presented a distributed system based on the Hyperledger Fabric that combines existing EHR. To protect a patient's privacy, they use a proxy re-encryption process when transferring data. Doctors can use the system to find patient records and check that the patient has given their consent to data access. The hospital has easy access to the patient's information. The access log is transparently and immutably maintained in the ledger for auditing purposes.

In Zhang et al. [5], Shamir's Secret Sharing was used to suggest a cloud storage solution for EHRs that fully secures data privacy by splitting the system into several components and distributing them across various cloud servers. Because reconstruction of a shared EHR might be time consuming during recovery, they propose a feasible cloud storage system that outsources it to a cloud computing service provider.

Kumar et al. [6] proposed a blockchain-based consortium structure for storing medical report details. They also describe a system for patient diagnostic reports that employs off-chain storage.

Sultana et al. [7] proposed a technique for dealing with vulnerabilities in medical/health data. The approach leverages blockchain's immutability, zero-trust principles' additional security, and the scalability of off-chain data storage via IPFS.

In paper Tanwar et al. [8], a blockchain-based approach for sharing electronic health records was presented. Many methods and configurations for block transactions are utilised on the network. A shared symmetric key and private key can be used to distribute the EHR to other users of the blockchain network.

3 Preliminary Knowledge

3.1 Blockchain

It is a decentralised digital ledger that records transactions without the need for a central authority in a public or private peer-to-peer network. It allows community members to record transactions in a shared ledger. Every transaction is recorded with a hash, which is a cryptographic signature that cannot be changed. Since the blockchain's inception, each blockchain node has a comprehensive record of all data recorded on it. In most cases, once a transaction has been published, it cannot be changed. Blockchain has a number of advantages, including no central administration, built-in transparency, distributed record-keeping, and immutability [9, 10].

3.1.1 Categorisation

Blockchain can be categorised into four categories: public, private, hybrid or consortium blockchain network.

1. **Public blockchain**: This has no restrictions as to who can use it. Anyone who wants to transmit, view, or modify to the database can do so. Anyone can send transactions to it and become a validator, hence anyone can use it.
2. **Private blockchain**: A private blockchain necessitates permissions. Only those who have been invited by the network's administrator are allowed to access it.
3. **Hybrid blockchain**: It is a hybrid of centralised and decentralised characteristics.
4. **Consortium blockchain**: A permissioned blockchain that allows many organisations to participate in decision-making, allowing for real decentralisation.

3.2 Hyperledger Fabric

Hyperledger Fabric is a private blockchain that allows organisations to collaborate in the formation of the blockchain network. It aims to advance blockchain technology. It is an open source framework implementation for private chain in which membership and roles are known to other members. Some of its features are [3, 4] (Table 1).

1. In a peer-to-peer blockchain network, it provides a private and adaptable framework for completing multiple transactions.
2. The adaptable, pluggable endorsement model aids in the realisation and attainment of consensus among network stakeholders.
3. It allows us to construct communication channels between different member organisations, allowing us to meet the privacy and security goals.
4. It uses channels to create a method that ensures transaction privacy and integrity.

Table 1 Distinction between permission-less blockchain and permissioned blockchain

	Permission-less	Permissioned
Access	Access to the database is enabled for both read and write operations	Database access with read/write permissions
Scale	Scale to a high number of nodes, but not at the expense of transaction throughput	Scale transaction throughput but not to a huge number of nodes
Consensus	Proof of work/proof of state	Algorithms for closed membership consensus
Identity	Anonymous/pseudonymous	Although node identities are known, transaction identities might be private, anonymous, or pseudonymous
Asset	Indigenous	Any data/state

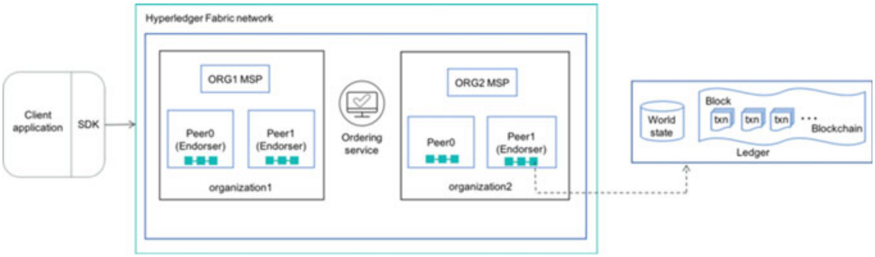


Fig. 1 Flow of operations/compositions in our proposed framework

The components and composition of HLF is shown in Fig. 1.

3.3 *InterPlanetary File System*

The InterPlanetary File System (IPFS) is a distributed file system protocol and peer-to-peer network. Content-addressing is used by IPFS to uniquely identify each file in a global namespace that connects all computing devices. In a similar way to BitTorrent, IPFS allows users to host and receive content. Rather than relying on a single server, IPFS is based on a decentralised system of user-operators who each hold a fraction of the overall data, resulting in a robust file storage and sharing system.

IPFS aims to establish a distributed and permanent web. This is accomplished by employing a content-addressed approach rather than HTTP's location-based mechanism. Instead of employing a physical address, IPFS addresses the content using a representation of the content itself. A file is broken into smaller bits, cryptographically hashed, and given a unique fingerprint called a content identifier (CID) when it is added to IPFS. This CID serves as a permanent record of the file at the moment it was created. A user can obtain this "beginning point" of data instead of talking to a server. When other nodes search up your file, they inquire as to which of their peer nodes is storing the content indicated by the CID.

IPFS uses a distributed hash table, or DHT, to store data. Data is transmitted between nodes in the network using BitTorrent-like processes. On the IPFS web, a user looking for material discovers neighbours who have access to that content. They then download little portions of the content from individuals who are close by. IPFS utilises a Merkle tree in addition to DHT and BitTorrent protocols. This is a data format that is akin to Git's version management system and the bitcoin blockchain technology. It is used to maintain source code versions in Git, but it is also used to track material across the Internet in IPFS.

IPFS and blockchains can operate well together due to their structural similarities. In fact, IPFS creator Juan Benet describes this as a "wonderful marriage." IPFS is one of a few initiatives that are part of Protocol Labs, an organisation that was formed by Benet as well [11, 12].

3.4 Secret Sharing Scheme

In secret sharing scheme, the secret is distributed among n users in such a way that, the generated share reveals zero information without all the random values. If all necessary and sufficient conditions are met, then only recovery of the original secret is possible [13–16].

4 Proposed System

4.1 System Conceptual Design

A user would be the one interacting with the system. Users could be patients, doctors, or an administrator. An administrator would be the one assigning roles to the other users and performing the backend operations. The basic tasks of the users would be to perform operations like create, update, query, and delete medical records. All the users would interact with the system with the help of a GUI (Fig. 2).

4.2 Workflow

4.2.1 Proposed Application

All the users of the system would be registered in the system by the admin and their respective IDs would be generated. Every user would also have their respective

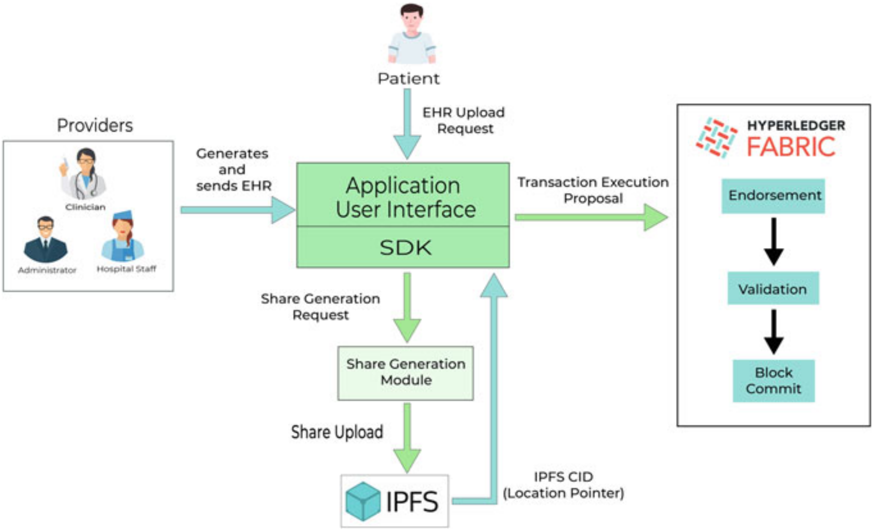


Fig. 2 System design

profiles which they would access through their login credentials. This login module would provide the outermost layer of security.

After successful login, a user can perform any CRUD operation to the network through the web UI. Let us say, a patient wants to view his details. He/she would make a request through the web page. This transaction would invoke the respective smart contract and make a transaction proposal to the fabric network. The Fabric SDK provides various APIs to interact with the network. The SDK will try to establish a connection to the network using the respective client identity. The client identity must be registered and enrolled in the network and recognised as a valid identity. After successful connection and invoking the respective smart contract, the network will return the response back to the SDK. The SDK will then parse the response accordingly and display it to the client through the webpage. In our proposed system, the large files like medical scan reports (image files) are not directly stored in the ledger, stored outside the blockchain network, known as off-chain storage. Here, we used another decentralised technology, the InterPlanetary File System or IPFS.

IPFS is a distributed file system protocol and peer-to-peer network. In a global namespace that connects all computing devices, IPFS uses content-addressing to uniquely identify each file. Along with the medical data fields available for the patient, the doctor can also upload/update the medical image files associated with the medical examination of the concerned patient. Before uploading the file needs to be encrypted in some sort. So in our proposed system, we are using the concept of secret sharing to hide the raw information of the EHR. The system would send a request to the share generation module. A secret share of the EHR would be generated and the system would then upload it to IPFS. After successfully adding the file to the IPFS network, a content identifier or CID (hash string) of the file is returned by the IPFS network. This CID acts as the pointer to the file and is actually stored in the blockchain.

4.2.2 Image Secret Sharing

An image sharing scheme is applied on the patient data to be stored in the IPFS platform. Multiple images of patient data are concatenated in one single image. Suppose $I_\mathrm{C} = I_1||I_2 \ldots ||I_K||$ is the set of concatenated images of all medical examination reports of Patient 1 generated by Hospital A. We consider ‘A’ to be the admin (trusted party) of Hospital A who acts as dealer and combiner who would send the secret share generated to IPFS. The concatenated images I_C is of size $W \times H$. ‘A’ creates access structure G which consists of authorised members and also generates random matrices for the members of G to be used during generation and reconstruction of a share. The secret share construction mechanism is as follows:

Initialisation and user verification:

1. ‘A’ generates I_C.
2. ‘A’ constructs the access structures G, which is the authorised subset of members $G = P_i|3$ ‘P_i’ k, $k =$ maximum number of members.
3. For each I_C, ‘A’ generates random matrices for each of the members in G.
4. First, ‘A’ encrypts the random matrices with his private key, then sends them to the concerned members after encrypting it with their public keys.

Algorithm 1: Share Generation

1. ‘A’ computes the secret share using random matrices of members by XORing it with I_C.

$$I_{\text{Secret Share}} = I_\mathrm{C} \oplus R_1 \oplus R_2 \oplus \cdots \oplus R_K.$$

2. ‘A’ uploads the secret share, $I_{\text{Secret Share}}$ on the IPFS network.

Example Let $I_\mathrm{C} = I_1||I_2 \ldots ||I_K||$ be the set of concatenated images and we are considering two access structures G_1 and G_2.

$$\begin{aligned}
&G_1 = \text{P1, D1, PTP} \\
&G_2 = \text{PTP, D1, D2, (To be used as backup if patient is not} \\
&\qquad \text{in the state to share its random matrix)} \\
&\text{D1} = \text{Doctor 1, D2} = \text{Doctor 2, P1} = \text{Patient 1, PTP} = \text{Patient Trusted Party}
\end{aligned}$$

‘A’ generates random number matrices R_{D1}, R_{D2}, R_{P1}, R_{PTP} for G_1 and G_2 using random functions. Secret share is generated $I_{\text{Secret Share}}$ by XORing I_C with the random matrices.

$$I_{\text{Secret Share}} = I_{\text{C}} \oplus R_{\text{P1}} \oplus R_{\text{D1}} \oplus R_{\text{PTP}} \text{ (using } G_1\text{)}$$
$$I_{\text{Secret Share}} = I_{\text{C}} \oplus R_{\text{D1}} \oplus R_{\text{D2}} \oplus R_{\text{PTP}} \text{ (using } G_2\text{)}$$

Algorithm 2: Image Recovery
Considering that all the participants agrees on recovering of data I_{C} by any member from the access structure, the mechanism of reconstruction is as follows:

1. Combiner will download the secret share, $I_{\text{Secret Share}}$ from the IPFS public network.
2. Each member will share their random matrices (R_1, R_2, …, R_K) to reconstruct the image I_{C}.

$$I_{\text{C}} = I_{\text{Secret Share}} \oplus R_1 \oplus R_2 \cdots \oplus R_K$$

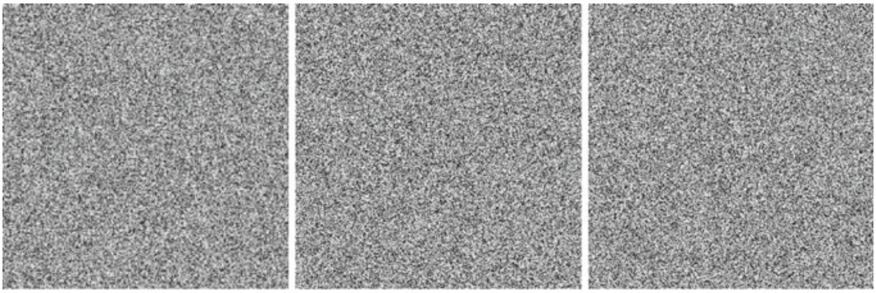
Example

$$\begin{aligned}
G_1 &= \text{P1, D1, PTP}\\
R &= R_{\text{P1}}, R_{\text{D1}}, \text{RPTP}\\
I_{\text{C}} &= I_{\text{Secret Share}} \oplus R_{\text{P1}} \oplus R_{\text{D1}} \oplus R_{\text{PTP}}\\
G_2 &= \text{PTP, D1, D2}\\
R &= \text{RPTP}, R_{\text{D1}}, R_{\text{D2}}\\
I_{\text{C}} &= I_{\text{Secret Share}} \oplus R_{\text{PTP}} \oplus R_{\text{D1}} \oplus R_{\text{D2}}
\end{aligned}$$

5 Discussion and Security Analysis

We discuss some of the features that our present framework proposes. For this, we have discussed some security measures with respect to the parameters (Figs. 3, 4 and 5):

- **Scalability**: The storing of large amounts of data on blockchain is a problem that demands a long-term solution. For storing patient's huge image files, our proposed model uses an off-chain storage mechanism. As a result, our methodology resolves the scalability issue.
- **Confidentiality**: In our architecture, we use a permissioned blockchain powered by the HLF network, which prevents access from any unauthorised third-party outside the network.
- **Integrity**: As we are leveraging tempar-proof blockchain technology, it is built specifically to protect data integrity.
- **Access Control**: Our suggested paradigm is patient-centric, which means that individuals have the power to grant access to their data to whoever they wish. Authorised parties are also permitted to see only a portion of the data, i.e. they are permitted to know only what they require.



(a) Patient Random Matrix (b) Doctor Random Matrix (c) PTP Random Matrix

Fig. 3 Random matrices

Fig. 4 I_{C}

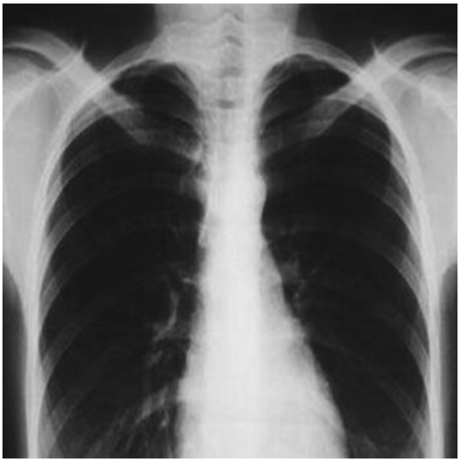
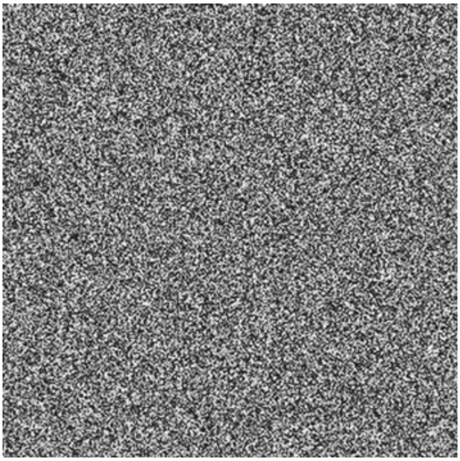


Fig. 5 $I_{\mathrm{Secret\ Share}}$



6 Conclusions and Future Scope

Our proposed architecture is a patient-centric framework that provides a secure, decentralised, and distributed EHR administration system that employs the IPFS method for off-chain data storage. The IPFS network ensures that a hash named CID is generated. As a result, the cryptographic hash generated for each file stored in IPFS protects its security. We intend to update the system in future by adding features such as verifiability to secret sharing method.

References

1. Shahnaz A, Qamar U, Khalid A (2019) Using blockchain for electronic health records. IEEE Access 7:147782–147795. https://doi.org/10.1109/ACCESS.2019.2946373
2. Tith D et al (2020) Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability. Healthc Inf Res 26(1)
3. Biswas S et al (2020) Blockchain for E-health-care systems: easier said than done. Computer 53(7):57–67
4. Aswin AV, Basil KY, Viswan VP, Reji B, Kuriakose B (2020) Design of AYUSH: a blockchain-based health record management system. In: Ranganathan G, Chen J, Rocha Á (eds) Inventive communication and computational technologies. Lecture notes in networks and systems, vol 89. Springer, Singapore. https://doi.org/10.1007/978-981-15-0146-3_62
5. Zhang H et al (2018) Cloud storage for electronic health records based on secret sharing with verifiable reconstruction outsourcing. IEEE Access 6:40713–40722
6. Kumar R, Marchang N, Tripathi R (2020) Distributed off-chain storage of patient diagnostic reports in healthcare system using IPFS and blockchain. In: 2020 International conference on communication systems networks (COMSNETS). IEEE
7. Sultana M et al (2020) Towards developing a secure medical image sharing system based on zero trust principles and blockchain technology. BMC Med Inf Decis Making 20(1):1–10
8. Tanwar S, Parekh K, Evans R (2020) Blockchain-based electronic healthcare record system for healthcare 4.0 applications. J Inf Secur Appl 50:102407
9. Aswin AV et al (2020) Design of AYUSH: a blockchain-based health record management system. In: Inventive communication and computational technologies. Springer, Singapore, pp 665–672
10. Choi Y-J, Kim K-J (2020) Secure healthcare data management and sharing platform based on hyperledger fabric. J Internet Comput Serv 21(1):95–102
11. Jeong J et al (2020) Design and implementation of a digital evidence management model based on hyperledger fabric. J Inf Process Syst 16(4)
12. Mukne H et al (2019) Land record management using hyperledger fabric and IPFS. In: 2019 10th International conference on computing, communication and networking technologies (ICCCNT). IEEE
13. Shamir A. How to share a secret. http://web.mit.edu/6.857/OldStuff/Fall03/ref/Shamir-HowToShareASecret.pdf
14. Ulutas M, Ulutas G, Nabiyev VV. Medical image security and EPR hiding using Shamir's secret sharing scheme. https://www.sciencedirect.com/science/article/pii/S0164121210003274
15. Chattopadhyay AK, Nag A, Singh JP et al (2020) A verifiable multi-secret image sharing scheme using XOR operation and hash function. Multimed Tools Appl. https://doi.org/10.1007/s11042-020-09174-0
16. Nag A, Singh JP, Singh AK (2020) An efficient Boolean based multi-secret image sharing scheme. Multimed Tools Appl 79:16219–16243. https://doi.org/10.1007/s11042-019-07807-7