International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Development of Fragmentation to Create a Big Data Security Framework using Blockchain Innovation

Dr. Shrihari M. R.[1], Ajay N.[2], Shwetha B. V.[3], Mahesh M. R.[4], Dr. Vikas Reddy S[5], Dr. Mohan H. S. [6], Muniraju M.[7]

Submitted: 27/08/2023 **Revised**: 15/10/2023 **Accepted**: 28/10/2023

Abstract: The volume and diversity of data created by both humans and robots is increasing. Although the exponential growth, the tools and technology implemented to control huge information volume are not designed to fulfill protection and data security standards. Furthermore, the majority of Current big data security systems are provided by a centralised third party, making them subject to a wide variety of security risks. This reading provides a novel large information security solution driven by Blockchain technology that integrates fragmentation, encryption, and access control mechanisms. Our suggested fragmentation method considers the data owner's request for encryption to be included in the fragmentation progression. Moreover, data pieces will be stored in the dispersed way provided by the big data atmosphere, adding an extra layer of data security. We want to improve large data security with suitable expense and eliminate encryption overhead for non-sensitive and low-sensitive data sections in classify to provide an ideal security clarification. We show the consequences of our developed approaches to demonstrate that the overheads (in terms of computing time) caused by our system are insignificant in comparison to the security and confidentiality improvements it provides.

Keywords: *Blockchain, Big data security, Auditing, Fragmentation, Access control*

1. Introduction

In today's world's data volume is rapidly increasing. According to the International Data Corporation (IDC) , worldwide data creation is predicted to exceed 180 zettabytes by 2025. Data is gathered and evaluated on a regular basis in the big data environment. Companies and organizations typically utilise the data gathered to customize services, enhance decision-making optimization, forecast future trends, and so on.

Data is becoming a critical aspect in the industry when massive data is kept, it is typically done so in a distributed le system or the cloud. To complete a certain task in distributed storage, numerous nodes must collaborate. As a result, targeting one or more nodes will have an impact on the dependability of computation results. The burden of security on storage nodes grows considerably with distributed data storage. If data encryption storage is used, the key. As a result, in large it is difficult to employ typical asymmetric data systems. As well as symmetric encryption schemes. In the Because of the enormous size volume, cloud storage is significant increase, and quick modifications, immediately encrypting large datasets will increase the danger of secret key organization and need a significant computing transparency [1]. Once the truth was revealed when a key is exposed, the entire dataset might be damaged. Blockchain technology has piqued the interest of academics and entrepreneurs alike. Because of various qualities like as impermanence, privacy, confidence, and authorization It is based on private and public key infrastructure (PKI), cryptographic, peer-to-peer networks, and consensus methods that guarantee the ledger's data integrity. There are various advantages to combining Blockchain with big data. Automation offered by Blockchain workflow, for example, will help firms become more effective and productive. Both technologies are important to the company, but Blockchain is still in its early phases of adoption in data science, thus more work is required. Any effective security system must have the essential data security: secrecy (preventing unauthorised and violation), accessibility (ensuring authorized access have access to resources in both normal and crisis conditions), and authenticity[2].

Data can be secured in a variety of ways, including: 1) secure access control methods 2) keeping secrets (through encryption, scrambling and steganography), and 3) making it understandable and accessible only to those with a map or key to effectively invert transformation and recover the original form. Data fragmentation for data security novel idea, and it has been used in several studies. Moreover, fragmentation is employed in combination with encryption

[1]*Associate Professor, Dept of CSE*,
SJC Institute of Technology, VTU, Karnataka, India,
[2]*Assistant Professor, Dept of CSE*,
SJC Institute of Technology, VTU, Karnataka, India.
[3]*Assistant Professor, Dept of CSD, SJC Institute of Technology, VTU, Karnataka, India., VTU, Karnataka, India*
[4]*Assistant Professor, Dept of ECE Nagarjuna College of Engineering and Technology, VTU, Karnataka, India*
[5]*Associate Professor, Dept of AI&ML*,
SJC Institute of Technology, VTU, Karnataka, India,
[6] *Professor, RNS Institute of Technology, VTU, Karnataka, India*
[7]*Assistant Professor, Dept of AI&ML*,
SJC Institute of Technology, VTU, Karnataka, India.
** Corresponding Author Email: shrihari.mr@gmail.com*

to allow for the simultaneous encryption of data fragments, which is more effective than conducting whole-sequence encryption. In order to eliminate considerable cost, recent research has recommended using fragmentation to substitute encryption for data security. Encryption, on the other hand, should be a foundation of any data security policy. [1].

2. Background

2.1 Technology in Big Data

In essence, big data could be viewed as the answer to the question of how challenging it is to deal with massive datasets that are organised, semi-structured, or unstructured, and challenging to handle using traditional approaches. Techniques of data analysis is alternative definition comes from "what big data is," and it focuses on a computational framework for acquiring and analysing key information from massive and varied databases. These descriptions reflect the primary aspects of huge data sets. also known as v Variety, velocity, veracity, and value are referred to as the "5Vs[3].

2.2 Security Issues In Big Data

A Big data raises significant security and privacy concerns. Its qualities posed several complications. particularly with the proliferation of IoT strategy that generate enormous amounts of data a large volume of data that must be processed in order to commence useful data with new On-demand information and technology services are examples of such services. As a result, public cloud service providers have vast quantities of power and storage accessible. Data protection has become a critical need for any large data endeavour. According to this article, conventional security techniques, such as firewalls and other perimeter security measures, are useless in a massive data environment. A overview of some large data security issues [5] follows:

Keeping attackers from producing and uploading malicious material data. Mechanisms for identity management, such as granular data access (giving different users differing levels of access), are no longer useful with large amounts of data.

The distribution of massive volumes of data storage and processing over several workstations creates several security issues. For example, discovering an attack might take a long time.

A Security auditing is a key technical examination for every organisation. Although, it is uncommon to see it used with large data systems. As a result, data management becomes difficult.

2.3 Technology in Blockchain

Large-scale data dispensation and security are unquestionably difficult tasks. Blockchain can be seen as the perfect answer to many problems with massive data management and analytics. Decentralization, integrity, and immutability are all crucial characteristics offered by Blockchain.

Blockchain: Blockchain [3] is a distributed database of documented transactions or events stored in blocks and linked by cryptographic techniques. Only producers, who are powerful processing nodes capable of validating transaction records and recording information on the ledger, may change the documents, which are shared and watched by all network nodes. Each block on the Blockchain is linked to the one before it and contains time-stamped transactions.

Blockchain Network: The foundational technology that A peer-to-peer network is what the blockchain is. every student in the network abides by definite peering procedure guidelines, operation ledger management, consent procedure, and processing. There are three varieties of blockchain: consortium, private, and public. Blockchains has different blockchain types can be distinguished depending on on a network node's degree of permission in words of accessing the blockchain data and acting as a validator. A public, permission less blockchain enables anonymous transactions. people to collaborate using their assets, such bitcoin Ethereum is also. Private blockchains are open and centralised. to specific people or organisations, but consortium blockchains are available to certain groups and only partially centralised. Specifically, private and consortium blockchains are regulated blockchains[4].

3. Related Work

Numerous studies have examined the use of blockchain technology in the access to healthcare data, which mandates that individuals own access their medical records, etc. Blockchain holds promise. to enable safe exchange of electronic health records (EHR) in here the true owners are the patients. Authors in [13] has been suggested that the blockchain only store relevant to health and medical events. Thus, greatly avoiding scalable blockchain technology to store all health data records. Social media user activity results in a large-scale databases. Depending on outside parties carries a significant risk. to safeguard sensitive and personal information. Therefore, it is necessary to allow people to monitor and manage their internet activity Blockchain can be applied in social media as a permission-filtering strategy[20].

A social networking platform built on top of the blockchain that gives users access credentials to manage their data. In order to supervise the user's circles and encrypt information before transmit; a PCA is issued for every position. Blockchain is used to store the user share transaction data. Although share, a customer - centric blockchain relevance, might provide users organize over

their information, implementing their explanation would be complex due to the difficulty of creating a large volume of transactions to be kept in the blockchain and encrypting the whole substance of users' data. The effort in [13] attempts to present a blockchain infrastructure for transactions including smart mobility. The goal is to safeguard and protect the individual data that has been gathered for various participants, government including businesses and institutions.

In this circumstances, each contestant sends encrypted information to a blockchain system that has a elegant agreement with business rules. This approach is only applicable to graceful mobility development.

Guan Z. et al. suggested a Blockchain and Trusted Security Model-based big data collecting and trading system (TSM). They merged much technology such as Substantial Unclonable Function (SUF), which employs the fingerprint as a projection identifier and the TSM model, which assures the trustworthiness of the data collecting development. They want to use blockchain to ease accounting and trading operations, which will allow them to trade confidential data and mitigate any assault. Trust is a critical concern in edge compute, particularly along with edge strategy that exchange massive data. However, their approach falls short when it comes to end-to-end data protection.

Xu C. et al. propose a blockchain architecture in mutual boundaries to consent to trusted massive information exchange though retaining competent reserve consumption in circumference strategy. Furthermore, to attain optimum processing power, this work introduced a consent technique dubbed Proof-of-Collaboration. Smart environments generate a tremendous quantity of data, which is often individual and insightful statistics that requires special care to safeguard and preserve. The study in offered a system for decentralised big data assessment in elegant cities based on blockchain to promote permanence and consistency to contribute in smart city building to prevent big data in auditing and centralization, which the third celebration routinely delivers..

The authors of presented a decentralised big data inspection strategy for smart cities along on blockchain in their purpose is to improve the stability and dependability of auditing schemes by eliminating the require for centralised third-party assessment, except the cost to consumers is significant. During the auditing process, their technique incurs additional expenses in blockchain ledger navigation. Similar work has been published in, however their approach was developed for cloud storage. Furthermore, they advocated storing information honesty authentication tags on blockchain to decrease the cost of announcement and estimation generated by the reliability authentication development.

4. Proposed Framework

This section will go over the proposed methodology. The scheme mechanism will be established first. Then comes the flow determination be explained. at last, we shall share our potential findings.

4.1 Framework of the Architecture

Information will be processed in our framework to determine sensitive parts. The sensitivity level is determined by the user's choices will be decided. The amount of compassion, information will be handled in a different way of the information will be fragmented. Prior to being placed in large data distributed storage, it is encrypted. The metadata (MD) produced during fragmentation The procedure and the authorization list will be stored on the permissioned blockchain for investigate and interfere resistance capabilities[11].

4.1.1 Data Sensitivity Detector(DSD)

There are three types of sensitivity detection techniques: automatic, semi-automated, and manual. Because of the high number of information, manual data compassion recognition necessitates significant work. As a result, automated solutions are advised. These solutions include complicated procedures that are beyond the scope of this research and will be addressed in future work. Province expert and neural system are two examples. In our scenario, sensitivity detection is subject to the policies and needs of the data owner (DO). The DO defines the data sensitivity level (high, low, or none) and identifies the sensitive properties that must be defended and Fig 3 depicts the run of DSD[7].

4.1.2 Data Splitter(DS)

Our approach makes use of fragmentation method to offer an further layer of security to sensitive data. Based on the needs of the user, data is separated into responsive and non-sensitive assortment. The confirmation is employed to information reliability by computing the SHA-512 for the innovative and evaluate the hashing consequence to the result after the renovation procedure. The system secures sensitive data based on its level of compassion. We employ scramble to strengthen the fragmentation development for low-sensitive information, which is paired with disseminated large information storage partition [6]. Moreover, to reduce the significant expense involved with encrypting the whole amount of data, our method encrypts solitary the big- sensitive section of the information. Subsection B contains more information on the suggested fragmentation strategies(see Fig1).

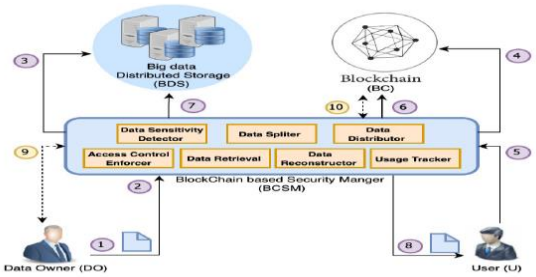

Fig 1. The planned framework architecture.

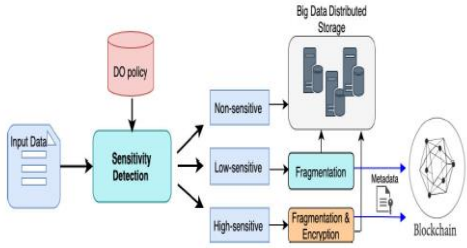

Fig 2. The flow of information compassion detector

4.1.3 Data Distributor(DD)

The module allocates and transfers the data-id to the combined. The MD is then created and encrypted by DD. The number of pieces determines the size of the MD information structure. This component also transmits the combined les to BDS. It also propel MD and PL to the blockchain register for storage.

4.1.4 Data Retrieval(DR)

To use the data-id, this module receives the data associated to the data as of the blockchain, and the request compound les are recover from the BDS. The metadata is subsequently decrypted by the DR before being sent to the Data Reconstructor.

4.1.5 Data Reconstructor(DRE)

To use the metadata acquired from the blockchain, this constituent restores the information to its innovative appearance. It reconstructs the original le using decryption and defragmentation algorithms.

4.1.6 Access Control Enforcer(ACE)

The module is in charge of the data owner's and user's authentication and permission. The client requests authentication from the ACE and once granted, the endorsement procedure is started. ACE uses multi-factor authorization to validate the user's identity. The ACL rules are used to ensure that only the privileges mentioned in PL are used to access the data. According to the PL, the demand information can only be accessed by a limited number of authorised persons. If the user is established, the ACE logs the demand on the blockchain for future reference[7].

4.1.7 Usage Tracker (UT)

The constituent receives inspection in sequence on information right to use and convention from the blockchain winning request from the data vendor or examiner, exploiting the blockchain's traceability capability.

The connections between these mechanisms are depicted in Fig 3 which depicts dynamic mechanism through the uploading procedure. Fig 4 depicts the dynamic mechanism through the analysis procedure. The technique for retrieving and uploading information will be outlined below. Data storage procedure begins with determining the sensitivity of the data, as shown in Fig 3.The data is then classified as elevated/low sensitive or nonsensitive. Nonsensitive information will be saved in BDS, while sensitive information will be transferred to the information splitter, which determination fragment the information and generate the combined file. In the case of very responsive material, Information fragments will be encrypted prior to the combined data files are created.

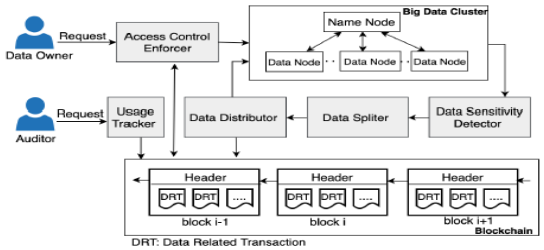

Fig3. The dynamic mechanism of the framework through uploading.

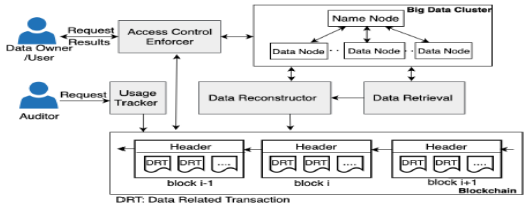

Fig 4. The dynamic components of the framework through reading

Furthermore, the MD and PL will be produced for storage and management by the blockchain.at last, the combined les will be separation by on chunk size and kept in BDS, adding another degree of security. Each data-related transaction, such as read and write events, will be record as a DRT (Data Related Transaction) on the blockchain ledger. Fig 5 depicts this procedure. Data Retrieval: To access data on the blockchain, the users have to be authentic and authorised by inspection the PL le linked with the desired data. Following authorisation, the manager retrieves the required data's details with the metadata acquired as of the blockchain[5], as illustrated in Fig 4.

4.2 Methodologies of Fragmentation Proposed

The fragmentation approaches attempt to improve the security of two categories of insightful data: low insightful data and high sensitive data. The degree of sensitivity is decided by the user desire for processing his or her data. Each section describes our fragmentation strategies, which include fragmentation, defragmentation and encryption. and defragmentation with decryption. The primary two procedures are used for low - sensitivity information while the third is used for high-sensitivity data.

4.2.1Fragmentation

Fragmentation allows for competent move quickly and encryption algorithms that take advantage of parallelism. It is critical to evaluate fragment size; a too-large segment may give too much data, but a confined part might produce excruciating transparency. To accomplish optimum fragmentation, the following guidelines must be followed:

Completeness: During the fragmentation process, no data should be lost. Every information item have to be represented by at slightest one fragment.

Reconstruction: If the information is broken down, it can still be united exclusive of modifying the data structure.

Disjointness: To avoid data duplication, information in the fragment must not be incorporated in the additional fragments.

The checksum is used to confirm that the aforementioned conditions are followed by compute the SHA-512 for the novel organizer and comparing the hashing result to the result of the file subsequent to the renovation development. Our system secures sensitive data based on its level of sensitivity. We employ scrambling to strengthen the fragmentation development for low-sensitive information, which is paired with disseminated large data storage segmentation[11].

To reduce the significant cost associated with encrypting massive amounts of data, we use an approach that encrypts just the most sensitive parts of a dataset. In any case, security requires fragmentation. Enhances existing system latency As a result, in order to decent fragmentation method, acceptable overall performance must be linked with parallelization of processing

The dataset's information and its fragmentation are store in the fragmentation. We plan a novel metadata data structure. stored in a blockchain ledger that takes use of blockchain immutability With tamper resistance to help with data integrity checks. The hashing algorithm is used to compute the data-hash.(SHA-512) as: message digest (md) = H (m).

The message digest (md) is computed and included in the MD. The pseudo-code for fragmentation is shown in Algorithm 1. There are two primary actions in this algorithm. To begin, divide the sensitive data le into fragments of predetermined size sufficient for the original le size. By saving the ID of each fragment, the original order of the pieces will be preserved in the Mapping Array.

4.2.2.1Defragmentation

Defragmentation is the opposite of fragmentation. It adheres to the same fragmentation rules. Algorithm 2 describes our defragmentation pseudo-code. The procedure begins by retrieving the Mapping Array encoded in the MD from the blockchain. The next step is to look for the fID in M in order to recreate the original le in the order obtained in Mapping Array. Decryption is required to create the original version of high-sensitive pieces.

Algorithm 1. Fragmentation

```
Input SenstvDataFile; fragmentize; File Path; IsHighSensitive
Output M:{ m1;m2;m3; : : :mn},MappingArray [ ]
procedure Fragmentation
SenstvDataFile  FileInputStream(FilePath)
F  SenstvDataFile:SplitFile(fragmentSize)
for i  0; fragmentsNumber do
MappingArray[i] Dfi:id
if IsHighSensitive then
for i  0; fragmentsNumber=2 do
mi D SenstvDataFile:merg(en(fj); en(fk ))
F fj; fk are selected arbitrarily
end for
mergeFileNumber D i
else
for i  0; fragmentsNumber=2 do
mi D SenstvDataFile:merg(fj; fk )
F fj; fk are selected arbitrarily
 end for
mergeFileNumber D i
end if
procedure Store
for i  0; mergeFileNumber do
insertHDFS(mi)
end for
send MappingArray [ ] to blockchain
```

Algorithm 2. Defragmentation

```
Input M:{ m1;m2;m3; :mn}; IsHighSensitive;
MappingArray [ ]
Output SensitiveDataFile
procedure Defragmentation
 get MappingArray[ ] from blockchain
Outputstream=FileOutputStream(SensitiveDataFile)
 for i  fragmentsNumber; 0 do
fID D MappingArray[i]
Search fID in M if match, read fragment and send
 to Outputstream
f D FileInputStream(fID)
f :read ()
if IsHighSensitive then
 outputstrem:write(decrypt(f ))
else
 outputstrem:write(f )
 end for
outputstream:close()
```

5. Results and Discussion

The research and findings resolve be explained in depth in this section.

5.1 Environmental Design

We used a effective Hadoop cluster comprised of one Name node and three data nodes to assess the suggested strategies in the BMBD framework. Furthermore conducted the trials in a well prohibited effective network to remove the variable system characteristics. A virtual machine represents a real node, and all effective technology run on the equivalent host mechanism. The congregation computer has an Intel Core i7 2.3 GHz processor, 16GB DDR6 memory, and a 1TB SSD hard disc. Virtual Box 6.1.26 is the virtual machine manager. The hardware configuration of the Name node is 8 CPU and 11.5 GB of RAM. We utilised 1 CPU and 1024MB for data nodes. Each host computer is running Ubuntu 18.04.5 LTS and Hadoop 3.3.0.

5.2 Data Used

To assess the presentation of the planned system, we employed structured datasets. Synthea [14], a artificial enduring information producer, was utilised. We utilised data les of various sizes (64MB, 128MB, 256MB, 512MB, and 1024MB). We created a Java model of our structure to interface through the Hadoop cluster.

5.2.1 Experiment Profile

We created a Java prototype of our system that can read and write to HDFS. To assess the efficacy of our approach, we evaluated different structured les sizes. We implemented multithreading to imitate fragmentation parallelism. Every research was performed four stage, and the standard was used as the intentional result.

Experiments in inscription and interpretation We calculated the speed in gigabytes per millisecond by measuring the calculation time of transferring a disc from the confined File system to HDFS. Furthermore, we compute the inscription and interpretation rates in gigabytes per millisecond and approximated the processing time of transferring a chunk as of HDFS to the neighboring file system. Five controlled statistics files of varying sizes are used in the studies. The lowest data extent was the equivalent as the HDFS block size of 64 MB. The bare minimum is two merged files. As a consequence, we chose two fragment sizes based on the data size. On behalf of data sizes of 64 MB and 128 MB, we used the same fragment size, resulting in two and four merged files, respectively. However, for 256 MB, 512 MB, and 1024 MB, we chose 64 MB as the fragment size.

5.2.2 Comparison Of The Proposed Structure With Standard Hdfs

To assess the transparency, we compare the planned architecture to the common HDFS. The concert overhead of the inscription trials as speed disparities between the generic HDFS and our implementations. The overhead incurred by the suggested fragmentation technique on a dataset.

The overhead depicted in this Fig 5 and Fig 5 is the instance required by the suggested fragmentation technique on a dataset range in size from (64MB to 1024 MB). This Figures has five points, each of which reflects the outcomes of data size. It also displays the transparency caused by the fragmentation method when functional to low-sensitivity information.

The transparency required by the fragmentation and encryption approach employed on high -insightful information illustrate that low-sensitive data fragmentation has a smaller overhead than high- insightful data fragmentation. Consequently, encrypting just the most sensitive data avoids the enormous cost compulsory by encrypting the whole dataset. Equally overhead of the suggested defragmentation technique on various data volumes is depicted. It demonstrates the burden compulsory on low - insightful information by a defragmentation process. The time required by the suggested defragmentation technique on the identical datasets used in preceding instances. This diagram depicts the transparency placed on highly sensitive data by defragmentation and decryption algorithm. The expense posed by low-insightful information defragmentation is smaller than that posed by high-sensitive data defragmentation and decryption. Tables 1 and 1 compare the calculation time of our suggested approaches to HDFS while writing and reading various data sizes. Table 1 shows a comparison between Generic HDFS, fragmentation, fragmentation, and encryption. Table 2 shows a assessment between frequent HDFS, defragmentation, defragmentation and decryption[10].

Table 1. Write Data research results

Data Size(mb)	HDFS (ms)	Fragmentation (ms)	Fragmentatin(ms) &Encryption(ms)
64	963	1022	1205
128	28925	29038	29225
256	34670	34886	35310
512	87142	87656	87955
1024	21112	212068	212502

Table 2. Read Data experiment results

Data Size(mb)	HDFS (ms)	Defragmentation (ms)	Defragmentatin(ms) & Decryption (ms)
64	4152	4180	4425
128	8556	8646	9032
256	14381	14482	15167
512	26184	26950	27266
1024	52580	53984	54551

5.3 Discussion on the Writing Experimentations

Table 1 summarizes the writing speeds in all experiments. As a starting point, believe the consequences of the general HDFS. In all writing trials, the overhead of fragmentation is smaller than the overhead of fragmentation and encryption. In order to analyse the findings, we intended the transparency ratio foundation by fragmentation and fragmentation & encryption in the writing overhead ratio.[8] For a 64MB data size, the transparency ratio was 1.0612 for fragmentation and 1.2510 for fragmentation and encryption, correspondingly.(see Fig5)

5.4 Debate on the Interpretation Research

Table 2 summarises the evaluation speeds in all testing. The findings of the generic HDFS were used as a benchmark. In all reading trials, defragmentation alone outperforms defragmentation paired with decryption. The overhead grows in proportion to the size of the tested data in each trial. As a consequence, considerable transparency will be circumvent based on user choices for disabling encryption when it is not required while preserving an adequate protection mechanism for low- insightful information[9]. We measured the transparency ratio foundation by defragmentation and defragmentation & decryption approaches in the reading trials when compared to the default HDFS. (see Fig6).

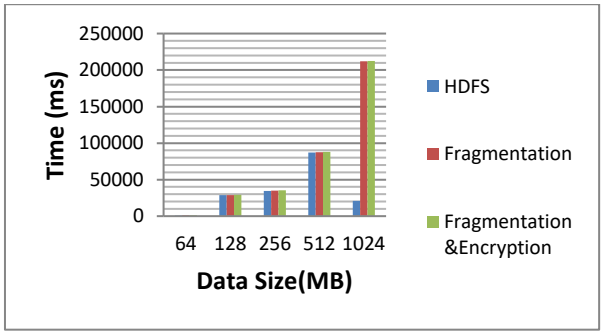

Fig 5. Performance assessment of proposed scheme with HDFS for dissimilar data size with Fragmentation & encryption overhead

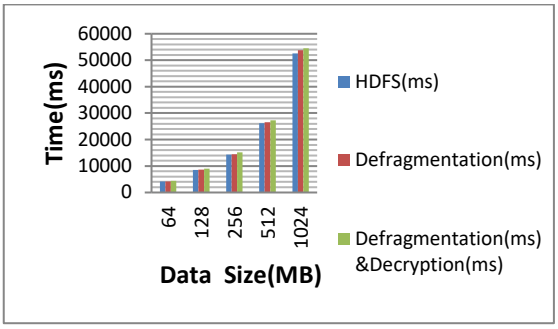

Fig 6. Performance assessment of proposed scheme with HDFS for different data size with Defragmentation & Decryption overhead

5. Conclusion and Future Work

Big data protection and confidentiality concerns are significant and must be addressed. For example, large data models like Hadoop are constructed with no security assumptions. Furthermore, the majority of available technologies rely on third parties, which raises important protection concerns. In this paper, we planned a big data security structure that uses blockchain technology and fragmentation to provide data security. The framework provides a safe environment for exchanging, storing, and sending large amounts of data. Blockchain is in charge of ensuring the security of huge data storage and retrieval procedures, as well as access control and auditing methods. Prior research has not fully addressed the security of massive data. As an example, Previous research has mostly concentrated on access control. Distribution and auditing of data in certain big data areas Smart homes and healthcare are two examples. However, our suggested A structure is a general resolution so as to may be used in a variety of situations. a variety of large data areas This research is still ongoing, and The authors feel that further information should be gathered. This will be published in a prospect work we are preparing for future projects to put our framework's entire scenario into action using blockchain technology. The Hyperledger framework The solution will be implemented using a platform. increased transaction throughput permissioned blockchain as well as more security than other blockchain platforms

References

[1] Kassen M. ``Blockchain and E-government innovation: Automation of public information processes," Inf. Syst., vol. 103, Jan. 2022,Art. no. 101862.

[2] Manzoor A. A. Braeken, S. S. Kanhere, M. Ylianttila, and M. Liyanage,``Proxy re-encryption enabled secure and anonymous IoT data sharing

platform based on blockchain," J. Netw. Comput. Appl., vol. 176,Feb. 2021, Art. no. 102917.

[3] Deepa. N., Q. Pham, D. C. Nguyen, S. Bhattacharya, P. B. T. R. Gadekallu,P. K. R. Maddikunta, F. Fang, and P. N. Pathirana, ``A survey on blockchain for big data: Approaches, opportunities, and future directions," CoRR,vol. abs/2009.00858, pp.1-4, Oct. 2020.

[4] Xiao Y., N. Zhang, W. Lou, and Y. T. Hou,``A survey of distributed consensus protocols for blockchain networks," IEEE Commun. Surveys Tuts., vol. 22, no. 2, pp. 1432-1465, 2nd Quart., 2020.

[5] J.Wu,G.Jiang J. Li, and T.Srikanthan, "`Blockchain-based public auditing for big data in cloud storage," Inf. Process. Manage.,vol. 57, no. 6, Nov. 2020, Art. no. 102382.

[6] N Ajay; H S Mohan; B V Shwetha; M R Shrihari," Access Control Framework in the Cloud based on Multi-Blockchain with Light Privacy Protection" ICDCECE, 23-24 April 2022.

[7] Z. Yang H. Yu, , and R. O. Sinnott, ``Decentralized big data auditing for smart city environments leveraging blockchain technology," IEEE Access, vol. 7, pp. 6288-6296, 2019.

[8] M R Shrihari, Manjunath T.N, R.A. Archana and Hegadi, Ravindra S," Development of Security Clustering Process for Big Data in Cloud", ICT Analysis and Applications ,Lecture Notes in Networks and Systems book series (LNNS,volume 154) 16 December 2020..

[9] Yang and Z. Liang Z., ``Automated identication of sensitive data from implicit user specication," Cybersecurity, vol. 1, no. 1, pp. 1-15, Dec. 2018.

[10] Lv, S. Zhu D., H. Xu, and R. Liu, ``Areview of big data security and privacy protection technology," in Proc. IEEE 18th Int. Conf. Commun. Technol. (ICCT), Oct. 2018, pp. 1082-1091.

[11] Santos and N. G. L. Masala, ``Big data security on cloud servers using data fragmentation technique and NoSQL database," in *Proc. Int. Conf. Intell. Interact. Multimedia Syst. Services*. Cham, Switzerland: Springer, 2018,pp. 5_13.

[12] Tankard C., ``Encryption as the cornerstone of big data security," Netw. Secur., vol. 2017, no. 3, pp. 57, Mar. 2017.

[13] M R Shrihari , Manjunath T.N, R.A. Archana and Hegadi, Ravindra S," Research Challenges in Big Data Security with Hadoop Platform" International Conference on (RTIP2R December 21-22 2018, RTIP2R 2018, CCIS 1037, pp. 550–560, 2019.

[14] M.R. Shrihari, R.A. Archana, T.N. Manjunath and Ravindra S. Hegadi," A Review on Different Methods to Protect Big Data Sets", 2018,issue-12&page-4.

[15] Chakravorty A. and C. Rong, ``Ushare: User controlled social media based on blockchain," in Proc. 11th Int. Conf. Ubiquitous Inf. Manage. Commun., Jan. 2017, pp. 1_6.