Optimized Secure Federated Learning for Event Detection in Big Data using Blockchain Mechanism

K. Prasanna Lakshmi[1], K.Swapnika[2]
Professor, Information Technology Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India[1]
Ph.D Scholar, JNTUH, Hyderabad, India[2]

***Abstract*—Currently, cloud storage in blockchain and federated learning technology provides better security among data transmission and file access. But, in some of the cases, security issues arose. So, to avoid security problems and offer better protection in a cloud environment, a novel optimized buffalo-based Homomorphic SHA blockchain (OBHSB). In this model for accessing the cloud storage data with the key matching method, if any of the unauthenticated users are trying to access the file initially, the system checks the key matching parameter. The proposed model was developed to provide better security in big data presented in the cloud environment. However, the parameters in the proposed model were compared with the existing models to make sure better performance was attained through the proposed model. Attack was considered as an event in this research. In the performance analysis, the performance rate of the proposed model was validated. Subsequently, the case study was developed in this research to explain the working procedure of the proposed design; model performs hashing, encryption, decryption, and key matching mechanisms. The results proposed model is observed to have 100% confidentiality rate after attack.**

Keywords*—Blockchain; cloud storage; decryption; encryption; federated learning; hashing; homomorphism

I. INTRODUCTION

Federated Learning is a Machine Learning (ML) concept that trains the program through a decentralized edge system having local data [1]. Fig. 1 describes the general architecture of federated learning. The training process does not involve the exchange of data [2]. In conventional ML approaches, the entire local datasets are transferred to a single server/ device, whereas in classical ML approaches, the local dataset is distributed identically [3]. In the federated learning approach [4], the local sample dataset is trained without distributing the information [5]. Therefore, it is widely used in applications to overcome data privacy and security issues [6]. Event detection (ED) involves the investigation of several events to provide a better understanding of social events [7].

The event investigation by analyzing massive heterogeneous datasets such as audio, images, and video is called Multimodal ED [8]. The development of different image processing approaches helps identify various types of events automatically [9]. Unstructured multimedia data are created in recent types to search the data more flexibly [10]. Many different technologies were developed to identify the event in various scenarios, such as ED in the smart city [11], ED in social media [12], ED in road traffic [13], etc. Recently, event identification has been carried out with the help of big data [14]. Big data is a massive collection of data whose size can be enhanced exponentially over time [15]. The conventional information management tool can store only a limited amount of data [16]. But big data can process and store massive amounts of data in it [17]. Moreover, event detection using big data is highly effective because of its high storage space [18]. However, multimodal event detection is one of the major concerns for machine learning approaches. In this research, the attack is considered as event. Here, we are performing homomorphism concept so single event detection has been performed. Thus, different machine learning-based multimodal ED such as message dissemination model with the assistance of blockchain [19], blade icing identification technique based on blockchain and imbalanced federated learning approach [20], incentive governor for FL approach based on blockchain technology [22] were developed for identifyingevents from audio, text, image and video. However, they cannot provide effective results. Hence, an optimized federated self-supervised learning for multimodal event classification in big data using blockchain is presented in this article. In this work, some of the recent literature related to this topic was discussed in the second Section. Moreover, the system model as well as the problem statement was presented in Section III, and in fourth section the proposed system model was discussed. Moreover, the result and a discussion of the proposed model were developed in Section V, and the paper was concluded with the conclusion in Section VI.

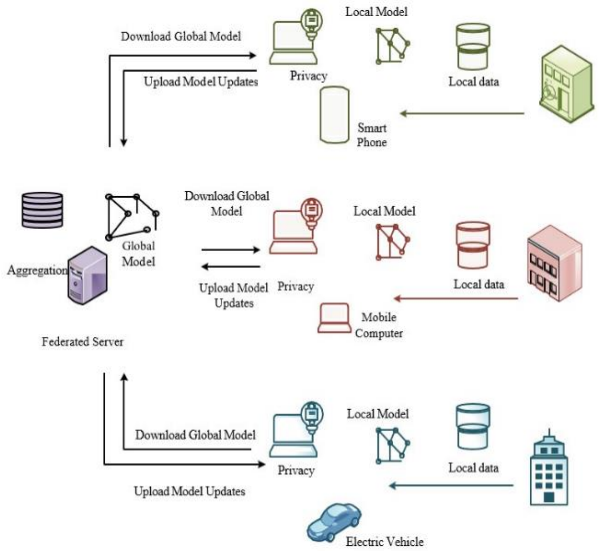

Fig. 1. General architecture of federated learning.

II. Related Work

The recent works associated with this research are summarized below.

Nowadays, message dissemination plays a significant role in providing safety to Electric Vehicles (EVS). The messages are transferred among EVs using broadcasting technology. Moreover, high mobility and density in incoming vehicles affect the message dissemination process. To overcome the issues in message dissemination, Ferheen Ayaz et al. [19] developed the message dissemination model with the assistance of blockchain. This technique minimizes the delay and improves the message delivery rate. However, the practical implementation of this model provides inefficient results. Recently, renewable energy resources such as solar, wind, etc., have been widely used in different sectors. Wind energy is one of the rapidly growing renewable sources among these renewable sources. Blade icing is one of the significant concerns in wind power generation. To overcome this issue, data-driven technologies are widely used for blade icing identification. They require substantial resource data. Hence, Xu Cheng et al. [20] developed a blade icing identification technique based on blockchain and an imbalanced federated learning approach.

This technique identifies blade icing issues accurately. However, the implementation cost is high in this technique. Federated learning (FL) is an ML concept mainly used to provide data privacy in different applications. In the federated learning approach, the data is trained in a distributed way where the training dataset is located on the user side. The major problem with the FL approach is that it requires vast resources, and the computation of resources is complex.Hence, LiangGao et al. [21] developed an incentive governor for the FL approach based on blockchain technology to overcome these issues. This technique provides high security by neglecting the attackers. But the computation time is more in this approach.

As network applications are developing faster, it is widely used in different sectors. Thus, providing trustworthy applications to the user is the primary task of the researchers. Therefore, SafaOtoum et al. [22] developed a technique to provide security and network trustworthiness. This technique integrates the features of blockchain architecture and a federated learning approach. Moreover, they provide network trustworthiness by considering the trust of every individual. This method provides high accuracy. However, the detection rate of false statementsis low in this approach. This technique integrates the features of blockchain architecture and a federated learning approach. Moreover, they provide network trustworthiness by considering the trust of every individual. This method provides high accuracy. However, the detection rate of false statements is low in this approach.

In a distributed network, the federated learning approach provides high data privacy by training data in a secure manner. It is a type of collaborative learning where the training dataset is located on the user's side to preserve privacy. YajingXu et al. [23] presented the FL approach based on blockchain methodology to identify malicious events in a unified model. This approach's experimental outcome improves FL's performance on data protection and negative identification. However, malicious events are not neglected in this approach.

The key contribution of this proposed work was described as follows:

- Initially, three different data (audio, image, text) is gathered and trained in the system.
- Moreover, a novel OBHSB has been developed with different security parameters and monitoring functions.
- Then the data transmission from the different users was encrypted to hide the raw data among third parties.
- The encrypted data was stored at central cloud server, the homomorphism function has been performed during file access.
- Finally, the performance of the designed monitoring crypto model is validated by launching a DoS attack.
- The robustness has been measured regarding computation time, confidential rate, resource usage, throughput, and data transfer time.

III. System Model and Problem Statement

Cloud computing is the primary memory resource for intelligent digital gadgets because the collected information by the smart devices is stored in the cloud environment. So, the cloud memory must become secure to maximize user trust.

Hence, to secure the data sharing process in the cloud environment, federated learning has been introduced with security mechanisms. Usually, the federated system has security features, but some harmful events have broken the security. So, the blockchain system has been introduced. The blockchain module's main reasons are homomorphic concepts and data integrity validation. But in some cases, the data became injected by malicious events. So, the present work has planned to design the monitoring of the homomorphic blockchain system for cloud environment. Fig. 2 illustrates the system's basic model and common problem. Subsequently,the data transmission to the system with the help of blockchain provides more security because through the use of blockchain, the transmitted data was safer and stored in the cloud. In this proposed model, three different types of data sets were used, so a large amount of data was transmitted in a single transmission.

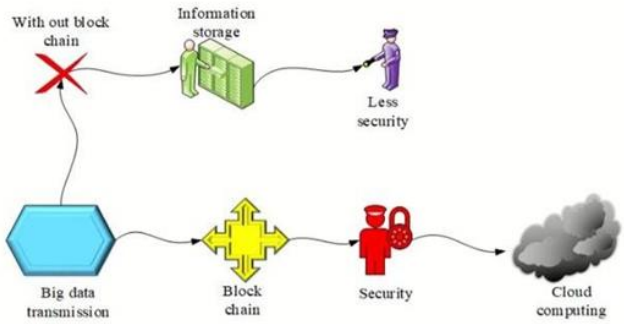

Fig. 2. System model with problem.

IV. PROPOSED METHODOLOGY

To enrich the security module of the federated learning, a novel Optimized Buffalo-based Homomorphic SHA Blockchain (OBHSB) has been designed with the required security parameters. Hence, to perform the federated learning concept, three different types of data have been considered. Initially, the collected data was trained to the system then a novel OBHSB was designed with the required data hiding and the homomorphic parameters. Here, the buffalo algorithm is incorporated to monitor the malicious activities in the designed federated learning system. Here, we provide three different types of data sets to the system. With the help of the proposed model, the system was monitored to detect malicious activities. The transferred data was stored in the cloud memory for better security. Here, finding the two hash values was termed homomorphism. After that, if the two hash values are the same, then through the help of a key, the file was accessed. Otherwise, the file was not able to access. However, the performance of the system was measured. Attack was launched to validate the proposed model; Fig. 3 illustrates architecture of the proposed model.

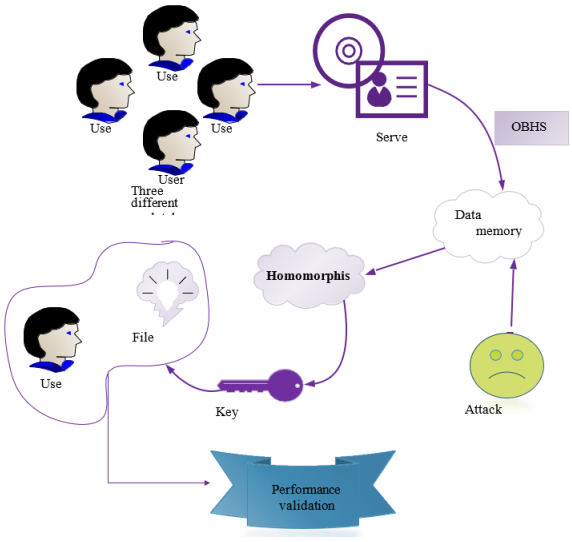

Fig. 3. Proposed architecture.

A. Design of OBHSB Model

The proposed model was designed based on the optimized buffalo algorithm and the Homomorphic SHA blockchain. At first, audio, image, and text data were collected and trained in the system to validate the implemented design. After that, the proposed model was designed to provide security for the ordered data sets. For better security, the data was stored in cloud storage. Moreover, with the help of the proposed model. encryption, decryption, and homomorphism were done. Initialization of the data set was declared through (1).

$$F[B_3] = (a_1 + i_1 + t_1, a_2 + i_2 + t_2, \dots\dots\dots\dots a_n + i_n + t_n \quad (1)$$

Where $F[B_3]$ defines the initializing parameter of the proposed model, $a_1 \dots a_n$ represents the amount of audio data presented in the data set, $i_1 \dots\dots . i_n$ defines the number of image data introduced in the initialized data set, and $t_1 \dots . t_n$ refers to the number of test data presented in the dataset.

B. Attack Prediction and Neglection

After initializing the data sets, the system was monitored through the proposed model for detecting the attack. For attack detection, Neglection, and encryption, decryption was done under the proposed technique. While initializing the data, the data set contains both the attacknode and the normal node. Among them, the attack nodes were removed from the system, andthe normal nodes were used for different processes. Here, the attack detection expression of the proposed model was declared in (2), (3).

$$Pa * (b_3) = \sum_{i=1}^{n} (a + i + t,\ a' + i' + t') \quad (2)$$

$$Na * (b_3) = b_3(a + i + t) - b_3(a' + i' + t') \quad (3)$$

Where $Pa * (b_3)$ defines the attack detecting parameters of the proposed model, $Na * (b_3)$ refers to the attack neglecting parameter of the implemented design, $a+i+t$ represents the normal data presented in the data set, and $a' + i' + t'$ defines the attack nodes of the proposed model.

C. Calculation of Hash 1

Hash 1 function was calculated after removing the attack nodes from the data set. After measuring the hash one value, the obtained value was stored in the cloud for more security. Subsequently, finding the hash one function of the proposed model was declared through the (4).

$$h^* = a \bmod b \quad (4)$$

Where h^* defines the hashing function of the proposed design. Moreover, a describes the plain text and b represents the prime number. Furthermore, the hash one value of the model was used for finding the key matching operation.

D. Data Encryption

In this research, SHA encryption was used to encrypt the data. At SHA encryption, Hexa decimal numbers were chosen for key values. Encryption was done in the data to protect the data from the unauthenticated user. Through the encryption process, the data were protected as private data and sensed data. Moreover, the SHA defies the secure Hash Algorithm used for hashing the data and securing the files. SHA encryption of the proposed model was declared through (5).

$$E^* = a * k \quad (5)$$

Where the parameter E^* defines the encryption parameter of the proposed model and a refers to the plain text and k critical parameter of the SHA encryption.

E. Homomorphism

For finding the hash, two values were mentioned as homomorphism. Hash 2 was measured by encrypting the initialized data. If the hash one and the hash two values are equal. it was mentioned as key matching so the user can access the file. Moreover, the system can access the file while displaying data injection if the two values are not equal. For finding the hash, two value of the proposed model was expressed in (6)

$$h^{**} = \frac{E^*}{k}(a \bmod b) \quad (6)$$

Where h^{**} refers to the parameter used for finding the homomorphism. E^* represents the encrypted data, k refers to the key parameter, and a defines the plain text, as well as the parameter b was the prime number.

F. Key Matching

Key matching was done to find the exact user to access the file. Moreover, if the hash one andthe hash two rates were the same, the system accomplishes the file accessing process. If the two values are not the same, then the system displays as data injected. The key matching expression of the proposed model was expressed through the 'if' condition and was declared in (7).

$$k_m = \begin{cases} if\ h^* = h^{**}; file\ accessed \\ else\ ; data\ injected \end{cases} \quad (7)$$

Where k_m represents the key matching parameter used in the proposed model. h^* Defines the hash one function of the system and h^{**} establishes the hash two parameters of the model.

G. Performance Validation and Attack Launching

The performance of the parameters was validated after finding the hash one and hash two values. Moreover, the proposed model attains a higher rate of parameters with better performance. After that, the robustness was validated by launching the attack. After launching, the attack's performance improved compared to the initial stage performances. After completing the file accessing process, the system was checked for malicious nodes. This research launched the Denial of Service (DoS) attack. Subsequently, the function of this attack was to shut down the system and quit the process when accessing the unauthenticated user, and here the unauthenticated user was considered the malicious node; after launching the attack, if any of the malicious nodes were presented.

The working procedure of the proposed model was developed in the pseudo-code format and represented in Algorithm 1 and the workflow diagram of the proposed model was shown in Fig. 4.

Algorithm 1: OBHSB

```
Start
{
    Initialization()
      int F[B3] = (a1 + i1 + t1, a2 + i2 + t2,   an + in + tn)
      //Initialize the audio, image and text data set
      Attack Prediction & Neglection( )
        {
          int P a *(b3) , Na *(b3)
          //initializing the attack detecting and neglecting
            parameters To neglect Na *(b3)
          //System was monitored, if the attack was predicted,
          then it was removed, and the attack detected node
          was considered as Na *(b3)
        }
      Hash1 ()
        {
          int h*, a, b;
          //initializing the parameters used for calculating the hash
              one function
          // hash one was calculated from the initialized data set,
             and it was stored in the cloud storage for better security
        }
      Data encryption ()
        {
          int E*, a, k;
          //initialize the data encryption parameters
          E* = a x k
          //SHA encryption method was followed for encryption
        }
      Homomorphism ()
        {
          int h**, E*, a, k, b;
          //initializing the parameters for homomorphism

          h** = E*/k (a mod b)
          //hash two value was calculated
        }
      Key matching ()
        {
          int h*, h**;
          //initializing the key matching parameters of the model
          km = { if h* = h** ; file accessed
               { else ; data injected
        }
}
Stop
```

Flow chart which explains the process of the developed model is illustrated in Fig. 4.

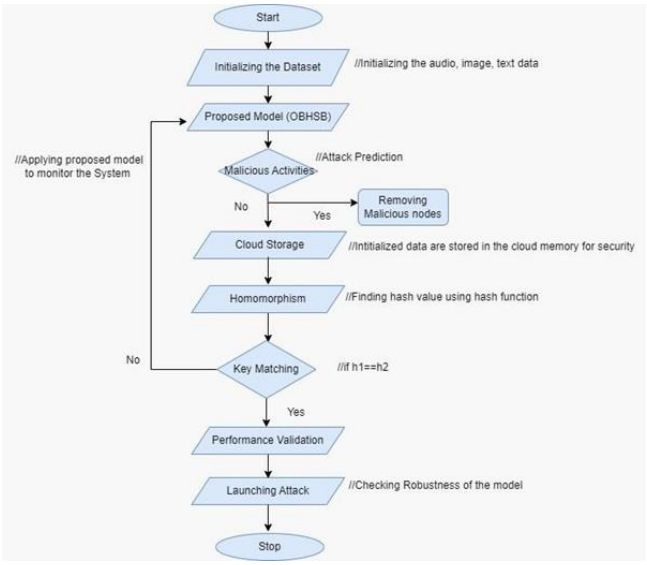

Fig. 4. Flow chart.

The primary aim of this present research was to provide security to the files presented in the cloud memory of the system. Through the proposed model, high security was attained. Consequently, the attack was launched on the cloud memory of the system.

H. Dataset Description

The proposed methodology uses datasets created manually using text dataset form Kaggle stocks data and audio mp4 has been collected randomly from google. The image dataset is created by using the images available on the Internet. Various kinds of images are available on the Internet, and the images suitable for event detection are collected for the proposed event detection process.

V. Result and Discussion

In this research, the proposed model was implemented and executed in the python platform to provide better security to the files. Here, the system provided three different types of data sets for analysis. After initializing the data sets with the help of the proposed model, the plan was monitored to predict the attack/event nodes. If any attack nodes were presented in the datasets,then the system neglected them. For more security, hash one and hash two functions were measured when both values were the same, and the user accessed the file. Subsequently, if not equal means the system displays data injection. However, the required parameters for developing the proposed design were tabulated in Table I.

TABLE I. Parameter Validation of the Developed Model

Parameters	Requirement
OS	Windows 10
Platform	Python
Version	3.10

The present work was designed in the python software 3.10 version. The performance analysis section mentioned the result after and before attack launching. Moreover, the comparative analysis of the model was evaluated and compared with the existing techniques for assuring the presentation of the implemented model.

A. Case Study

Here, the working procedure of the implemented design was explained in detail. The primary aim of this particular research was to provide security for the files from the attackers. Initially, the system was monitored through the proposed model. The file was accessed only when both the hash values were the same. Consequently, from this method, more security was provided for files. After that, the DoS attack was launched to validate the system's performance. At the performance analysis, the version of the proposed model, such as encryption time, decryption time, confidential rate, throughput, and data transfer time, were calculated. Fig.5. presents the calculated performance metrics.

The total time it takes the system to encrypt and decrypt the data is called encryption and decryption time. The encryption time, decryption time as well as the confidential rate of the proposed model is shown in Fig. 6. Here, the encryption time of the audio signal was about 10.35ms, the encryption time of the image data was approximately 17.56, and the encryption time of the text data was about 5.29ms. Subsequently, in the proposed model, the decryption rate of the audio signal was about 1.6ms, the decryption rate of the image data was about 28.2ms, decryption rate of the text data was about 6.233ms. Data transfer time was the time needed to transfer the data, and in the proposed model, less time was required to move the high amount of data. Moreover, the data transfer rate of the proposed model is 18.8ms for audio signal, 19.2ms for image data and 3.3ms for text data.

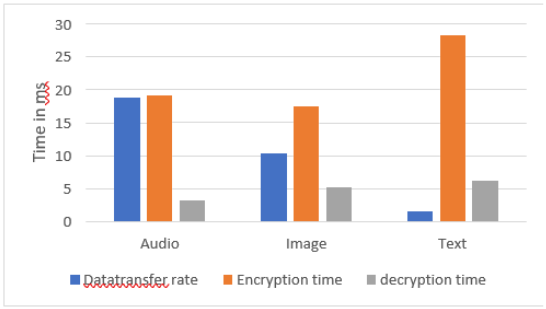

Fig. 5. Performance metrics of the proposed model.

Fig. 6 depicts the rate of throughput in the model.

The system's throughput was calculated based on the rate of exactness and time needed to complete the process. Here, the throughput range was varied at image, audio, and text data sets. However, the throughput range of the image data set was about 73%, throughput attained through the audio data set was about 94% and 85% of the throughput was attained through the text data set.

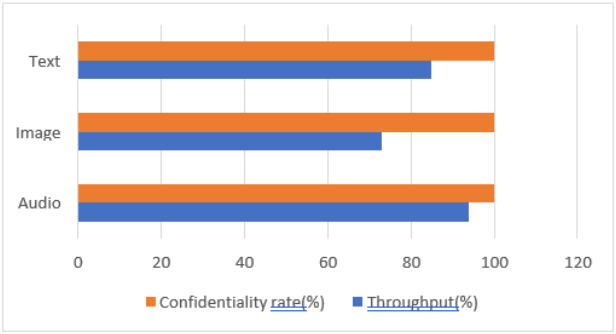

Fig. 6. Confidential rate and throughput of the model.

B. Comparison Analysis

A higher performance rate was attained through the proposed model, and validating the performance of the proposed model comparison analysis was done. In this section, the proposed model's stability, scalability, and resource usage were compared with the existing models. Moreover, the parameters of the developed model were comparing with the current models like DNN+GRM+MFO, DNN+GRM, CNN+HHO+GRM+MFO, and CNN+GRM+MFO as well as the scalability and the stability rate of the proposed model was compared to the existing models such as Recurrent neural network (RNN), Convolution neural network (CNN), Deep convolution Neural network (DNN), Deep belief network (DBN) as well as the proposed MLFC.

1) Resource usage: Here, the resource usage of the proposed model was compared with the existing models such

as DNN+GRM+MFO, DNN+GRM, CNN+ HHO+GRM+MFO, and CNN+GRM+MFO. But, the proposed model uses a lower rate of resources. Moreover, the proposed model uses 8.5MB of audio data, 7.3MB of image resources, and 9.4MB of text data. Based on the time and the performance rate, resource usage was measured. The resource usage was measured through (8).

$$R_u^* = P_\gamma * \Delta t \quad (8)$$

Where R_u^* defines the model's resource usage function, P_γ represents the proposed model's performance rate, and Δt defines the time required to complete the process. Consequently, the resource usage comparison of the proposed model was illustrated in Fig. 7. After comparison, the proposed model needs a lower rate of resources to perform the function.

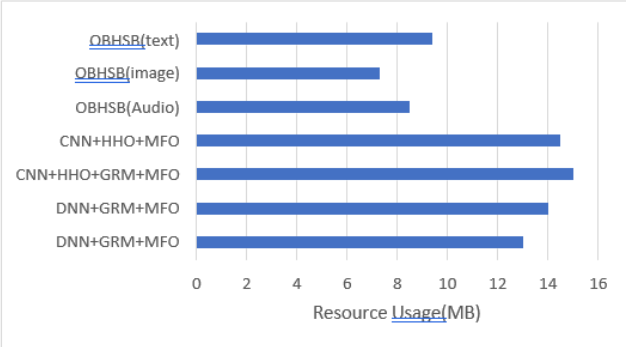

Fig. 7. Comparison of resource usage.

2) Scalability: Scalability refers to the overall capability to co-operate and perform well with data given to the model.

$$\sigma = R_u^* * \beta_t \quad (9)$$

Where σ defines the scalability function of the proposed model R_u^* was the resource usage function and β_t represented the model's throughput. Moreover, the scalability of the proposed model was about 153 MB. The scalability of the implemented model was comparing through the existing models like CNN, DCNN, RNN and DBN. Among them, the proposed model attains a better rate of scalability. The stability comparison of the proposed model is illustratedin Fig. 8.

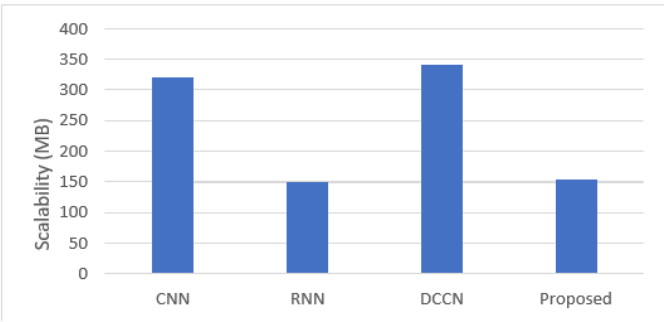

Fig. 8. Scalability comparison.

3) Stability: Stability refers to the steadiness of the system as well as the condition of the system. Subsequently, the Stability range of the developed model was comparing with the existing models such as CNN, RNN, DCNN, and DBN However, the proposed model staining a better stability range was about 99%. Stability helps to find the men's performance within a particular time. For evaluation, the exact increment of performance stability will be used. The stability comparison of the proposed model was shown in Fig. 9.

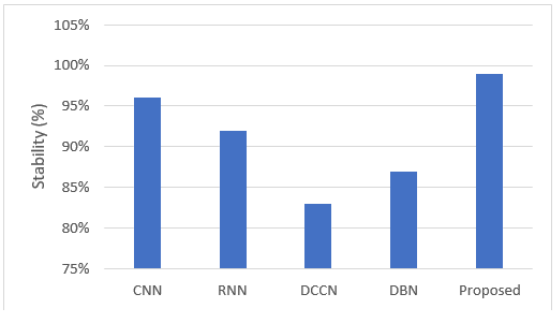

Fig. 9. Stability comparison.

C. Discussion

In this section, the performance parameters of the proposed model were discussed. In this research, the proposed model attained higher rate of throughput about lower rate of encryption time was achieved through the audio, text, and image data was about 10.35ms, 17.56ms, and 3.29ms. Moreover, the decryption time of the proposed model, also low through audio, image, and text data, was about 11.66ms, 28 2ms, and 6.233ms. A higher rate of confidential rate was achieved through the proposed model through the three different datasets, and the confidential rate was 100%. The proposed model used fewer number resources. The text data rate of resource usage was about 8.5MB, the image resource used was about 7.3MB, and 94MB resources were used through the text data. Moreover, the overall performance rate of the developed model was based on three different data sets were tabulated in Table II.

TABLE II. OVERALL PERFORMANCE OF THE IMPLEMENTED DESIGN

Parameters	Audio	Image	Text
Encryption time	10.35ms	17.ms	5.29ms
Decryption time	1.66ms	28.2ms	6.23ms
Confidential rate	100%	100%	100%
Throughput	94%	73%	85%
Data transfer rate	18.95ms	19.6ms	4.20ms
Resource usage	8.5MB	7.3MB	9.4MB

VI. CONCLUSION AND FUTURE SCOPE

This work was presented to provide security for file access. Moreover, the blockchain and homomorphism methods were used to secure the data from the unauthenticated node. Through this proposed model, more security was provided in the cloud environment. In this work, key matching and homomorphism were considered essential features. Due to

these two features, the data in the cloud was safe without accessing the unauthenticated user. In this model, the encrypted data was stored in the cloud environment for better security, and the percentage of improvement was also added in this section. This proposed model measured the output based on image, text, and the audio data set. Subsequently, through the audio, image, and text data, the proposed model needs a low encryption time was about 10.35ms, 17.6ms, and 5.29ms, andthe decryption time of the proposed model is also low with three different datasets was about 11.66ms, 28.2ms and 6.23ms. The amount of resource usage of the proposed model with the audio, image, and text data was about 8.5MB, 7.3MB, and 9.4MB. On comparing with existing models, the developed design needs a lower rate of resources to perform the function. The system attains 94% throughput with the audio data set, 73% of throughput in the image data set, and 85% in the text data set. The data transfer rate of the implemented design along with audio, image, and text data was 18.9MB, 19.6MB, and 4.20MB. Subsequently, the proposed model attains 100% of the confidential rate while processing the audio, image, and text data set. Resource usage of the proposed model was 8.5 MB in the audio data set, 7.3 MB in the image data set, and 9.4 MB in the text data set. Among them, the proposed model developed by 4% in resource usage. The Stability of the system was about 99% comparing with the existing techniques; the implemented design improves 3% of stability rate. Consequently, the overall system scalability range obtained through the proposed model was about 153MB; compared with the existing models, the scalability of the proposed model was developed by 5%. Due to the high confidentiality rate, the system provides more security to the files presented in the cloud environment. In future, the model can be checked with other attacks also to check the stability of the model.

REFERENCES

[1] Alazab, Mamoun, et al. "Federated Learning for Cybersecurity: Concepts, Challenges, and Future Directions." IEEE Transactions on Industrial Informatics 18.5 (2021): 3501-3509.

[2] Huo, Weiwei, et al. "Performance prediction of proton-exchange membrane fuel cell based on convolutional neural network and random forest feature selection." Energy Conversion and Management 243 (2021): 114367.

[3] Chen, Mingzhe, et al. "Distributed learning in wireless networks: Recent progress and future challenges." IEEE Journal on Selected Areas in Communications (2021).

[4] Zhu, Hangyu, et al. "Federated learning on non-IID data: A survey." Neurocomputing 465 (2021): 371-390.

[5] Wink, Tobias, and Zoltan Nochta. "An approach for peer-to-peer federated learning." 2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W). IEEE, 2021.

[6] Long, Guodong, et al. "Federated learning for privacy-preserving open innovation future on digital health." Humanity Driven AI. Springer, Cham, 2022. 113-133.

[7] Rezaee, Khosro, et al. "A survey on deep learning-based real-time crowd anomaly detection for secure distributed video surveillance." Personal and Ubiquitous Computing (2021): 1-17.

[8] De, Suparna, et al. "Analysing environmental impact of large-scale events in public spaces with cross-domain multimodal data fusion." Computing 103.9 (2021): 1959- 1981.

[9] Sahoo, Somya Ranjan, and Brij B. Gupta. "Multiple features based approach for automatic fake news detection on social networks using deep learning." Applied Soft Computing 100 (2021): 106983.

[10] Hassan, Mohammad A. "Relational and NoSQL Databases: The Appropriate Database Model Choice." 2021 22nd International Arab Conference on Information Technology(ACIT). IEEE, 2021.

[11] Komninos, Nicos, et al. "Towards high impact smart cities: A universal architecture based on connected intelligence spaces." Journal of the Knowledge Economy 13.2 (2022): 1169-1197.

[12] Leite, Emilene. "Innovation networks for social impact: An empirical study on multi- actor collaboration in projects for smart cities." Journal of Business Research 139 (2022): 325-337.

[13] Csukás, Máté S., and Roland Z. Szabó. "The many faces of the smart city: Differing value propositions in the activity portfolios of nine cities." Cities 112 (2021): 103116.

[14] Corsi, Alana, et al. "Big data analytics as a tool for fighting pandemics: a systematic review of literature." Journal of ambient intelligence and humanized computing 12.10 (2021): 9163-9180.

[15] Pika, Anastasiia, et al. "Using big data to improve safety performance: an application of process mining to enhance data visualisation." Big Data Research 25 (2021): 100210.

[16] Iqbal, Naeem, et al. "A novel blockchain-based integrity and reliable veterinary clinic information management system using predictive analytics for provisioning of quality health services." IEEE Access 9 (2021): 8069-8098.

[17] Naeem, Muhammad, et al. "Trends and future perspective challenges in big data." Advances in intelligent data analysis and applications. Springer, Singapore, 2022. 309-325.

[18] Chen, Jie, L. Ramanathan, and Mamoun Alazab. "Holistic big data integrated artificialintelligent modeling to improve privacy and security in data management of smart cities." Microprocessors and Microsystems 81 (2021): 103722.

[19] Ayaz, Ferheen, et al. "A blockchain based federated learning for message dissemination in vehicular networks." IEEE Transactions on Vehicular Technology 71.2 (2021): 1927-1940.

[20] Cheng, Xu, et al. "A Blockchain-Empowered Cluster-based Federated Learning Model for Blade Icing Estimation on IoT-enabled Wind Turbine." IEEE Transactions on Industrial Informatics (2022).

[21] Otoum, Safa, Ismaeel Al Ridhawi, and Hussein Mouftah. "Securing critical IoT infrastructures with blockchain-supported federated learning." IEEE Internet of ThingsJournal 9.4 (2021): 2592-2601.

[22] Gao, Liang, et al. "FGFL: A blockchain-based fair incentive governor for Federated Learning." Journal of Parallel and Distributed Computing 163 (2022): 283-299.

[23] Xu, Yajing, et al. "BESIFL: Blockchain Empowered Secure and Incentive Federated Learning Paradigm in IoT." IEEE Internet of Things Journal (2021).

AUTHORS' PROFILE

K Prasanna Lakshmi is working as a Professor and Dean Academics in Information Technology Department in Gokaraju Rangaraju Institute of Engineering and Technology (GRIET), Hyderabad. She received the Bachelor's Degree in Mechanical Engineering (2000), Master's Degree in Computer Science (2002) and Ph.D in Computer Science (2016) in the area of Data Stream Mining. She has total 20 years of teaching and 13+ years of research experience. She is a reviewer for various conferences, journals and books, Editorial Board Member for IGI Global Publications. Her research interests include Data Science, Artificial Intelligence, Data Stream Mining, Web Mining, Big Data Analysts, Social Networking. Published many research articles in renowned national and international journals and conferences.

K. Swapnika pursuing Ph.D in Data Mining and Information Retrieval Systems stream at Jawaharlal Nehru Technological University Hyderabad. She completed M. Tech in Software Engineering from Jawaharlal Nehru Technological University Hyderabad and has 10 years of academic experience. Her Research Interest includes Information Retrieval Systems, Deep Learning, Cloud Computing and Blockchain Technology.