

**International Journal of Simulation and Process Modelling**

ISSN online: 1740-2131 - ISSN print: 1740-2123

<https://www.inderscience.com/ijspm>

---

**Secured big data storage in cloud by intelligent authentication and privacy preservation via intelligent deep learning-aided heuristic strategy**

Calistus Mohandhas Varun, Raphel Pankiras Anto Kumar, Michael Raj Judith Reya, Rangini Murugan

**DOI:** [10.1504/IJSPM.2024.10066645](https://doi.org/10.1504/IJSPM.2024.10066645)

**Article History:**

Received:	11 March 2023
Last revised:	22 December 2023
Accepted:	27 December 2023
Published online:	04 October 2024

---

## Secured big data storage in cloud by intelligent authentication and privacy preservation via intelligent deep learning-aided heuristic strategy

---

Calistus Mohandhas Varun\*

R.M.K. Engineering College,  
Kavaraipettai-601206, Tamil Nadu, India  
Email: cmvarun87@gmail.com  
\*Corresponding author

Raphel Pankiras Anto Kumar

St. Xavier's Catholic College of Engineering,  
Chunkankadai-629003, Tamil Nadu, India  
Email: antokumar@sxcce.edu.in

Michael Raj Judith Reya

Saveetha Engineering College (Autonomous),  
Saveethanagar, Sriperumbadurtaluk,  
Chennai, Tamil Nadu-602105, India  
Email: judithreya1712@gmail.com

Rangini Murugan

Velammal Engineering College,  
Ambattur Red Hills Rd, Velammal Nagar,  
Surapet, Chennai, Tamil Nadu 600066, India  
Email: ranju020790@gmail.com

**Abstract:** This paper is exploring a new way of offering the secure authentication and privacy maintenance of big data over the cloud. It ensures the safeguard of big data using deep learning methods. The biometric information is used for offering secure authentication to the cloud data to avoid malicious entries. The secured key is extracted via encryption, where the modification is done by the PDGCBO algorithm. Extracted key is used for protecting the privacy of cloud data through the CMECHE better than other traditional algorithms regarding performance and security. The computational results show the effectiveness of the developed algorithms in this security framework by optimising the parameters of LSTM and fuzzy network using the PDGCBO algorithm. The experimental validation takes place in order to provide better outcomes when compared with state-of-the-art methods.

**Keywords:** big data; cloud computing; key optimisation; long short-term memory with fuzzy network; cascaded modified elliptic-curve cryptography with homomorphic encryption; probability-based darts game colliding bodies optimisation.

**Reference** to this paper should be made as follows: Varun, C.M., Anto Kumar, R.P., Judith Reya, M.R. and Murugan, R. (2024) 'Secured big data storage in cloud by intelligent authentication and privacy preservation via intelligent deep learning-aided heuristic strategy', *Int. J. Simulation and Process Modelling*, Vol. 21, No. 2, pp.90–107.

**Biographical notes:** Calistus Mohandhas Varun is presently working as an Assistant Professor in Department of Computer Science and Business Systems at R.M.K. Engineering College, since August 2022. He is in the teaching profession since 2011. He obtained his BE in Computer Science and Engineering from Anna University, Chennai and ME in Computer Science and Engineering degree from Anna University, Tirunelveli. He is currently pursuing PhD in Anna University, Chennai. His area of interest includes computer programming, web application, cloud computing and data science. He is a Premium Instructor in Udemy. He has been the trainer for various workshops with topics related to computer programming, information security, cloud infrastructure, multimedia and web designing.

Raphel Pankiras Anto Kumar received his PhD degree from Bharathiyar University, Coimbatore, in September 2014. He is currently a Professor in the Department of Computer Science and Engineering at St. Xavier's Catholic College of Engineering, Nagercoil, Tamil Nadu, India. His research interest includes image processing, biometrics, cloud computing and security.

Michael Raj Judith Reya is an Assistant Professor at Saveetha Engineering College with two years of teaching experience. Her expertise lies in database management systems, data structures, and machine learning. She is dedicated to fostering academic excellence and innovation among her students.

Rangini Murugan is an Assistant Professor with two years of experience. She specialises in image processing and deep learning. She integrates fieldwork with advanced analytical techniques and theoretical learning for students.

## 1 Introduction

In recent times, massive amounts of data's are stored in the cloud which is utilised to secure the data. According to other research, roughly 75% of digital data are the same (or replicate) (Tao et al., 2020), and backup and archive storage systems have data redundancy levels that are substantially higher than 90% (Zhou et al., 2022). There are costs for the maintenance, management, and managing of such massive data, even while the storage price is very low and developments in cloud storage outcomes allows to store a rising count of data (Yang et al., 2020). Therefore, it is unexpected that attempts have been made to decrease administrative costs brought on by data duplication. By maintaining just one copy of redundant data, the data de-duplication technique aims to locate and eliminate duplicate data. To put it another way, data de-duplication methods can drastically decrease bandwidth and storage needs (Cui et al., 2019). Data (especially sensitive information) are expected to be encrypted ahead of outsourcing, nevertheless, as owners of the data and users might not entirely believe cloud storage suppliers (Zaghloul et al., 2020). One of the most popular favours offered by the cloud storage is sharing the data. Users can exchange their data with other users using a data-sharing service, which lessens the need for local data storage. However, whenever users share their data in the cloud, they forfeit physical contribution over it. Any error (human negligence or hardware/software malfunction) could result in data loss or damage (Deng et al., 2020). Plans for sharing data have been put out (Zhu et al., 2019). Users should be removed from groups when they act inappropriately or left the group. Revocation of the user is thus a frequent and practical requirement in cloud storage audits for shared information.

Big data is a large volume, high velocity, high diversity information benefit that calls for novel processing techniques in order to improve decision-making, uncover new insights, and streamline processes (Zeng and Choo, 2018). Big data handling with available database management technologies is challenging because of its complexity and size (Zhang et al., 2017). Outsourcing the information to a server with the ability to store large amounts of data and quickly handle user access requests is a practical approach (Zhang et al., 2019). Standard data

management systems are put to the test by the velocity, huge volume, and different data being created by many scientific and commercial areas, necessitating their scaling while assuring dependability and safety. The fundamental difficulty is evaluated based on where and how to keep the enormous count of data that is being simultaneously created. For many companies, private infrastructures are their first choice (Yang et al., 2020). Data centres must be built and maintained, which is expensive, involves specialised labour, and can be problematic for sharing (Chen et al., 2020).

The majority of currently used methods for protecting outsourced big data in clouds are either ABE or secret sharing based. ABE-based systems (Wang et al., 2021), give a data owner the freedom to specify in advance the user group who are permitted access to the data, but they are limited by how difficult it is to effectively update the access control policies and cipher text. A secret can be shared and rebuilt by a group of cooperative users using mechanisms for secret sharing (Xu et al., 2016). However, they frequently require asymmetric public key cryptography, such as 'Ron Rivest, Adi Shamir', and 'Leonard Adleman (RSA)', for user validity verification, which has a high computational cost (Senthilnathan et al., 2018). Additionally, it is a difficult problem to dynamically and effectively change the access policies in accordance with the new demands of the owners of the data in covert sharing techniques. Distributed storage, one of the key cloud applications, has made it possible to store large amounts of remote data using the STaaS model (Khan et al., 2023). Along with the growth of Internet networks and services, this cloud service paradigm has mostly come to be accepted as a big data strategy (Dehghani et al., 2020). 'Google Drive and Microsoft's One Drive' are just two of the popular storage service providers that provide users with large, expandable cloud-based storage spaces. Among all the techniques that have been used, this paper introduces a novel idea about secured data transmission.

The main offerings of the suggested methodology are explained below.

- To plan a method of secured big data in the cloud by intelligent authentication and privacy preservation via a heuristic-aided deep learning strategy that helps the user to access and store the data safely in the cloud.

- To encrypt the data by CMECHE model, which is developed by the hybrid technique of HE and ECC with the help of an optimised key that helps to obtain the standard encryption. Here, double encryption is occurred while processing the data.
- To propose the PDGCBO algorithm for tuning the parameters called, private key in ECC, 'hidden neurons in LSTM', 'epochs in LSTM', and exponential bound in fuzzy which gives the optimal solutions.
- To obtain the optimised key, using LSTM with fuzzy network method. It provides the optimal key for the authenticated users.
- To estimate the efficiency of the methodology utilising distinct factors and compared with the help of other traditional approaches of optimisation and the classifiers of deep learning.

The framework of the work is explained here. Section 2 explains the conventional model of secured big data storage model. Section 3 explains the secured big data storage cloud in the cloud sector via an intelligent deep learning sector. Section 4 elaborates on the PDGCBO algorithm for the secured big data in the cloud. Section 5 explains the CMECHE in secured big data storage. Section 6 reviews the outcomes and discussions of the proposed model. Finally, Section 7 summarises the proposed model.

## 2 Literature review

### 2.1 Related works

In 2017, Li et al. (2017) have explained a method of clever cryptography that ignores cloud facility providers from directly working partial data. The suggested model separated the file and kept the information on the scattered cloud servers independently. To reduce the operation time, a different method was developed to assess if the data packets needed to be split. The SA-EDS model was the name of the suggested scheme, and it was primarily maintained by the recommended algorithms AD2, SED2, and the EDCon. The experiments have demonstrated that our technique can successfully fight against the key risks from clouds and needs with an adequate calculation time. Their experimental findings have tested both privacy and efficiency abilities.

In 2021, Mendes et al. (2021) have proposed to meet the regulatory requirements for sensitive information, a cloudbacked storage system that could store and share massive data in a safe, dependable, and effective manner utilising a variety of cloud providers and storage repositories. Charon developed three distinctive features:

- 1 it did not need client-managed servers
- 2 it does not need any confidence in a single entity
- 3 it effectively handles huge files across a number of geo-distributed storage providers.

Additionally, in order to prevent write to write problems between clients accessing shared repositories, they have created a 'novel Byzantine resilient data-centric leasing' mechanism. They tested Charon utilising micro-and benchmarks based on the applications that simulated typical workflows from the well-known large data field of bioinformatics. The findings demonstrate that not just was innovative architecture workable, but it also offered superior end-to-end performance compared to competing for cloud-based systems.

In 2022, Yang et al. (2022) have examined a three tier cross domain infrastructure and suggested effective and privacy maintaining in big data redundancies in cloud storage. EPCDD had accomplished data availability, privacy preservation, and resistance against 'brute-force attacks'. Additionally, accept responsibility into account to provide schemes with higher privacy guarantees. It shown that, in terms of computing, storage overheads, and communication, EPCDD beats the currently used rival techniques. Additionally, the EPCDD duplication search had logarithmic temporal difficulty.

In 2020, Zhang et al. (2020) have suggested an innovative storage auditing approach. This was accomplished by investigating a revolutionary generation of a key mechanism and a fresh method for updating private keys. By employing this technique and method, they have achieved user revocation by only upgrading the private keys of the members who have not had their access to the system revoked, as opposed to their authenticators. Whenever the authenticators were not changed, integrity auditing of the data belonging to the revoked user could still be correctly carried out. The suggested technique, removed the challenging certificate administration in conventional public key infrastructure (PKI) systems because it was based on identity-based encryption. Through analysis and testing findings, the recommended scheme's effectiveness and safety were confirmed.

In 2018, Hu et al. (2018) have explained a safe and measurable control access approach that was based on the NTRU cryptosystem for massive data storage in clouds. In order to fix the decryption issues with the original NTRU, they first presented a new NTRU decryption method. They then went on to describe their plan and examine its accuracy, security, and computing efficiency. Their method enabled the server of the cloud to effectively upgrade the cipher text whenever the owner of the data specifies a new access strategy. The owner of the data was also able to evaluate the update to prevent fraudulent cloud activity. Additionally, it allowed:

- 1 A user to examine the data provided by others for accurate plaintext retrieval and the data holder and qualified users to successfully authenticate the validity of a person for obtaining data.
- 2 A thorough investigation revealed that their system can stop cheating by qualified users and withstand numerous assaults, including the collusion attack.

In 2020, Prabhu Kavın et al. (2020) have developed an improved security strategy for protecting cloud users' data in the cloud environment. The 'access control mechanism', 'encryption or decryption methods', and 'digital signature algorithms' make up the new security infrastructure. Here, a brand-new key generation procedure based on ECC was suggested for producing highly safe keys. Additionally, a brand new method Id-EAC was also put out in this work to limit the access to data of cloud users to various types of information. To secure the data of cloud users in the cloud infrastructure, a new binary value based two phase encryption and the decryption technique which related to the ECC-based values of the key was developed. To safeguard the integrity of this suggested security architecture, novel lightweight digital signature algorithms based on modulo functions were also presented. High levels of data protection, accessibility, and integrity for user data were offered by this security architecture. The experimental findings have demonstrated that the proposed methods in this security architecture were more efficient and secure than other old techniques.

## 2.2 Research gaps and challenges

Cloud storage technique has been considered an attractive service that has an optimal method to tackle a sufficient count of data. It has been used by various types of end users such as individual users, enterprises as well as organisations for the purpose of storing personal cloud data environment. SA-EDS (Li et al., 2017) technique has secured data in processing time as well as has a positive relation among various data sizes. The computational cost is also less than other active techniques. Data duplication securing is not explored in this model. Charon (Mendes et al., 2021) technique does not acquire any of the client's managed servers. It is also considered an effective model as it has the ability to deal with large files. But it is regarded as a non-feasible model and it is limited to real-time execution. EPCDD (Yang et al., 2022) method has provided better privacy assurance when assimilated over other models. It also has the potential to minimise duplicate information disclosures. But, this technique consumes more time while duplicate search. The storage auditing scheme (Zhang et al., 2020) method has offered enhanced efficiency as well as security on both the group user side and the cloud side. But, this model faces difficulty, while performing on large-scale data. The NTRU cryptosystem (Hu et al., 2018) technique has offered a verification process to the users to validate data. The computational difficulty and security strength of this technique are better. The threshold secret sharing involves a limited access structure in this model. Id-EAC (Prabhu Kavın et al., 2020) technique has provided the data integrity, retrieval as well as storage of the data in a secure manner. The encryption and decryption keys are also provided to the users. But, it is limited by temporal constraints on decision-making in the data communication process. AR-RRNS (Tchernykh et al., 2019) technique is considered effective for correcting and detecting the error in the residues. It is also effective and secured over other

models. But, this model faces the issue of redundancy and speed. SADS-Cloud (Narayanan et al., 2020) method is effective in terms of compression ratio and it is regarded as a beneficial technique. But, this model needs to speed up the encryption and decryption operation as it degrades the operation. To solve the limitations in the existing method, this paper delivers a new methodology for secured big data storage via intelligent deep learning-aided heuristic strategy.

## 3 Secured big data storage in the cloud sector through an intelligent deep-learning model

### 3.1 Proposed privacy preservation-based big data storage in cloud

A new method has been implemented. Figure 1 displays the framework of the proposed method.

In this work, to safeguard the information in the cloud PDGCBO-based deep learning algorithm is used. In CMECHE, the input cloud data is first sent to the HE model where the encryption process is processed. Next, the data which is encrypted is sent to the ECC where the encrypted data is again encrypted to produce the double encrypted data. To gain the optimal solutions, the optimised key in the CMECHE model is determined by the PDGCBO algorithm. The secure key is used to access the cloud storage. The secure key is obtained for the user by using the optimised LSTM with the fuzzy model. In this model, the iris image is given as input, where the outcome is acquired in form of a key that is for the authenticated user. Some of the hyper factors are tuned by the PDGCBO method for removing the complexity and increasing the accuracy value. Finally, it creates the secured private key for the particular user. Then the data is accessed in the cloud storage.

### 3.2 User authentication using iris biometric information

The iris images are taken as input for the secured big data in the cloud. The representation of this iris image is denoted as  $I$ . Here five sample images are taken for the process. They are further evaluated in this proposed model for the optimal solutions. The sample iris image for the proposed secured big data in the cloud is given in Figure 2.

### 3.3 LSTM with fuzzy-based user authentication

LSTM (Sherstinsky, 2020) is an artificial NN utilised in the fields of deep learning and AI. Here the iris image  $I$  is given as input. Not like the other general feed forward NN, LSTM has the connections for feedback. These characteristics make LSTM networks ideal for processing and predicting data. LSTM networks have the ability of understanding long-term dependencies in sequential data, which makes them well suited for tasks such as data access and storage. The LSTM method is developed to neglect the exploding and vanishing gradient issues in the data. Each time LSTM takes input  $y_u$ , the candidate cell gate is referred to as  $d_u$ , and

the three gates are  $j_u$ ,  $g_u$  and  $p_u$  expressed below equation (1) to equation (4):

$$j_u = \sigma(X_v [i_{u-1}; y_u] + c_v) \quad (1)$$

$$g_u = \sigma(W_g [i_{u-1}; y_u] + c_g) \quad (2)$$

$$p_u = \sigma(X_p [i_{u-1}; y_u] + c_p) \quad (3)$$

$$\hat{d}_u = \tanh(X_d [i_{u-1}; y_u] + c_d) \quad (4)$$

Here the sigmoid function is indicated as  $\sigma$ .  $i_{u-1}$  refers to the hidden state at times  $(u-1)$  or  $u$ . This is utilised for the next layer input or output in the LSTM.  $g_u$  is for the outcome of forget gate.  $j_u$  is the input gate result.  $\hat{d}_u$  is the outcome of the member cell gate.  $p_u$  is the hidden gate result.  $W_g$  and  $X_p$  are the same set of parameters that helps to decrease the speed of the network to learn. Utilising the outcomes of the cell state and the cell gate, the state of the cell at present has updated in equation (5)

$$d_u = g_y * D_{u-1} + j_u * \hat{d}_u \quad (5)$$

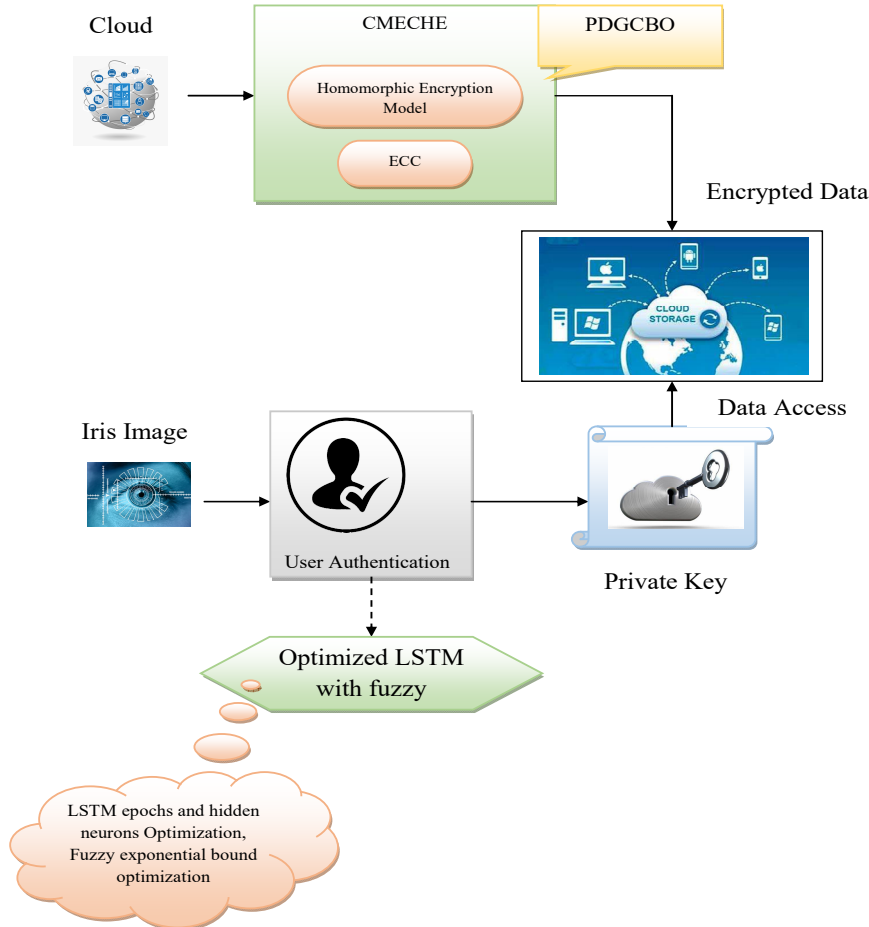
Utilising the outcomes of the cell gate and output gate the hidden state is updated which is expressed in equation (6).

$$i_u = p_u * \tanh(d_u) \quad (6)$$




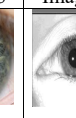
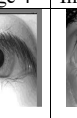
The output is obtained which is gained from the LSTM technique that will be used for finding the authenticated key together with fuzzy network output data.

*Fuzzy network* (Sun and Jang, 2018) is a learning machine that helps to find the fuzzy system parameters. Here also the iris image  $I$  is given as input. ‘Fuzzy min-max NN’ is a unique type of a neuro fuzzy method which has high effectiveness contrasted to the other models. This neural network utilised a multiple level tree shape structure where every overlapped region of the first node is managed by a sub node in the next level. A fuzzy neural network is a feed forward network which is made up of three different layers: an input layer for fuzzy, a hidden layer contains the fuzzy conditions, and the last fuzzy output layer. A fuzzy neural network concatenates the advantages of both the neural networks and Fuzzy logic making them a standard hybrid method. They permit the integration of expert knowledge into the system and are assumed inherently more knowable because of their use of human like fuzzy inference. The typical classification of fuzzy is expressed below.

**Figure 1** Fundamental architecture of the secured big data storage in the cloud method (see online version for colours)



**Figure 2** The sample iris images for the proposed secured big data in the cloud (see online version for colours)

Description	Image 1	Image 2	Image 3	Image 4	Image 5
Sample iris images					

If  $Y_1$  is  $A$  and  $Y_2$  is  $B$  then  $X$  is  $C$ . Here,  $Y_1$  and  $Y_2$  are input variables.  $A$  and  $B$  are linguistic terms featured by membership functions which explain the characters of  $C$  is the class of the object.

The fuzzy network output gained and concatenated with the LSTM network output to obtain the optimal key.

#### 4 Probability-based darts game colliding bodies optimisation for secured big data storage in the cloud sector

##### 4.1 Probability-based darts game colliding bodies optimisation

To get an optimal solution for this model, PDGCBO is proposed. This recommended PDGCBO algorithm is a combination of two algorithms called DGO and CBO. The functionalities of these two algorithms are explained here. The DGO algorithm is implemented depending on the simulation of the noticing darts play. In this algorithm, the members of the population are referred to as the darts players who try to gather more amounts of points in their throws toward the board of the game. Being the first player to strike each number on the board in order from 1 to 20 is the goal. Once, twice, or thrice hits on a number count, and the player must hit a number in order to go on to the next. After the completion of the third throw, the players switch roles. The winner is the first person to score a 20. The second method called CBO is one of the search algorithms. It is motivated based on the energy and the law of momentum in physics. In this method, collisions are happening between solid bodies. The benefits of using DGO are this, optimisation has very simple equations. It has good exploration abilities and also good exploitation capacities. It is utilised to overcome multi-objective issues. Also, the CBO does not depend on any of the internal parameters. The algorithms are very easy to understand. It displays a very fast converging feature. However, DGO lacks accuracy. The complexity of this method is high. Also, the CBO has low robustness and efficiency. To overcome these two method's drawbacks new proposed PDGCBO algorithm has been implemented.

In conventional DGO, the parameter  $P$  is considered for the probability of players. It takes the random value as lie between 0 and 1. So it lacks in accuracy. So proposing a new formulation for the probability expressed as  $Q$  in the new suggested algorithm. The parameter  $Q$  represents the probability of the player in the darts game. When the probability is  $Q > 0.5$  then, the DGO algorithm is updated. Otherwise, the CBO algorithm is updated.

The mathematical expression of the existing two algorithms is given below.

- *DGO* (Dehghani et al., 2020): search agents in DGO are the players of the game and the main goal of this game is to result in the optimal score. The calculation model of DGO is explained below.

The proposed algorithm is designed with a matrix, every row denotes one player and every column indicates the distinct features of every player. Always the count of columns in the matrix is similar to the count of problem variables. The matrix for the players has represented in equation (7).

$$Y = \begin{bmatrix} Y_1 & y_1^1 & \cdots & y_1^e & \cdots & y_1^o \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_j & y_j^1 & \cdots & y_j^e & \cdots & y_j^o \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_M & y_M^1 & \cdots & y_M^e & \cdots & y_M^o \end{bmatrix} \quad (7)$$

Here, the matrix of the player denoted by  $Y$ .  $y_j^e$  is the  $e^{\text{th}}$  dimension of the  $i$  player. The count of variables is denoted as  $o$  also.  $M$  is the count of the players.

To find out the fitness function  $y_j$  is used which is shown in equation (8) to equation (13).

$$G_{best} = \min(\text{fit})_{M \times 1} \quad (8)$$

$$Y_{best} = Y(\text{locof} \min(\text{fit}), 1:n) \quad (9)$$

$$G_{worst} = \max(\text{fit})_{M \times 1} \quad (10)$$

$$Y_{worst} = Y(\text{locof} \max(\text{fit}), 1:n) \quad (11)$$

$$G^m = \frac{\text{fit} - G_{worst}}{\sum_{i=1}^M (\text{fit}_i - G_{worst})} \quad (12)$$

$$Q_j = \frac{G_j^m}{\max(G^m)} \quad (13)$$

Here,  $G_{best}$  denotes the value of the best fitness function,  $Y_{best}$  represents the value of best variables,  $G_{worst}$  is the value of the worst fitness function,  $Y_{worst}$  and denotes the value of the worst variable. The normalised value of the fitness function is denoted as  $G^m$ , and also the  $Q_j$  is the function of the probability of the  $i^{\text{th}}$  player.

The score of the throw is designed and evaluated for every player utilising the equations from equation (14) to equation (17).

$$D_j = \text{rnd}(82 \times 1 - Q_j) \quad (14)$$

$$TD_j = \begin{cases} T(1:D), & \text{ran} < Q_j \\ T(D+1:82), & \text{else} \end{cases} \quad (15)$$

$$t_j = TD_j(l) \text{ \& } 1 \leq l \leq 82 \quad (16)$$

$$t_j^m = \frac{\sum_{throw=1}^t t_j^{throw}}{180} \quad (17)$$

Here,  $TD_j$  indicates the candidate score for the  $j^{\text{th}}$  player, the score matrix represented as  $T$ , which is arranged from max scores to min scores,  $t_j$  denotes the score for every row of  $j^{\text{th}}$  player, and the normalised score is represented as  $t_j^m$  for the  $j^{\text{th}}$  player. At last, the status of every player has upgraded using equation (18).

$$Y_j = Y_j + \text{ran}(1, n) \times (Y_{\text{best}} - 3t_j^m y_j) \quad (18)$$

The CBO algorithm is implemented to find out a new simple and efficient solution.

- *CBO (Kaveh and Mahdavi, 2014)*: the primary aim of the proposed algorithm is to evaluate a tuning process that is frequently needed in experimental options. In the CBO method, every solution candidate  $W_k$  have a large amount of variables (i.e.,  $W_k = \{W_{k,j}\}$ ) are represented as a colliding body.

The starting places of CBs are estimated with arbitrary population initialisation of the separate terms in the space of search which is given in equation (19).

$$w_k^0 = w_{\min} + \text{ran}(w_{\max} - w_{\min}), k = 1, 2, \dots, m \quad (19)$$

where  $w_k^0$  indicates the starting value vector of the  $k^{\text{th}}$  CB value,  $w_{\max}$  and  $w_{\min}$  are the maximum and minimum permitting vectors of variables. The arbitrary number in the limit of  $[0, 1]$  is represented as  $\text{ran}$ , and also the CB number is denoted as  $m$ .

Every CBs magnitude of the body boss is evaluated using equation (20).

$$n_i = \frac{\frac{1}{\text{fit}(i)}}{\sum_{k=1}^m \frac{1}{\text{fit}(k)}}, i = 1, 2, \dots, m \quad (20)$$

where  $\text{fit}(k)$  is represented as the agent  $k$ 's objective function. The population size is indicated as  $m$ . For maximisation the  $\text{fit}(k)$  is replaced by  $\frac{1}{\text{fit}(k)}$ .

The arranged CBs are splitted into two groups.

- The 'stationary CBs': the velocities of these CBs are zero and they are good agents that are stationary explained in equation (21).

$$u_k = 0, k = 1, \dots, \frac{m}{2} \quad (21)$$

- The mobile CBs: these CBs are moving towards the lower half. Before the collision, the change of the body place is denoted in equation (22).

$$u_k = w_k - w_{k-\frac{1}{2}}, k = \frac{m}{2} + 1, \dots, m \quad (22)$$

where  $w_k$  and  $u_k$  are the position vector and the velocity of the  $k^{\text{th}}$  CB. CB pair position of the  $k^{\text{th}}$  is denoted as  $k - \frac{1}{2}$ .

The colliding body's velocity in every group was evaluated in equation (25) after the collision using equations (23) and (24).

$$U'_1 = \frac{(n_1 - \sigma n_2)u_1 + (n_{21} + \sigma m_2)u_2}{n_1 + n_2} \quad (23)$$

$$U'_2 = \frac{(n_2 - \sigma n_1)u_2 + (n_{11} + \sigma m_{12})u_{12}}{n_1 + n_2} \quad (24)$$

$$U'_k = \frac{(n_{ki} - \sigma n_{k-\frac{m}{2}})u_k}{n_k + n_{k-\frac{m}{2}}}, k = \frac{m}{2} + 1, \dots, m \quad (25)$$

where  $U'_k$  and  $U_k$  were the velocities of the  $k^{\text{th}}$  mobile CB after and before the collision accordingly.  $n_k$  is denoted as the mass of the  $k^{\text{th}}$  CB and also  $n_{k-\frac{m}{2}}$  is represented as the mass of the  $k^{\text{th}}$  CB pair. The every stationary velocity for CB after the collision is given in equation (26).

$$U'_k = \frac{(n_{ki+\frac{m}{2}} + \sigma n_{k+\frac{m}{2}})u_{k+\frac{m}{2}}}{n_k + n_{k+\frac{m}{2}}}, k = 1, \dots, \frac{m}{2} \quad (26)$$

where  $k + \frac{m}{2}$  and  $U'_k$  were the velocities of the  $k^{\text{th}}$  motion CB pair before and the  $k^{\text{th}}$  stationary CB after the collision accordingly. The COR parameter is denoted as  $\sigma$ .

The new places of the CB are estimated utilising the developed velocity after the collision in CB is expressed in equation (27).

$$w_k^{\text{new}} = w_{k-\frac{m}{2}} + \text{ran} \cdot u'_k, k = \frac{m}{2} + 1, \dots, m \quad (27)$$

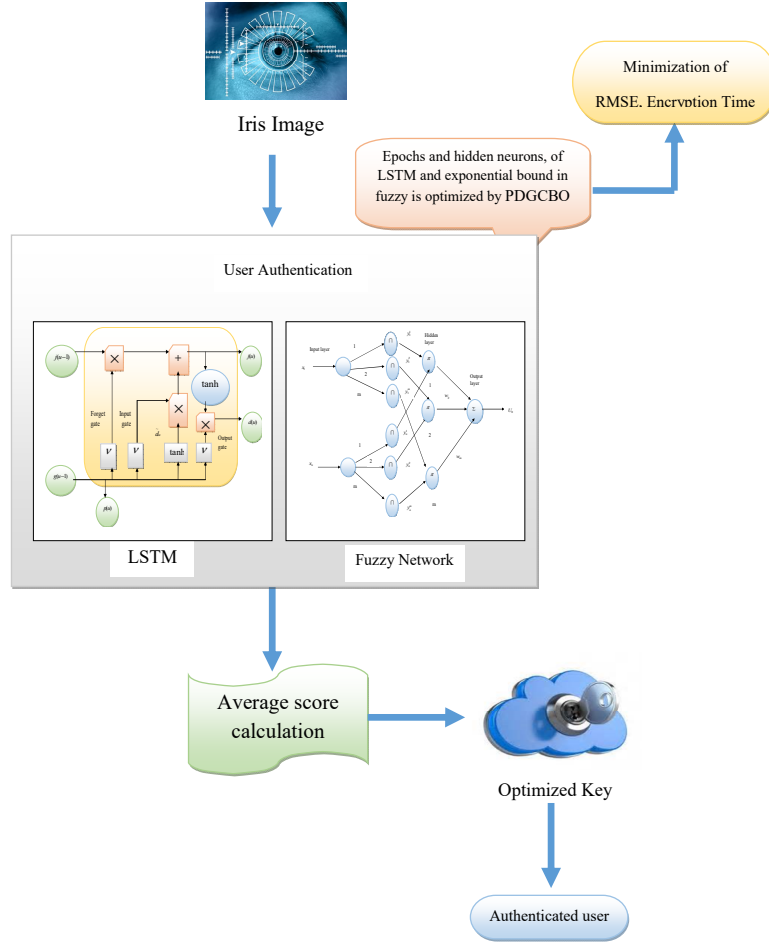
where  $u'_k$  and  $w_k^{\text{new}}$  are the velocity after the collision and the new position of the  $k^{\text{th}}$  motion CB accordingly. The old place of the CB pair is represented as  $k - \frac{m}{2}$ . The new place of the CB pair is given in equation (28).

$$w_k^{\text{new}} = w_k + \text{ran} \cdot u'_k, k = 1, \dots, \frac{m}{2} \quad (28)$$

where the new places are indicated as  $w_k^{\text{new}}$ ,  $w_k$  and  $u'_k$ .

Until the termination criteria, the optimisation is iterated from equation (26).



**Figure 3** The pictorial representation of the optimised LSTM with fuzzy for user authentication (see online version for colours)**Algorithm 1** PDGCBO

---

Take the overall population and looping count  
 Find out the fitness function of every search agent  
 The probability  $Q$  is estimated newly using equation (13)  
 Do while  
 If  $Q > 0.5$   
     DGO approach  
     Find the fitness of all search agents.  
     Update the  $G_{best}$ ,  $Y_{best}$ ,  $G_{worst}$  and  $Y_{worst}$  using equations (8) to (12).  
     Update the  $Q_j$ , of equation (13)  
      $j = j + 1$   
 Else  
     CBO approach  
     To get the value  $U'_k$  using equation (25).  
     Evaluate the new places  $w_k^{new}$  using equation (27)  
      $k = 1$   
     End  
 Iterate every mentioned steps until reach the highest count  
 Returns the optimal results

---

**4.2 Optimised LSTM with fuzzy-based user authentication**

Even though LSTM with fuzzy based user authentication is used in several ways, it has disadvantages also. LSTM takes a long time to train the data. To train the data LSTM requires more amount of memory. There is a high possibility to overfit. Dropout function is much harder to develop in LSTM. Also, the fuzzy network is very slow when it comes to running time. The results are always inaccurate. To overcome that, proposed a new optimised LSTM with fuzzy-based user authentication. The optimisation parameters are hidden neurons in LSTM, epochs in LSTM, exponential bound in fuzzy, and the private key in ECC encryption. The 'objective function'  $OF$  is expressed in equation (29).

$$OF = \arg \min_{\{hn^{lstm}, ep^{lstm}, eb^{fuzzy}, pk^{ecc}\}} [RMSE, Enctime] \quad (29)$$

The hidden neuron in LSTM is denoted as  $hn^{lstm}$ . The epochs in LSTM are indicated as  $ep^{lstm}$ . The  $eb^{fuzzy}$  is represented for exp. bound in fuzzy and the private key in fuzzy is denoted as  $pk^{ecc}$ .

The hidden neurons in LSTM range from 5 to 255. The epochs in LSTM are limited from 50 to 100. The exponential bound in fuzzy ranges from 0.01 to 0.99.

Finally, the ECC encryption private key is limited from 0 to 9. The population size is 10 for the optimised user authentication.

Further, the term *RMSE* refers to the ‘root mean square error (RMSE)’ which is ‘The mean distance between the predicted values from the model and true value of the dataset’. It is given in the equation (30).

$$RMSE = \sqrt{\sum (Q_j - S_j)^2 / n} \quad (30)$$

The term  $Q_j$  denoted as an expected value of the  $j^{\text{th}}$  monitoring in the dataset.  $S_j$  is pointed out by the observed value of the  $j^{\text{th}}$  monitoring in the dataset. And  $n$  is the sample size.

The word *Enctime* refers the encryption time that means, the time taken to perform the encryption. The diagrammatic presentation of the optimised LSTM with fuzzy model for user authentication is visualised in Figure 3.

The process of authentication using optimised LSTM with fuzzy is explained here. The iris image is denoted as  $I$ . The iris image is given to the input for the LSTM with the fuzzy network. The LSTM with the fuzzy network is already trained to find the optimised key. Here, the optimised parameters are used while the iris image is processed with the LSTM with the fuzzy network. The final optimised key  $K$  for authenticated users is gained from calculating the average of the gained iris data.

## 5 Secured big data storage in the cloud sector using cascaded modified elliptic-curve cryptography with homomorphic encryption

### 5.1 ECC model

ECC (Pan et al., 2017) is utilised to develop the public key cryptography. It is much faster and more efficient. The elliptic curve which is represented by  $F/G_r$ . Here,  $F$  denoted as the elliptic curve and  $r$  represented as the prime. The finite field of the prime is represented as  $G_r$ , which is given as:  $x^2 = (y^3 + by + c) \bmod r$ ,  $b, c \in Z_r^*$ . The five basic operations of ECC are explained below:

- 1 The addition of two points in the EC: consider the two points called  $B$  and  $C$ , on the elliptic curve, so the sum of these two points will be  $B + C = D$ , the two lines  $B$  and  $C$  joins at the curve  $-D$ , that is in the  $x$ -axis reflection of  $D$ .
- 2 Subtraction of two points in the EC: assume the two points are called  $B$  and  $C$ , on the elliptic curve, so the curve of these two points will be  $B = -C$ , i.e.,  $B + C = B + (-B) = 0$ . The two points  $B$  and  $C$  intersects at the abstract point 0, which is named as the point of infinity.
- 3 Doubling of elliptic curve point: the point  $B$  is added by itself which gives the new point called  $C$  so that makes  $C = 2B$ , that is the reflection of the intersected point  $B$  with tan respect to the  $x$ -axis drawn at the point.

- 4 Scalar point multiplication in the elliptic curve: an elliptic curve such as  $q.B = B + B + \dots + B$  ( $q$ -times)  $= \sum_1^q B$ , where,  $q \in Z_q^*$  is a scalar value in the point  $B$ .
- 5 Point order: the element order  $B$  in  $H_r$  is explained as  $m$ , where  $m > 0$  is an integer so that  $m.q = 0$ .

### 5.2 Homomorphic encryption model

HE (Cheon and Kim, 2015) is one of the encryption methods which allow users to perform computational tasks on its data which is encrypted without decrypting the data first. The final results of the computation are in encrypted form after decrypted the result will be the output identically which generate the operations that have been performed on the data unencrypted. Assume  $N$  indicates the set of plain texts. An encryption model is called as HE if for any given encryption key  $l$  and the encryption function  $F$  which explains in equation (31):

$$\forall n1, n2 \in N, F(n1 * n2) = F(n1) * F(n2) \quad (31)$$

If we assume addition operators then it is called as additively HE, or if we assume multiplication operators it is expressed as multiplicatively HE. The ring or algebraic homomorphisms are explained in equation (32).

$$\begin{aligned} \forall n1, n2 \in N, F(n1 + Nn2) &\leftarrow F(n1) \\ + DF(n2), F(n1 \times Nn2) &\leftarrow F(n1) \times DF(n2) \end{aligned} \quad (32)$$

This means that for fixed keys  $l$ , it is similar to performing functions on the plain texts before performing encryption, or after performing encryption for the particular cipher texts.

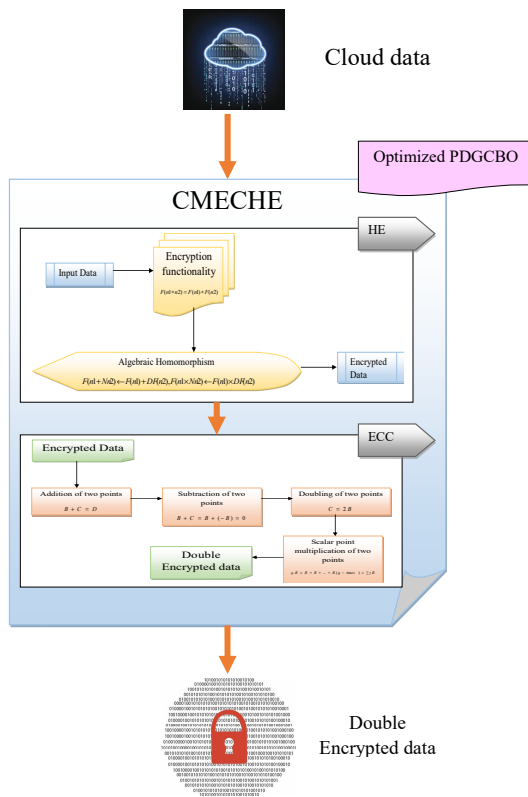
### 5.3 Cascaded modified elliptic-curve cryptography with homomorphic encryption

In this method, the data is send to the HE model. Here the data is encrypted. To prevent unauthorised access, the data will be converted into some sort of code which is known as encryption. It gives authentication and integrity. The encrypted data is then given to the input for the ECC method. Here also the encrypted data is again encrypted which will strengthen the authentication of the information. The output of the ECC method was optimised with the help of PDGCBO for gaining the optimal data. This is called Cascaded HE with ECC. Because of the cascading technology, both techniques can influence their characters. It concatenates the classifiers utilising all the data gathered from the output. Cascading utilises the efficient usage of resources by using data to extend the total biomass availability of the given model. The diagrammatic representation of the CMECHE in the cloud is shown in Figure 4.

The encrypted data is saved in the cloud storage which helps to store the data by transferring it over the internet. To access the data, it should contain the authenticated user key. This authenticated user key is obtained from the optimised

LSTM with the fuzzy technique where the data is trained to generate the authenticated key. The LSTM with fuzzy utilising the optimisation parameters while generating the authenticated user key. After gaining the authenticated user key the cloud data will be decrypted. Decryption means transferring the data that is already encrypted which is in the format of unreadable into readable, i.e., the raw data. The process of decryption that occurs in the proposed model is explained here. First, the double time encrypted data is sent to the ECC decryption. Here, the doubly encrypted data is decrypted once. Then, it is sent through the HE decryption model where again the decryption process is carried out. Finally, the encrypted data that is the original data is obtained.

**Figure 4** The diagrammatic representation of the CMECHE in the cloud (see online version for colours)



## 6 Results and discussion

### 6.1 Experimental setup

This suggested secured data storage in the cloud was run in the Python platform and particular outcomes were taken. The suggested algorithm has 25 highest looping counts and 10 as the size of the population. Distinct measurements were utilised to evaluate the execution. So, the comparison algorithms such as SFO (Gomes et al., 2019), DHOA (Brammya et al., 2019), DGO (Dehghani et al., 2020) and CBO (Kaveh and Mahdavi, 2014) were taken.

Consecutively, old classifier models were RSA (Koc et al., 2021), ECC (Pan et al., 2017), HE (Cheon and Kim, 2015), and ECC\_HE (Pan et al., 2017; Cheon and Kim, 2015) respectively.

### 6.2 Performance measures

- **CPA:** in CPA, it is attacker selecting the random plaintext to be encrypted and obtains the corresponding plaintext.
- **MD:** it denotes the ‘highest difference’ value which is evaluated using equation (33).

$$md = \left\lceil \frac{100}{len(J)} \right\rceil * sum(v) \quad (33)$$

Here, the word could be evaluated in equation (34).

$$v = abs \left[ \frac{K - J}{J} \right] \quad (34)$$

In equation (33) and equation (34), the variable  $J$  and  $K$  denotes the true and assumed measurement of images, and the absolute value is expressed by  $abs$ .

- **SMAPE:** SMAPE is a correctness factor based on % errors.

$$SMAPE = \frac{100\%}{W} \sum_{b=1}^B \left( \frac{j_b - k_b}{j_b + k_b / 2} \right) \quad (35)$$

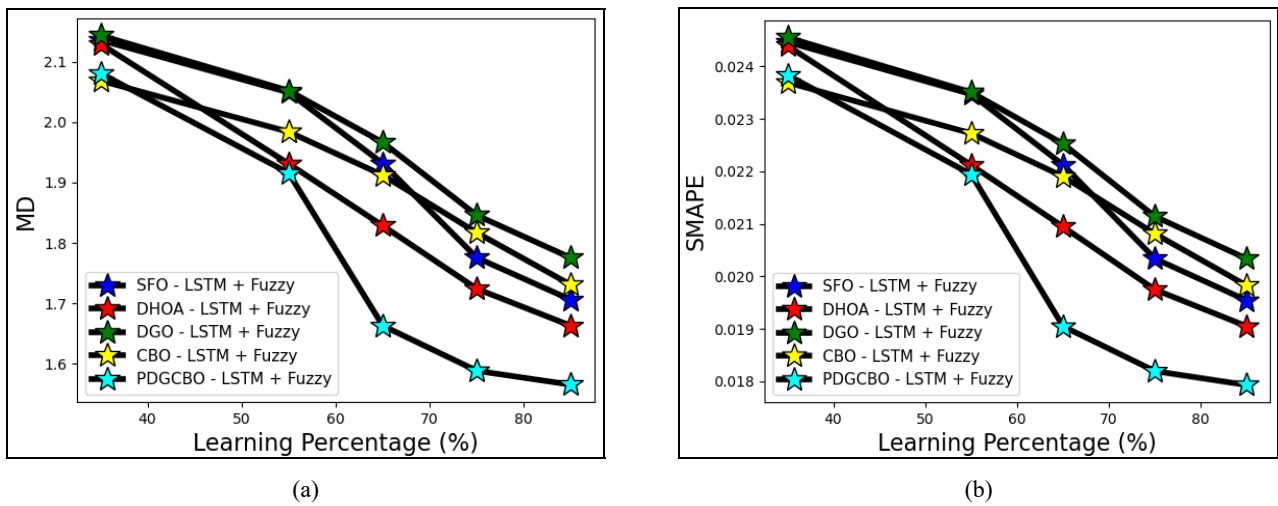
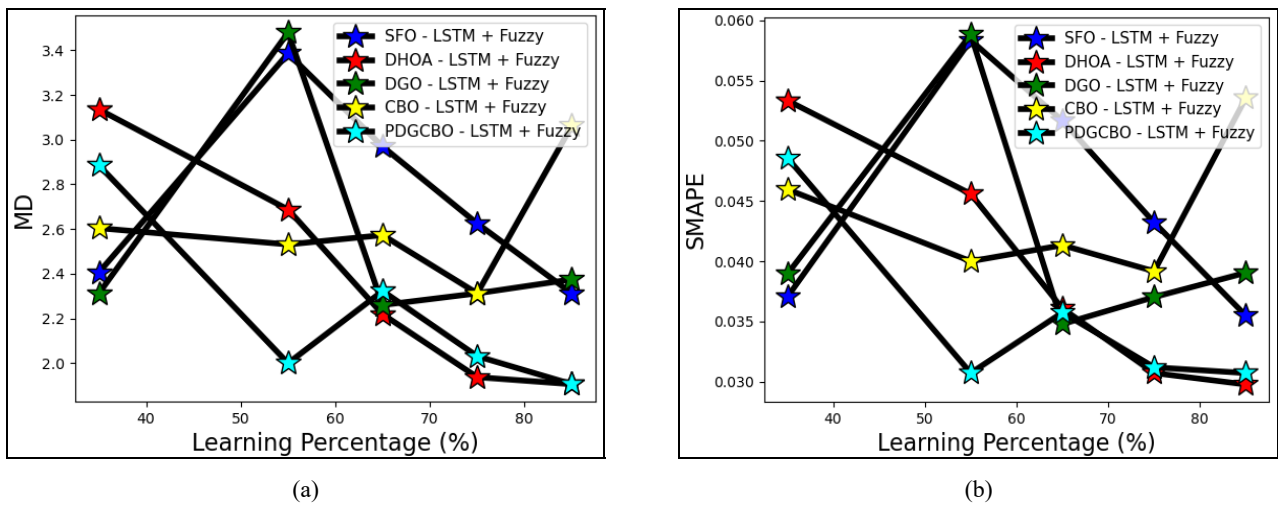
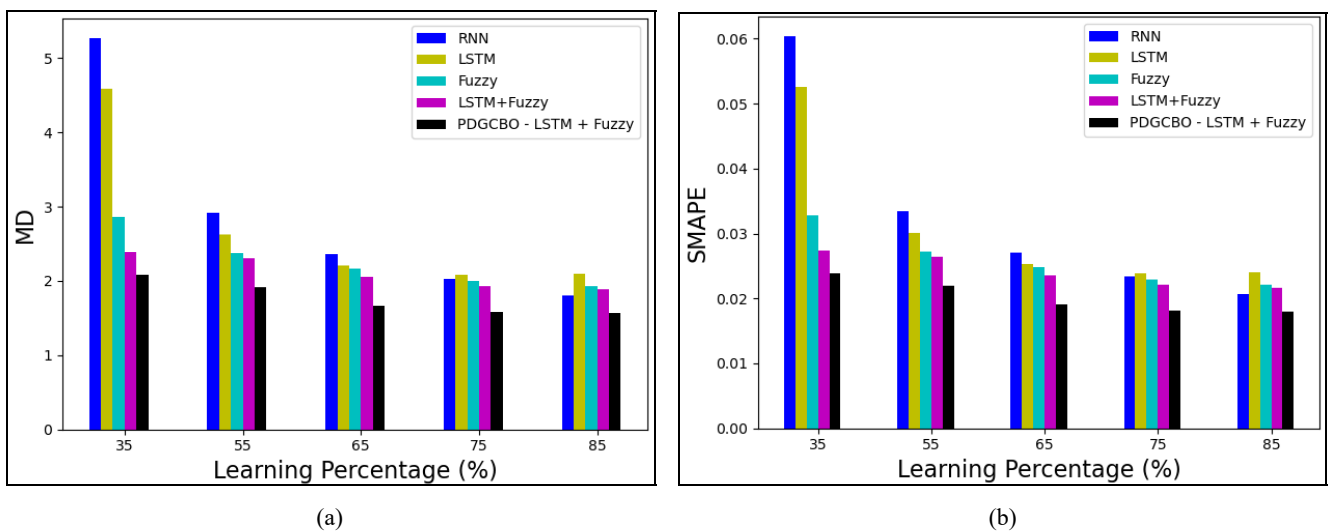
- **KPA:** it is an attack method for cryptanalysis where the attacker has access to both the plaintext and its encrypted version.
- **Data 1 and data 2:** for evaluating the proposed secured big data in the cloud here taken two medical data called data 1 and data 2. Using these two data performance of the suggested model was established.

### 6.3 Performance analysis of the proposed secured big data in the cloud for data 1 and 2

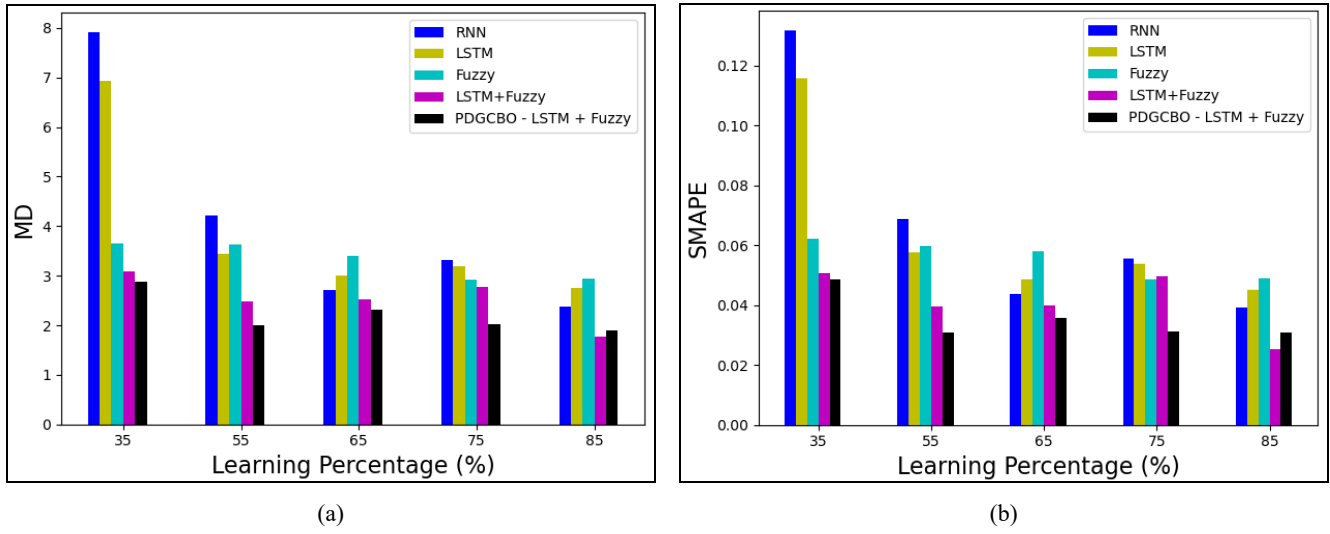
Figures 5 and 6 displays the performance evaluation of the suggested method in compared with distinct algorithms. Figures 7 and 8 displays the performance analysis of the suggested method in compare with classifier models.

### 6.4 Attack analysis of the proposed secured big data in the cloud for data 1 and 2

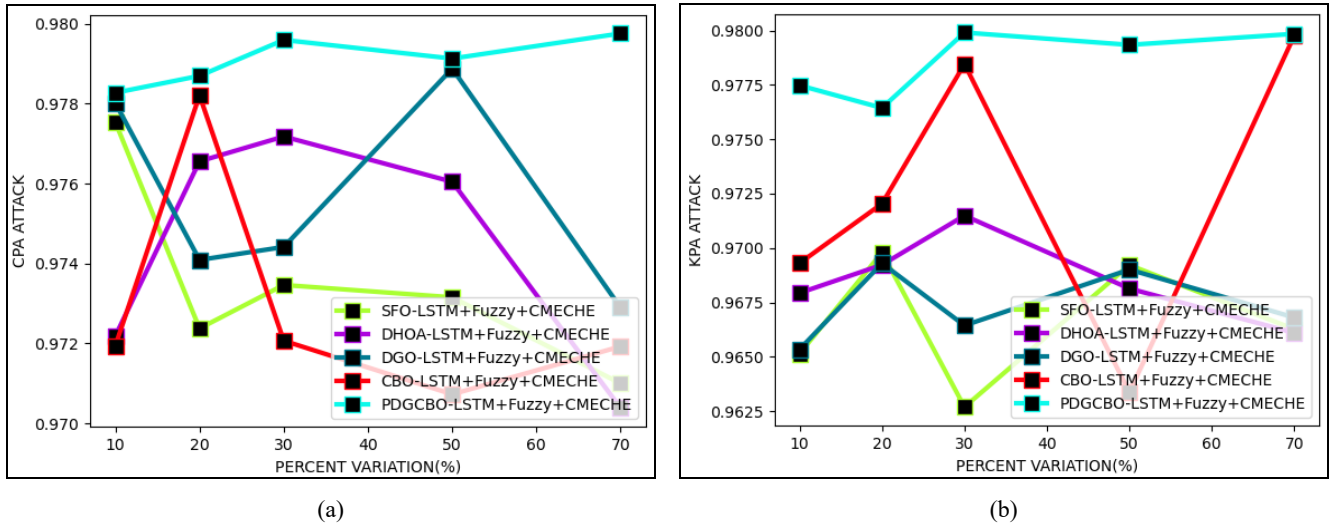
Figures 9 and 10 displays the attack evaluation of the suggested model in compared with distinct algorithms. Figures 11 and 12 displays the CPA attack analysis for the data 1 and 2.

**Figure 5** Performance analysis of novel authentication model to secure the big data in cloud for data 1 compared with traditional algorithms in terms of ‘(a) MD, (b) SMAPE’ (see online version for colours)**Figure 6** Performance analysis of novel authentication model to secure the big data in cloud for data 2 compared with traditional algorithms in terms of ‘(a) MD, (b) SMAPE’ (see online version for colours)**Figure 7** Performance analysis of novel authentication model to secure the big data in cloud for data 1 compared with conventional classifier models in terms of ‘(a) MD, (b) SMAPE’ (see online version for colours)

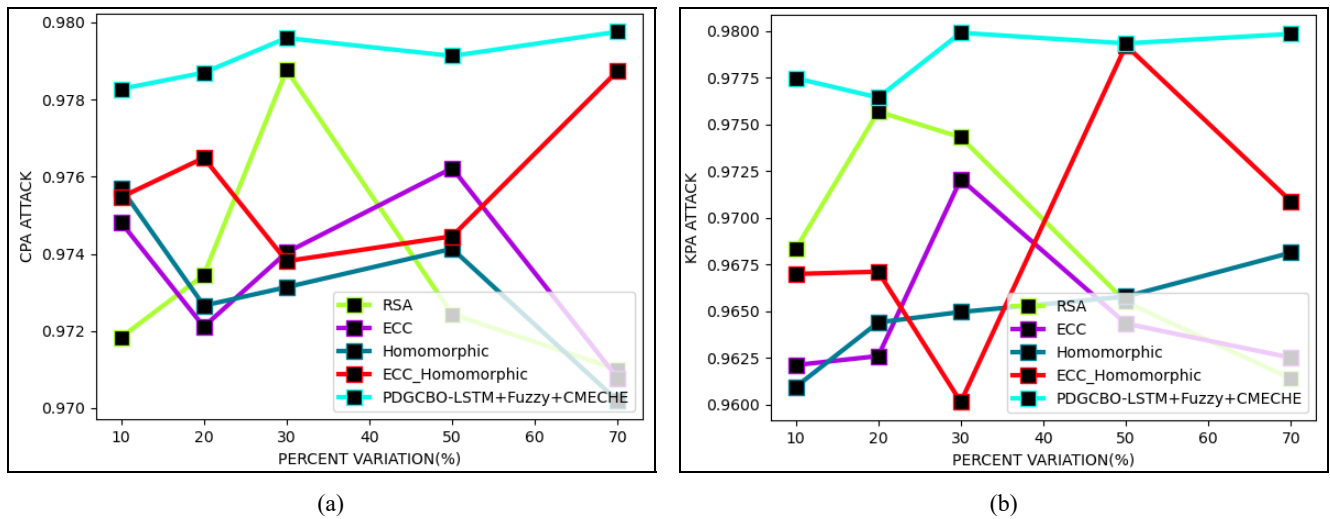
**Figure 8** Performance analysis of novel authentication model to secure the big data in cloud for data 2 compared with conventional classifier models in terms of ‘(a) MD, (b) SMAPE 2’ (see online version for colours)

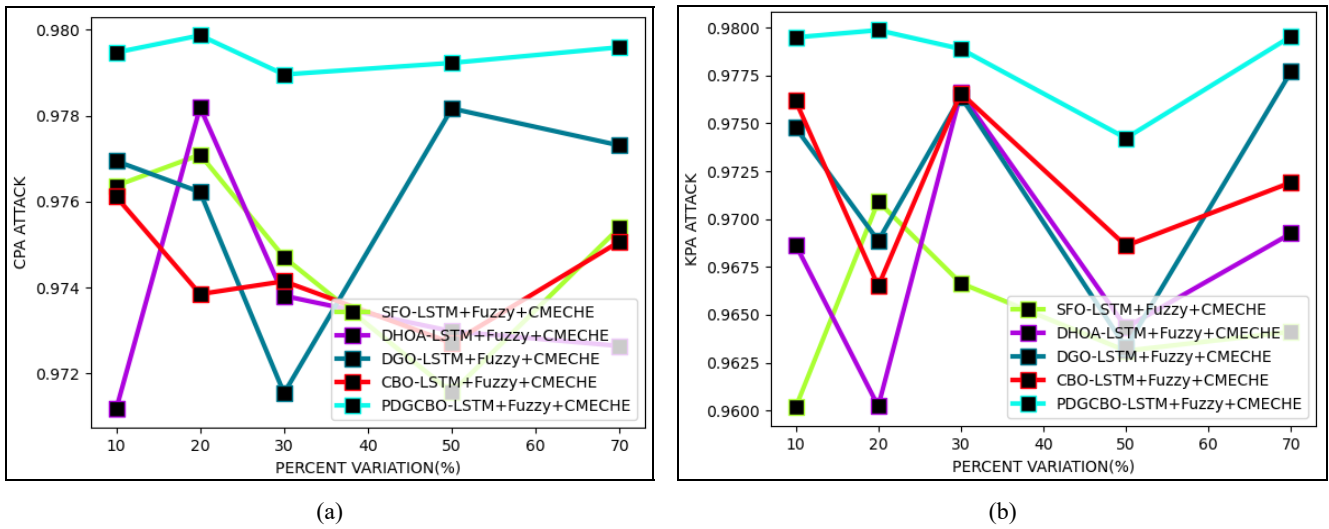
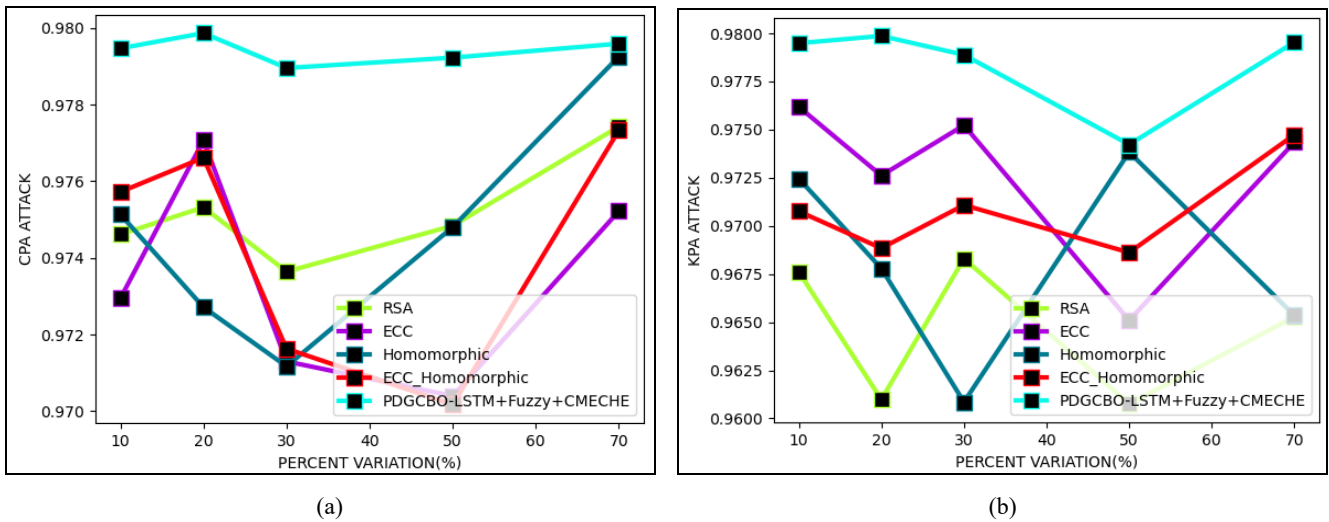
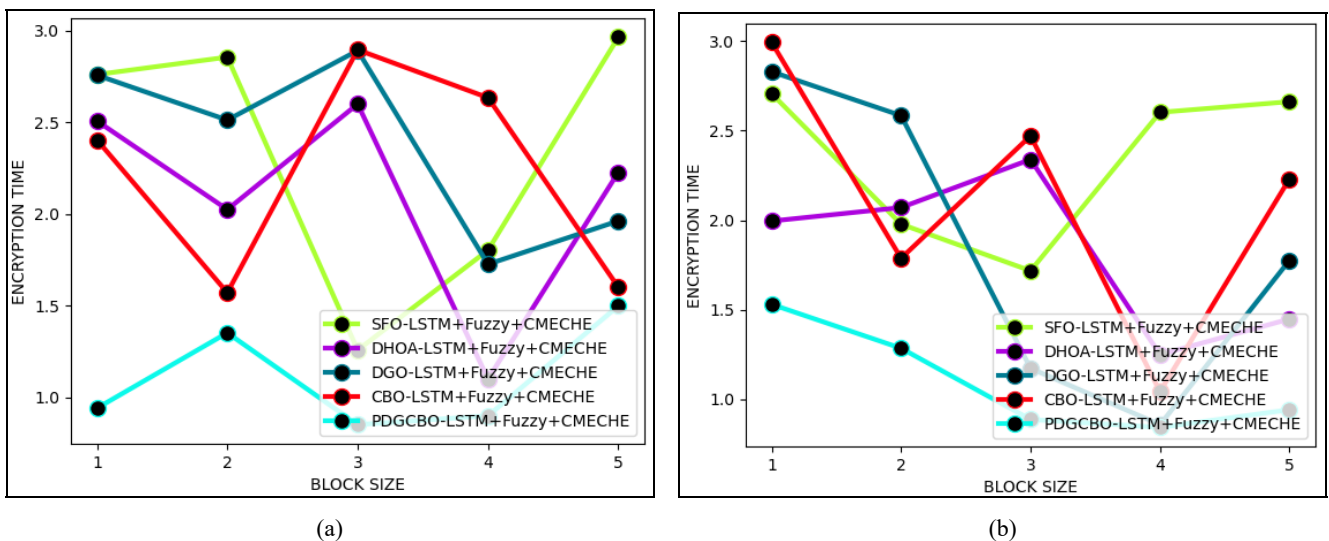


**Figure 9** Attack analysis of the privacy-preservation methodology for securing the big data in the cloud for data 1 compared with traditional algorithms regarding ‘(a) CPA attack, (b) KPA attack’ (see online version for colours)



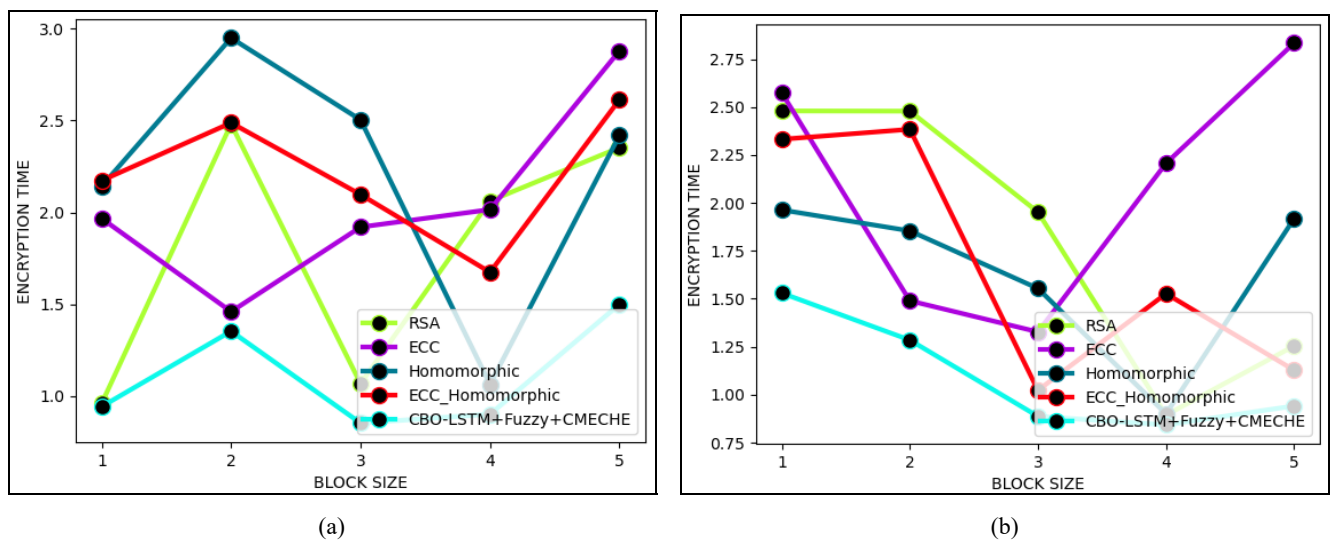
**Figure 10** Attack evaluation of the privacy-preservation model for securing the big data in the cloud for data 2 comparison with traditional algorithms in terms of ‘(a) CPA attack, (b) KPA attack’ (see online version for colours)



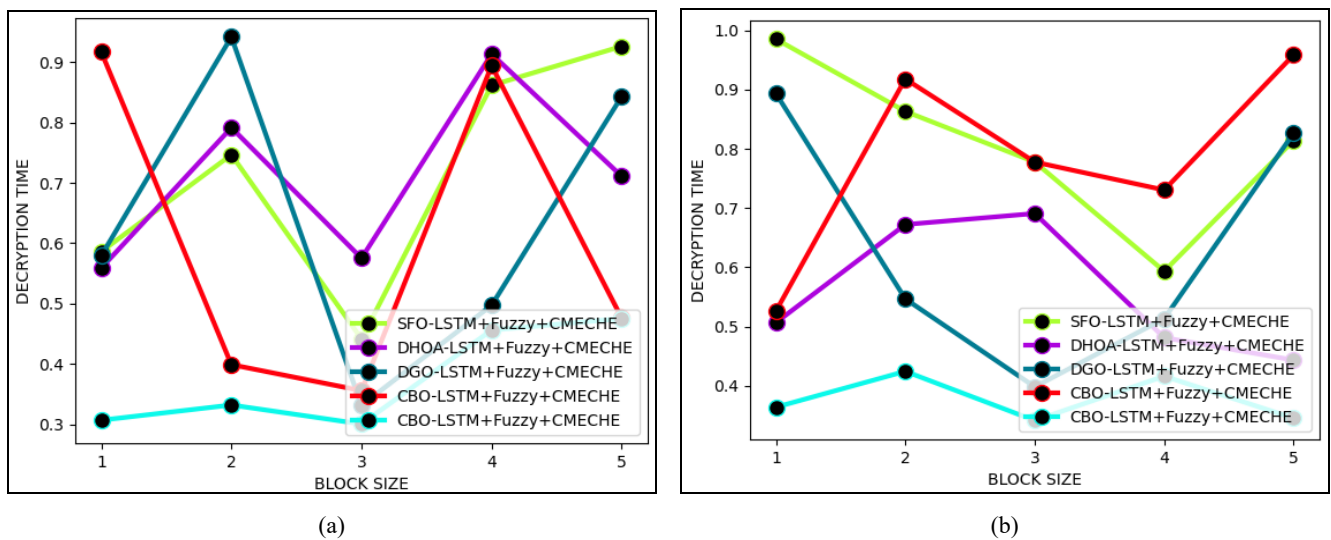
**Figure 11** Attack evaluation of secured big data in the cloud for data 1 comparison with traditional classifier models regarding ‘(a) CPA attack, (b) KPA attack’ (see online version for colours)**Figure 12** Attack analysis of the privacy-preservation methodology for securing the big data in the cloud for data 2 compared with traditional classifier models in terms of ‘(a) CPA attack, (b) KPA attack’ (see online version for colours)**Figure 13** Encryption time of privacy-preservation methodology for securing the big data in the cloud comparison with traditional algorithms in terms of ‘(a) Algorithm 1, (b) Algorithm 2’ (see online version for colours)



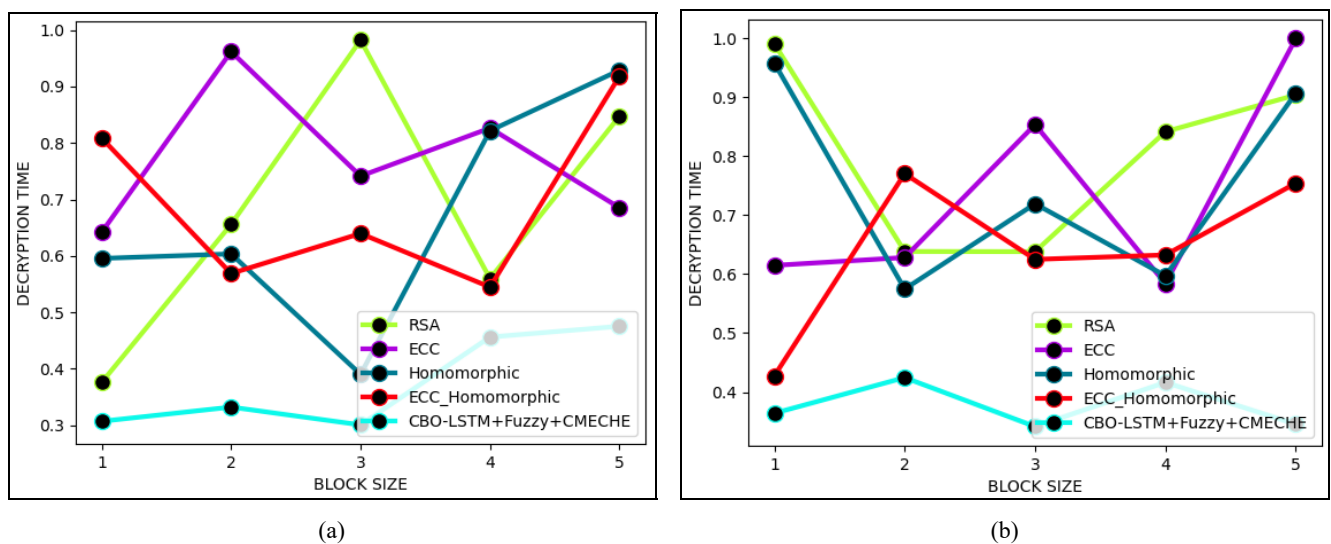
**Figure 14** Encryption time of privacy-preservation methodology for securing the big data in the cloud compared with traditional classifier models in terms of 'a) method 1, (b) method 2' (see online version for colours)

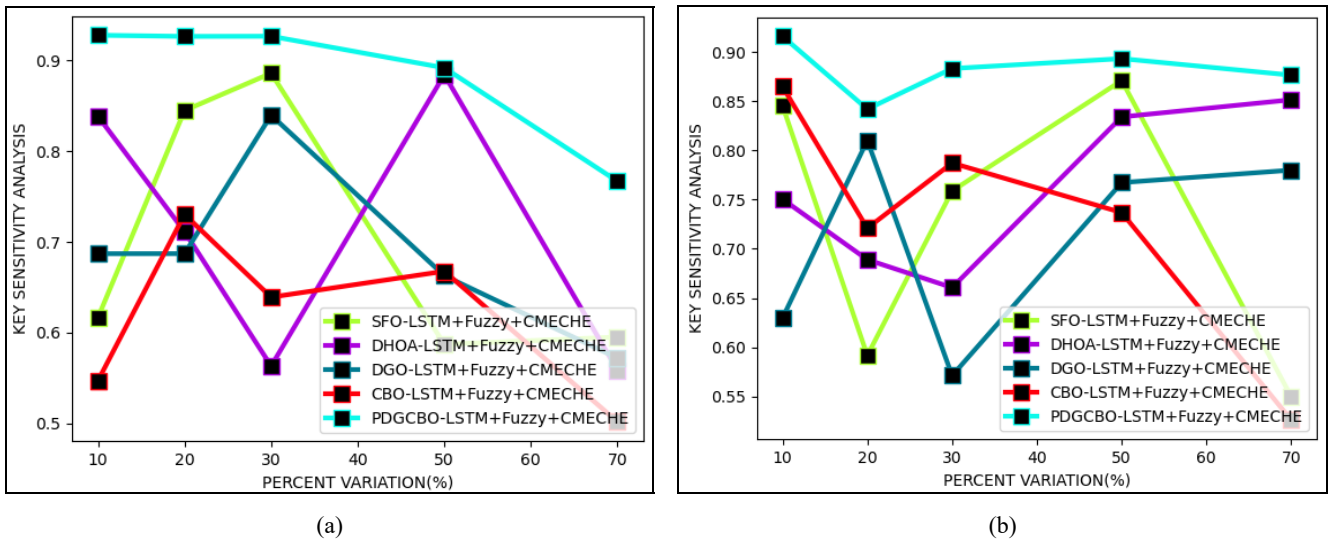
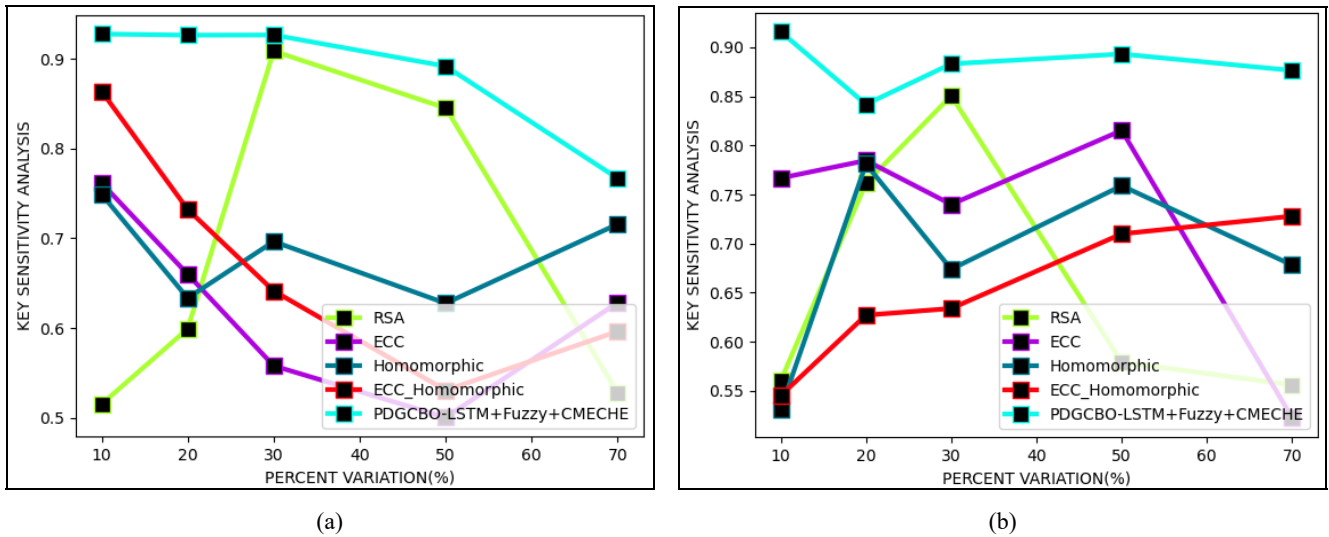


**Figure 15** Decryption time of privacy-preservation method for securing the big data in the cloud comparison with traditional algorithms in terms of 'a) Algorithm 1, (b) Algorithm 2' (see online version for colours)



**Figure 16** Decryption time of privacy-preservation model for securing the big data in the cloud compared with traditional classifier models regarding 'a) method 1, (b) method 2' (see online version for colours)



**Figure 17** Optimised key evaluation of the privacy-preservation model for securing the big data in the cloud compared with traditional algorithms regarding ‘(a) Algorithm 1, (b) Algorithm 2’ (see online version for colours)**Figure 18** Optimised key evaluation of the privacy-preservation model for securing the big data in the cloud compared with traditional classifier models regarding ‘(a) method 1, (b) method 2’ (see online version for colours)**Table 1** Overall comparison analysis of the suggested secured bigdata storage in the cloud over different algorithms for two datasets

Terms	SFO-LSTM+Fuzzy (Gomes et al., 2019)	DHOA-LSTM+Fuzzy (Brammya et al., 2019)	DGO-LSTM +Fuzzy (Dehghani et al., 2020)	CBO-LSTM +Fuzzy (Kaveh and Mahdavi, 2014)	PDGCBO- LSTM +Fuzzy
<i>Dataset I</i>					
‘MD’	1.70582	1.662824	1.775213	1.730516	1.565421
‘SMAPE’	0.01954	0.019048	0.020335	0.019822	0.017931
‘MASE’	6,904.012	6,696.312	7,131.891	6,987.962	6,351.661
‘MAE’	1.875171	1.827454	1.956338	1.91017	1.720339
‘RMSE’	7.653598	7.548873	7.822574	7.738637	7.324377
‘ONE-NORM’	135,498	132,050	141,363	138,027	124,310
‘TWO-NORM’	2,057.366	2,029.215	2,102.789	2,080.226	1,968.868
‘INFINITY-NORM’	61	58	66	59	61
<i>Dataset II</i>					
‘MD’	2.3125	1.90625	2.375	3.0625	1.90625
‘SMAPE’	0.035476	0.029762	0.039048	0.053571	0.030714



**Table 1** Overall comparison analysis of the suggested secured bigdata storage in the cloud over different algorithms for two datasets (continued)

Terms	<i>SFO-LSTM+Fuzzy</i> (Gomes et al., 2019)	<i>DHOA-LSTM+Fuzzy</i> (Brammya et al., 2019)	<i>DGO-LSTM</i> +Fuzzy (Dehghani et al., 2020)	<i>CBO-LSTM +Fuzzy</i> (Kaveh and Mahdavi, 2014)	<i>PDGCBO-LSTM +Fuzzy</i>
<i>Dataset II</i>					
‘MASE’	29.84127	22.69886	27.71098	28.32276	21.02632
‘MAE’	0.05	0.0375	0.045	0.0475	0.035
‘RMSE’	0.223607	0.193649	0.212132	0.217945	0.187083
‘ONE-NORM’	40	30	36	38	28
‘TWO-NORM’	6.324555	5.477226	6	6.164414	5.291503
‘INFINITY-NORM’	1	1	1	1	1

**Table 2** Overall comparison evaluation of the recommended secured bigdata storage in the cloud over diverse classifiers for two datasets

Terms	<i>RSA (Koc et al., 2021)</i>	<i>ECC (Pan et al., 2017)</i>	<i>HOMOMORPHIC</i> (Cheon and Kim, 2015)	<i>ECC HOMOMORPHIC</i> (Pan et al., 2017; Cheon and Kim, 2015)	<i>PDGCBO-LSTM</i> +FUZZY
<i>Dataset I</i>					
‘MD’	1.800697	2.103433	1.929672	1.88394	1.565421
‘SMAPE’	0.020625	0.024095	0.022104	0.02158	0.017931
‘MASE’	7,300.849	8,363.649	7,747.638	7,542.65	6,351.661
‘MAE’	1.999239	2.319061	2.132869	2.070012	1.720339
‘RMSE’	7.928132	8.509641	8.173943	8.037252	7.324377
‘ONE-NORM’	144,463	167,573	154,119	149,577	124,310
‘TWO-NORM’	2,131.164	2,287.48	2,197.24	2,160.496	1,968.868
‘INFINITY-NORM’	59	58	64	58	61
<i>Dataset II</i>					
‘MD’	2.385417	2.75	2.9375	1.760417	1.90625
‘SMAPE’	0.03919	0.045238	0.049048	0.025143	0.030714
‘MASE’	27.44656	33.13478	31.07222	28.42884	21.02632
‘MAE’	0.045	0.055	0.0525	0.0475	0.035
‘RMSE’	0.212132	0.234521	0.229129	0.217945	0.187083
‘ONE-NORM’	36	44	42	38	28
‘TWO-NORM’	6	6.63325	6.480741	6.164414	5.291503
‘INFINITY-NORM’	1	1	1	1	1

### 6.5 Encryption time of the proposed secured big data in the cloud for data 1 and 2

Figures 13 and 14 display the time of the encryption analysis of the suggested method in compared with distinct algorithms and classifier models.

### 6.6 Decryption time of the proposed secured big data in the cloud for data 1 and 2

Figures 15 and 16 display the decryption time of the evaluation recommended method in contrasted with distinct algorithms and classifier models.

### 6.7 Key sensitivity analysis of the proposed secured big data in the cloud for data 1 and 2

Figures 17 and 18 display the key sensitivity evaluation of the recommended method in comparison with distinct algorithms and classifier models.

### 6.8 Overall comparative analysis of the proposed model over algorithms and classifiers

The last estimation of the proposed method is compared with the conventional methodologies of different algorithms and distinct classifiers that are shown in Table 1 and Table 2.

## 7 Conclusions

This research work has implemented a new secure authentication and preservation of privacy in big data over the cloud. It assured the safety of big data using deep learning techniques. At first, the biometric information was utilised for providing secure authentication to the cloud data for neglecting the malicious allowances. The secure key was extracted through encryption, where the alteration was done by the same DG-CBO algorithm. At last, the final outcomes showed that the effectiveness of the suggested algorithms in this security method was confirmed as optimised by LSTM with fuzzy network, where a new PDGCBO algorithm was implemented for tuning the parameters in LSTM and fuzzy network. The extracted secured key was utilised for assuring the privacy of data cloud via the CMECHE better than other conventional algorithms regarding security and performance. Figure 18(a) shows the key sensitivity evaluation of the suggested method in compared with the conventional classifier method for data 2. When it has considered the 20% for the key sensitivity, the recommended model has raised by 97% of RSA, 96.65% of ECC, 96.8% HE, and 96.3% of ECC\_HE accordingly. Therefore, the implementation of the model was estimated utilising distinct factors and contrasted with other conventional methods. The outcomes have explained that it has improved the security of the data access process.

## References

- Brammya, G., Praveena, S., Ninu Preetha, N.S., Ramya, R., Rajakumar, B.R. and Binu, D. (2019) *Deer Hunting Optimization Algorithm: A New Nature-Inspired Meta-heuristic Paradigm*, 24 May.
- Chen, F., Meng, F., Xiang, T., Dai, H., Li, J. and Qin, J. (2020) 'Towards usable cloud storage auditing', *IEEE Transactions on Parallel and Distributed Systems*, 1 November, Vol. 31, No. 11, pp.2605–2617.
- Cheon, J.H. and Kim, J. (2015) 'A hybrid scheme of public-key encryption and somewhat homomorphic encryption', *Transactions on Information Forensics and Security*, May, Vol. 10, No. 5, pp.1052–1063.
- Cui, H., Deng, R.H., Li, Y. and Wu, G. (2019) 'Attribute-based storage supporting secure deduplication of encrypted data in cloud', *IEEE Transactions on Big Data*, 1 September, Vol. 5, No. 3, pp.330–342.
- Dehghani, M., Montazeri, Z., Givi, H., Guerrero, J.M. and Dhiman, G. (2020) 'Darts game optimizer: a new optimization technique based on darts game', *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 5, pp.286–294.
- Deng, L., Yang, B. and Wang, X. (2020) 'A lightweight identity-based remote data auditing scheme for cloud storage', *IEEE Access*, Vol. 8, pp.206396–206405.
- Gomes, G.F., da Cunha Jr., S.S. and Ancelotti Jr., A.C. (2019) 'A sunflower optimization (SFO) algorithm applied to damage identification on laminated composite plates', *Engineering with Computers*, Vol. 35, pp.619–626.
- Hu, C., Li, W., Cheng, X., Yu, J., Wang, S. and Bie, R. (2018) 'A secure and verifiable access control scheme for big data storage in clouds', *IEEE Transactions on Big Data*, 1 September, Vol. 4, No. 3, pp.341–355.
- Kaveh, A. and Mahdavi, V.R. (2014) 'Colliding bodies optimization: a novel meta-heuristic method', *Computers & Structures*, July, Vol. 139, pp.18–27.
- Khan, A.Q., Nikolov, N., Matskin, M., Prodan, R., Roman, D., Sahin, B., Bussler, C. and Soylu, A. (2023) 'Smart data placement using storage-as-a-service model for big data pipelines', *Sensors*, Vol. 23, No. 2, p.564.
- Koc, C.K., Ozdemir, F. and Ozger, Z.O. (2021) 'Rivest-Shamir-Adleman algorithm', *Partially Homomorphic Encryption*, pp.37–41.
- Li, Y., Gai, K., Qiu, L., Qiu, M. and Zhao, H. (2017) 'Intelligent cryptography approach for secure distributed big data storage in cloud computing', *Information Sciences*, Vol. 387, pp.103–115.
- Mendes, R., Oliveira, T., Cogo, V., Neves, N. and Bessani, A. (2021) 'Charon: a secure cloud-of-clouds system for storing and sharing big data', *IEEE Transactions on Cloud Computing*, 1 October–December, Vol. 9, No. 4, pp.1349–1361.
- Narayanan, U., Paul, V. and Joseph, S. (2020) 'A novel system architecture for secure authentication and data sharing in cloud-enabled big data environment', *Journal of King Saud University – Computer and Information Sciences*, No. 6, pp.3121–3135.
- Pan, W., Zheng, F., Zhao, Y., Zhu, W.-T. and Jing, J. (2017) 'An efficient elliptic curve cryptography signature server with GPU acceleration', *Transactions on Information Forensics and Security*, January, Vol. 12, No. 1, pp.111–122.
- Prabhu Kavin, B., Ganapathy, S., Kanimozhi, U. and Kannan, A. (2020) 'An enhanced security framework for secured data storage and communications in cloud using ECC, access control and LDSA', *Wireless Personal Communications*, Vol. 115, pp.1107–1135.
- Senthilnathan, T., Prabu, P., Sivakumar, R. and Sakthivel, S. (2018) 'An enhancing reversible data hiding for secured data using shuffle block key encryption and histogram bit shifting in cloud environment', *Cluster Computing*, Vol. 22, pp.12839–12847.
- Sherstinsky, A. (2020) 'Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network', *Special Issue on Machine Learning and Dynamical Systems*, March, Vol. 404, p.132306.
- Sun, C.-T. and Jang, J.-S. (2018) 'A neuro-fuzzy classifier and its applications', *Department of Computer and Information Science*, Vol. 56, pp.102–105.
- Tao, Y., Xu, P. and Jin, H. (2020) 'Secure data sharing and search for cloud-edge-collaborative storage', *IEEE Access*, Vol. 8, pp.15963–15972.
- Tchernykh, A., Miranda-López, V., Babenko, M., Armenta-Cano, F., Radchenko, G., Drozdov, A.Y. and Avetisyan, A. (2019) 'Performance evaluation of secret sharing schemes with data recovery in secured and reliable heterogeneous multi-cloud storage', *Cluster Computing*, Vol. 22, pp.1173–1185.
- Wang, H., Wang, Q. and He, D. (2021) 'Blockchain-based private provable data possession', *IEEE Transactions on Dependable and Secure Computing*, 1 September–October, Vol. 18, pp.2379–2389.

- Xu, X., Zhou, J., Wang, X. and Zhang, Y. (2016) 'Multi-authority proxy re-encryption based on CPABE for cloud storage systems', *Journal of Systems Engineering and Electronics*, February, Vol. 27, No. 1, pp.211–223.
- Yang, X., Lu, R., Choo, K.K.R., Yin, F. and Tang, X. (2022) 'Achieving efficient and privacy-preserving cross-domain big data deduplication in cloud', *IEEE Transactions on Big Data*, 1 February, Vol. 8, No. 1, pp.73–84.
- Yang, X., Pei, X., Wang, M., Li, T. and Wang, C. (2020) 'Multi-replica and multi-cloud data public audit scheme based on blockchain', *IEEE Access*, Vol. 8, pp.144809–144822.
- Yang, Y., Zheng, X., Rong, C. and Guo, W. (2020) 'Efficient regular language search for secure cloud storage', *IEEE Transactions on Cloud Computing*, 1 July–September, Vol. 8, No. 3, pp.805–818.
- Zaghloul, E., Zhou, K. and Ren, J. (2020) 'P-MOD: secure privilege-based multilevel organizational data-sharing in cloud computing', *IEEE Transactions on Big Data*, 1 December, Vol. 6, No. 4, pp.804–815.
- Zeng, P. and Choo, K-K.R. (2018) 'A new kind of conditional proxy re-encryption for secure cloud storage', *IEEE Access*, Vol. 6, pp.70017–70024.
- Zhang, Y., Huang, H., Xiang, Y., Zhang, L.Y. and He, X. (2017) 'Harnessing the hybrid cloud for secure big image data service', *IEEE Internet of Things Journal*, October, Vol. 4, No. 5, pp.1380–1388.
- Zhang, Y., Yu, J., Hao, R., Wang, C. and Ren, K. (2020) 'Enabling efficient user revocation in identity-based cloud storage auditing for shared big data', *IEEE Transactions on Dependable and Secure Computing*, 1 May–June, Vol. 17, No. 3, pp.608–619.
- Zhang, Z. et al. (2019) 'Achieving privacy-friendly storage and secure statistics for smart meter data on outsourced clouds', *IEEE Transactions on Cloud Computing*, 1 July–September, Vol. 7, No. 3, pp.638–649.
- Zhou, L., Fu, A., Yang, G., Wang, H. and Zhang, Y. (2022) 'Efficient certificateless multi-copy integrity auditing scheme supporting data dynamics', *IEEE Transactions on Dependable and Secure Computing*, 1 March–April, Vol. 19, No. 2, pp.1118–1132.
- Zhu, H. et al. (2019) 'A secure and efficient data integrity verification scheme for cloud-IoT Based on short signature', *IEEE Access*, Vol. 7, pp.90036–90044.

## Notation

Abbreviation	Description
ABE	Attributed-based encryption
AD2	Alternative data distribution
CMECHE	Cascaded modified elliptic-curve cryptography with homomorphic encryption
CBO	Colliding bodies optimisation
CPA	Chosen plaintext attack
DGO	Darts game optimiser
ECC	Elliptic-curve cryptography
DHOA	Deer hunting optimisation algorithm
EDCon	Efficient data conflation
EPCDD	Efficient and privacy-preserving cross-domain big data deduplication
HE	Homomorphic encryption
Id-EAC	Identity-based elliptic curve access control
KPA	Known-plaintext attack
LSTM	Long short-term memory
NTRU	Number theory research unit
NN	Neural network
PDGCBO	Probability-based darts game colliding bodies optimisation
RSA	Rivest-Shamir-Adleman
SA-EDS	Security-aware efficient distributed storage
STaaS	Storage as a service
SED2	Secure efficient data distributions
SFO	Sunflower optimisation
SMAPE	Symmetric mean absolute percentage error