ORTHOGONAL REGRESSED STEEPEST DESCENT DEEP PERCEPTIVE NEURAL LEARNING FOR IOT-AWARE SECURED BIG DATA COMMUNICATION

S. L. Swapna[1*] and V. Saravanan[2]

(Received: 30-Nov.-2022, Revised: 21-Jan.-2023 and 15-Feb.-2023, Accepted: 19-Feb.-2023)

ABSTRACT

The Internet of Things (IoT) is a collection of interconnected intelligent devices that exist within the larger network known as the Internet. With the increasing popularity of IoT devices, massive data is generated day by day. The collected data needs to be continuously uploaded to the cloud server. Besides, the transmission of data in the cloud environment is performed via the Internet, which faces numerous threats. However, the security issue always lacks effective big data communication. Therefore, a novel technique called Orthogonal Regressed Steepest Descent Deep Structured Perceptive Neural Learning based Secured Data Communication (ORSDDSPNL-SDC) is introduced with higher accuracy and lesser time consumption. The ORSDDSPNL-SDC technique comprises three phases; namely, registration, user authentication, and secure data communication. In the ORSDDSPNL-SDC technique, the registration phase is carried out for creating a new ID and a password for each user in the cloud. The IoT device's data is then sent to a cloud server by the cloud user for storage. After that, the orthogonal regressed steepest descent multilayer deep perceptive neural learning is applied to examine the user_ ID with already registered ID based on Szymkiewicz–Simpson coefficient. Then, the Maxout activation function is to classify the user as authorized or unauthorized. Finally, the steepest descent function is applied for minimizing the classification error and increasing the classification accuracy. In this way, the authorized or unauthorized user is identified. Then, the secured communication is performed with the authorized cloud users. Experimental evaluation is carried out on the factors, such as classification accuracy, classification time and error rate, and space complexity for several users. The qualitative results and discussion indicate that the proposed ORSDDSPNL-SDC offers an elevated performance by achieving a higher classification accuracy and a minimum error as well as computation time when compared to the existing methods.

KEYWORDS

Big data, IoT, Multilayer deep perceptive neural learning, Orthogonal regression analysis, ***Szymkiewicz–Simpson coefficient, Maxout activation function, Steepest descent function.*

1. INTRODUCTION

The Internet of Things (IoT) has gained broad acceptance and increased in various aspects of human life. Heterogeneous big data security is a great challenge in IoT due to the rapid growth of data. Therefore, novel security approaches are needed for data communication and outsourcing to storage systems through the authorized entity. Hence, an IoT and deep learning-based secure data analytics system is proposed introduced to address this security issue.

A Homomorphic Block-Ring Security System (HBRSS) was developed in [1] for enhancing the security of data communication. HBRSS brings higher security and more comprehensive performance compared with some mainstream security systems. HBRSS is proposed for high security data transmission and data processing in public networks and mistrusted cloud environments. However, a higher security level was unable to achieve. A trusted and Collaborative Framework for Deep Learning enabled IoT was designed in [2] to enhance data transmission and computation security. Though the framework minimizes the overhead, time complexity was not reduced.

Deep learning and IoT-based data processing frameworks were introduced in [3] in various security concerns to include malicious code tracking, intrusion, privacy issues, vulnerability detection, and fault diagnosis.

In [4] a reliable lightweight authentication scheme for IoT-based secure data sharing was developed. But, the machine learning technique was not applied for improving the authentication accuracy. In [5],

1. S. L. Swapna (Corresponding Author) is with Department of Computer Science, Hindusthan College of Arts and Science (Autonomous), Coimbatore, India. Email: swapnamartin2003@gmail.com
2. V. Saravanan is with Department of Information Technology, Hindusthan College of Arts and Science (Autonomous), Coimbatore, India. Email: vsreesaran@gmail.com

a secure authentication protocol for cloud big data was introduced to reduce computing power. The storage costs were not lowered. A lightweight authentication technique was developed in [6] for ensuring device and cloud server security. But, accurate authentication was not performed with minimum time.

A lightweight identity-based authenticated data-sharing protocol was designed in [7] for enhancing the security of data distribution between physical devices and clients. But, the designed protocol failed to evaluate in a real-world setting. AI-enabled lightweight, a secure communication method for an IoMT called ASCP-IoMT was developed in [8] to improve accuracy. But, it failed to consider the functionality of features in the presented scheme.

In [9], a lightweight and secure communication technique was developed for safe data transmission among healthcare infrastructure devices. However, the registration process was completed offline. An ultra-lightweight device-to-device secure transmission protocol was introduced in [10] to protect data transmission. However, when dealing with a large number of devices, it was unable to guarantee the complete security of the IoT environment. In [11], two optimization schemes were designed to prevent user access from leaking data while also protecting users' privacy. The designed schemes have higher computation costs.

The following are the research contributions of the ORSDDSPNL-SDC technique.

- A novel privacy-preserving technique called ORSDDSPNL-SDC is introduced for secure big data transmission by identifying the authorized user with a better accuracy and in less time.
- First, the registered cloud users put their information with the cloud service and receive a user _ID and a password each for further processing. This ID is used for identifying the authorized user with minimum time as well as storage overhead.
- To improve secure big data transmission between the server and the users, orthogonal regressed steepest descent multilayer deep perceptive neural learning is applied in ORSDDSPNL-SDC to analyze the received user ID with registered user ID through the Szymkiewicz–Simpson coefficient. The Szymkiewicz–Simpson coefficient is a mathematical function used to match the ID of the user for identifying authorized users. The Maxout activation function identifies the user as authorized or unauthorized based on classification results. This helps improve classification accuracy.
- Finally, the steepest descent function is used to reduce the error rate with aid of the weight updating rule.
- A series of experimental assessments is carried out to quantify the proposed ORSDDSPNL-SDC method in comparison to existing algorithms and performance metrics.

The rest of the paper is organized as follows. Section two presents a literature review on the IoT security of application domains. The third section describes the proposed ORSDDSPNL-SDC for IoT security. Section four provides the experimental settings and dataset description. Section five presents the outcomes and discussion generated by various performance indicators and the paper is concluded in Section six.

2. RELATED WORKS

Big data security is a major concern due to its critical role in the widespread adoption of cloud architectures. In [12], a case-based security-by-design framework for big data deployment over cloud computing was developed. The study's findings show the effectiveness of raising security awareness in cloud-based data environments.

Multifactor authentication and lightweight cryptography encryption methods were developed in [13] to protect big data systems by using a cloud-enabled IoT environment. But, those methods failed to perform mutual authentication between gateway devices and IoT devices. To identify the genuine or fraudulent user, a deep learning-based IoT data analytics method was devised [14]. But, despite employing larger datasets, it was unable to reduce the time and cost limits.

For safe big-data transmission, the Data-centric Authentication technique was introduced in [15]. However, neither the time nor the complexity of the storage was reduced. Secure Authentication and Data Sharing in Cloud (SADS-Cloud) were introduced in [16] for data security with big data. However, they failed to speed up the operations of secure authentication and data sharing in the cloud.

According to [17], MSCryptoNet is a model based on multi-scheme fully homomorphic encryption in the privacy-preserving scenario that allows for scalable execution and conversion of the trained neural network. MSCryptoNet was claimed to be secure as a deep learning-based privacy-preserving scheme over aggregated encrypted data. Furthermore, it was claimed that the MSCryptoNet does not require any correspondence between data providers and the cloud server to provide privacy-preserving predictions. However, there is no relevant evidence in the literature to support the second claim.

An enhanced Diffie Hellman when combined with Elliptic Curve (E-ECDH) was developed in [18] for secure and lightweight communications of connected devices in IoT. However, it failed to improve the security of IoT by considering applications in different domains. A Chebyshev Chaotic Map-based lightweight multi-server authentication method was developed in [19] for secure data transmission with minimum computation cost. However, the storage overhead was not reduced.

In [20], a lightweight and privacy-aware fine-grained access control method for secure IoT-oriented health data transmission was developed. However, authentication-based access control was not implemented. Privacy-preserving hierarchical fuzzy neural network (PPHFNN) was developed in [21] for improving the security of heterogeneous big data. However, the performance of the accuracy level was not improved.

A three-factor user access control mechanism for data usage in the IoT environment is presented in [22]. Despite claims that the suggested approach is secure, it was unclear what kinds of threats the method handled. By combining attribute-based encryption with edge computing, [23] implemented a mechanism for a healthcare IoT system. It was asserted that the suggested mechanism offers a reliable, adaptable, secured access control with data verification and permits data consumers to take use of the lightweight decryption. In [24], a crypto-deep neural network cloud security (CDNNCS) method for increasing cloud user trust was developed. However, the other deep learning architectures were not implemented for the dynamic behavior of communication in the cloud environment. Multifactor authentication and lightweight cryptography encryption methods were developed in [25] based on cloud-enabled IoT environments to protect big data systems. However, mutual authentication between devices and IoT devices was not performed.

3. METHODOLOGY

The Internet of Things (IoT) system is composed of several internetworked smart equipment that shares information *via* the world wide web. Such self-directed devices are dispersed in specific fields to sense and collect data. The data collected from IoT sensors is transferred to cloud servers for storage and further processing in IoT cloud-based systems. The data is then accessed by authorized users from the cloud server. However, security concerns are growing in conjunction with the widespread adoption of IoT-based infrastructure. As a result, there is an urgent need to propose novel security mechanisms to protect IoT systems from potential threats. Based on this motivation, a novel technique called ORSDDSPNL-SDC is introduced. The ORSDDSPNL-SDC technique accurately identifies the authorized or unauthorized user to get the data from the cloud service through orthogonal regressed steepest descent multilayer deep perceptive neural learning. The advantage of multilayer deep perceptive neural learning is to provide better results with a large volume of big data.

The architecture diagram of the proposed ORSDDSPNL-SDC technique for secure big data communication is shown in Figure 1. The proposed technique comprises two types of entities; namely, cloud users '$CU = CU_1, CU_2, \ldots, CU_n$ who want to store their big data '$D = D_1, D_2, \ldots., D_n$' collected from various IoT devices. The proposed ORSDDSPNL-SDC technique includes three major processes; namely, registration, user authentication and secure data communication.

The IoT device allows the user to collect the data simply using the different sensors installed at various locations. When the users want to store their collected data from IoT devices on the cloud server, they first perform the registration process. The user has to fill up the registration form which is provided by the cloud server. It contains information about the user, like name, date of birth, age, gender, mobile number and so on. After filling out the registration form, the user presses the submit button. The cloud server stores the user details in its database and creates the new user_ID.

After successful registration, the user stores the data in the cloud server database for further processing. Whenever the users want to access their desired data from the cloud-server database, they first verify their authenticity. The proposed ORSDDSPNL-SDC technique uses the orthogonal regressed steepest descent multilayer deep perceptive neural learning for identifying the authorized or unauthorized user through classification.

An orthogonal regressed steepest descent multilayer deep perceptive neural learning is a family of machine learning methods that work based on artificial neural networks. The word deep learning refers to the use of numerous layers utilized in the network for learning the given input. The orthogonal regression is applied to deep learning for verifying the authenticity of the user by analyzing the current ID and the already stored ID in the server database based on the Szymkiewicz–Simpson coefficient. The Szymkiewicz–Simpson coefficient is a similarity function that is used to match the IDs of the cloud users. If the two IDs get matched, the max-out activation function returns the output as an authorized user or the unauthorized user. The cloud server offers the user the requested data for enhancing secure communication after identifying him/her as an authorized user. Otherwise, the cloud server denies access to unauthorized users. Secure communication between cloud users and servers is thus accomplished. A brief description of the proposed ORSDDSPNL-SDC technique with different processes is given below.

3.1 New User Registration

The registration phase is the initial step in the proposed ORSDDSPNL-SDC technology. The users must provide the necessary identification during the registration step to register their information in the server database. The user has to fill out a registration form which contains information about username, date of birth, age, gender, mobile number, mail_ID and so on.

Figure 2 exhibits the flow process of new user registration. First, the users enter their personal information in the registration form. The user has to give a valid username at the time of registration. The server checks the created username against the availability of that username in its database. If the username is not matched with the existing username in the server database, then the server sends an error and try again message. Otherwise, a new user ID is generated and stored in the server database for later use. Following registration, the users uploads their data '$D = D_1, D_2, \ldots.., D_n$' collected from various IoT devices into the server database.

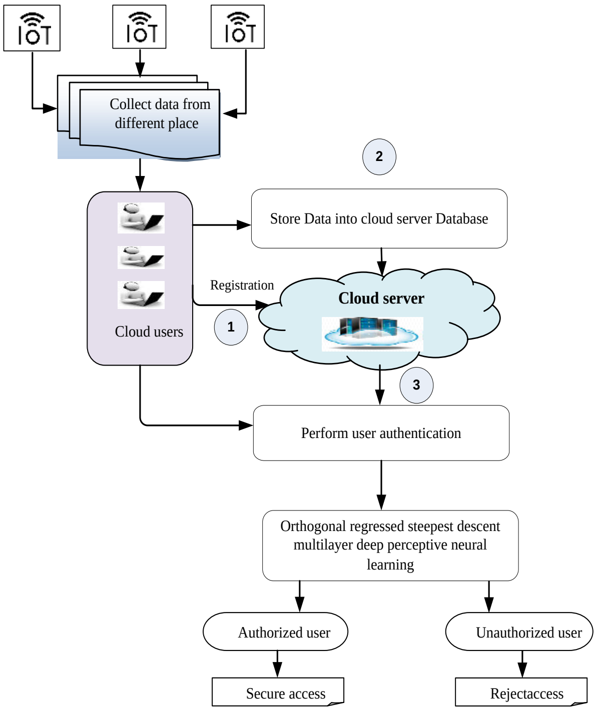
IoT
IoT
IoT
Collect data from different place
2
Store Data into cloud server Database
Registration
1
Cloud server
Cloud users
3
Perform user authentication
Orthogonal regressed steepest descent multilayer deep perceptive neural learning
Authorized user
Unauthorized user
Secure access
Rejectaccess

Figure 1. Architecture diagram of the proposed ORSDDSPNL-SDC technique.

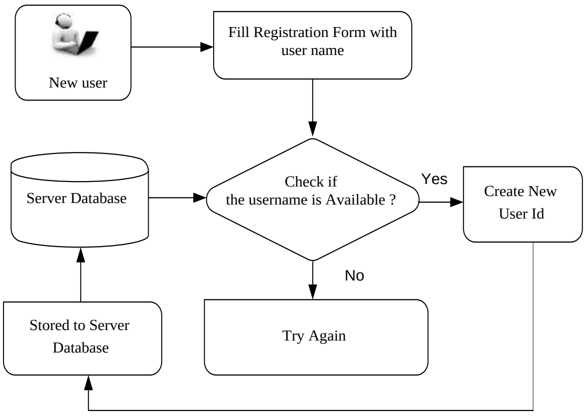
New user
Fill Registration Form with user name
Server Database
Check if the username is Available ?
Yes
Create New User Id
No
Stored to Server Database
Try Again

Figure 2. Flow process of new user registration.

The input dataset was separated into two parts; namely, the training and testing. Most cloud users (70%) were used for training and the remaining cloud users (30%) were taken for testing. Here, MHEALTH (Mobile HEALTH) dataset is considered for collecting the data. The different sensors (i.e., IoT devices) are located on the candidate's chest, right wrist and left ankle for recordings of body motion and vital indications while performing physical activities. The activities are standing still, sitting and relaxing, lying down, walking, climbing stairs, waist bends forward, frontal elevation of arms, knees bending, cycling, jogging, running, jumping front and back, respectively. The big data collected from IoT devices is sent to a cloud data center for further processing.

3.2 Orthogonal Regressed Steepest Descent Multilayer Deep Perceptive Neural Learning-based User Authentication

Whenever the users access their data from the cloud server database, they first verify authenticity for guaranteeing secure communication. User authentication is the process of verifying the identity of the user. The proposed ORSDDSPNL-SDC technique uses the deep learning technique called orthogonal regressed steepest descent multilayer deep perceptive neural learning for verifying the authenticity of the user by classifying the user as an authorized user or as an unauthorized user. Orthogonal regressed steepest descent multilayer deep perceptive neural learning is a fully connected feed-forward artificial neural network. A multilayer perceptron includes at least three layers of nodes such as an input layer, a hidden layer and an output layer.

Figure 3 reveals the schematic structure of the deep multilayer perceptive learning technique for user authentication. The input layer is initially given the number of cloud users. The network has many layers, including an input layer, hidden layers and an output layer. Each layer in deep learning learns the given input and transforms it into the next layer.

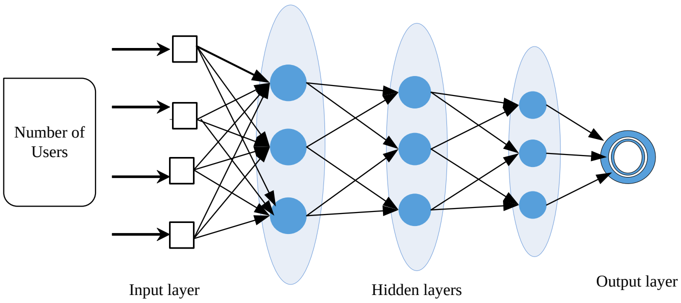

Figure 3. The orthogonal regressed steepest descent multilayer deep perceptive neural learning architecture.

To form the neural network architecture, the nodes between one layer and another are linked in a feed-forward way by regulating weights. The input layer accepts the cloud users $CU = CU_1, CU_2, \dots, CU_n$.

$$X_{input}(t) = B + \left[\sum_{i=1}^{n} CU_i * \varphi_{i\_h}\right] \quad (1)$$

where, $X_{input}(t)$ indicates the input unit, φ_{ih}denotes a controlling weight between the input layer and the hidden layer, CU_i indicates a cloud user and B denotes a bias in which the integer value stored is +1. The input is sent to the first hidden layer.

When a registered cloud user desires to retrieve big data from the cloud service, he/she must first log in to the cloud server and enter a valid ID on his/her log-in page. Then, the server identifies the authorized user or unauthorized user with the help of orthogonal regression.

Orthogonal regression is a statistical technique for estimating relationships and determining the best fit between two variables, such as the currently received user ID and the already stored ID. The Szymkiewicz–Simpson coefficient is a similarity function that is used to calculate the degree of overlap between two finite sets; namely, the currently entered user ID and the already stored ID. The Szymkiewicz–Simpson coefficient is used to identify the

$$Cff_{ss} = \frac{|RID \cap RegID|}{|S_{RID}||RegID|} \quad (2)$$

where, Cff_{ss} represents the Szymkiewicz–Simpson coefficient, RID indicates a received ID and $RegID$

indicates a registered ID, $|S_{RID}|$ denotes a set containing a string of received IDs, $|RegID|$ denotes a set containing a string of registered IDs, $RID \cap RegID$ designates mutual dependence between the two IDs.

In the second hidden layer, Maxout activation is applied for verifying the similarity coefficient value. The activation function of a node in the layer defines the output of that node based on a given set of inputs. Maxout activation finds the maximum value of the similarity coefficient.

$$\beta = \begin{cases} 1; & argmax\ Cff_{ss} \\ 0; & otherwise \end{cases} \tag{3}$$

where, β denotes the Maxout activation function, the activation function (β) returns '1' when the two IDs get matched and the user is classified as an authorized user. The activation function (β) returns '0', which means that the two IDs are not matched and the user is classified as an unauthorized user. The hidden layer output is formulated as given below,

$$Z = \left(\sum_{i=1}^{n}\sum_{i=1}^{n} CU_i * \varphi_{i\_h}\right) + [Z' * \varphi_h] \tag{4}$$

where, 'Z' indicates the output of the hidden layer, 'φ_{i_h}' denotes the regulating weight between the input and the hidden unit, φ_h indicates the regulating weight of the hidden unit and Z' denotes an output of the previously hidden layer.

$$W = \varphi_{h\_o} Z \tag{5}$$

where, 'W symbolizes the final results of the output layer, φ_{h_o} represents the weight between the hidden layer and the output layer and Z denotes the output of the hidden layer. After classification, the degree of training error in an output node is calculated as follows:

$$T\varepsilon = \frac{1}{2}(T_o - W)^2 \tag{6}$$

where, $T\varepsilon$ denotes a training error, T_o indicates a target output value and W denotes an output produced by the multilayer perceptron. Applying the steepest descent function leads to minimize the error in the entire output by changing the weight between the layers. There is a mathematical steepest descent function that is used to minimize the error by using the update rule.

$$\Delta\varphi = -\gamma\ \frac{\partial T\varepsilon}{\partial\varphi} W \tag{7}$$

where, $\Delta\varphi$ updating the weight, γ indicates a learning rate, $T\varepsilon$ denotes an error and W indicates an output of classification result. In this way, accurate classification of the authorized and unauthorized users is determined with minimum error. Algorithm 1 describes the step-by-step process of secure data communication using Orthogonal Regressed Steepest Descent Deep Structured Perceptive Neural Learning.

The users first register their information with the cloud service. The cloud server then generates a new user ID for accessing the service. After receiving the new ID, the user stores the data collected from the IoT device in a cloud server. Every time a user accesses data from the cloud server, it first confirms the user's ID. The Szymkiewicz–Simpson coefficient is applied for verifying the user ID. If the two IDs get matched, the max-out activation function returns '1' and the user is classified as authorized. Otherwise, the Maxout activation function returns '0' and the user is classified as unauthorized. Finally, the training error of classification is minimized by updating the weight between the layers using the steepest descent function. Then, the cloud server allows for accessing the data to the authorized user and denied access to the unauthorized user. In this way, secure big data communication is performed.

4. Experimental Setting

In this section, the CloudSim simulator is used to perform an empirical analysis of the proposed ORSDDSPNL-SDC technique as well as the existing HBRSS [1], TCFDL [2] and deep learning and IoT-based data processing frameworks [3]. Secure big data communication is performed using the MHealth dataset taken from https://www.kaggle.com/datasets/gaurav2022/mobile-health. The MHEALTH (Mobile HEALTH) dataset comprises recordings of 10 individuals' body motions and vital signs as they engaged in twelve different physical activities. Sensors placed on the subject's chest, right

Algorithm 1: Orthogonal Regressed Steepest Descent Deep Structured Perceptive Neural Learning based Secured Data Communication

```
Input:  Dataset ‘$DS$’, IoT device ‘$S = S_1, S_2, \dots, S_n$’, cloud user’s ‘$CU = CU_1, CU_2, \dots, CU_n$’,
big data ‘$D = D_1, D_2, \dots., D_n$’, cloud server ‘$CS$’
Output:   Improve the secure communication
Begin
Step 1:    For each user
Step 2:        Fill in the details in the registration form
Step 3:         $CS$ creates a new user_ID
Step 4:         $CU$ stores the data on a cloud server
Step 5:   End for
Step 6:       If $CU$ accesses data from the cloud then
Step 7:           $CS$ verifies the authenticity of the user
Step 8:     End if
Step 9:       User login to the system with User_ID
Step 10:       server uses the  Szymkiewicz–Simpson coefficient  ‘$Cff_{ss}$’
Step 11:           Apply max out activation function
Step 12:         $\boldsymbol{if}$  $(argmax\ Cff_{ss})$   then
Step 13:              $\beta$  returns  ‘1’
Step 14: the user is classified as authorized
Step 15:                    Secure data access
Step 16:              else
Step 17:                $\beta$  returns  ‘0’
Step 18: The user is classified as unauthorized
Step 19:       Denied  access
Step 20:       Compute error rate ‘$T\varepsilon$’
Step 21:         Apply the steepest descent to minimize the error
Step 22:         update the weight ‘$\Delta\varphi$’
  End
```

wrist and left ankle are utilized to track the acceleration, rate of turn and magnetic field orientation of various body parts. The sensor, which is placed on the chest, may also record a 2-lead ECG to check for various arrhythmias and record the heartbeat. Each sensing modality is recorded on a sampling rate of 50 Hz that is focused enough to capture human activity. All sessions were recorded with a video camera. The dataset is found to simplify general behaviors of the day-by-day living, given the variety of body parts involved in each one (e.g. the frontal elevation of arms *vs.* knees bending), the intensity of the actions (e.g. cycling *vs.* sitting and relaxing) as well as the execution speed or dynamicity (e.g. running *vs.* standing still). The dataset summary, such as 12 activities, 3 sensor devices and 10 subjects, is considered. The dataset consists of 14 attributes (i.e., features) and 12, 15,745 instances (i.e., samples).

Table 1. Attributes’ descriptions.

S. No	Attributes	Description
1	alx	Acceleration from the left-ankle sensor (X-axis)
2	aly	Acceleration from the left-ankle sensor (Y-axis)
3	alz	Acceleration from the left-ankle sensor (Z-axis)
4	glx	Gyro from the left-ankle sensor (X-axis)
5	gly	Gyro from the left-ankle sensor (Y-axis)
6	glz	Gyro from the left-ankle sensor (Z-axis)
7	arx	Acceleration from the right-ankle sensor (X-axis)
8	ary	Acceleration from the right-ankle sensor (Y-axis)
9	arz	Acceleration from the right-ankle sensor (Z-axis)
10	grx	Gyro from the right-ankle sensor (X-axis)
11	gry	Gyro from the right-ankle sensor (Y-axis)
12	grz	Gyro from the right-ankle sensor (Z-axis)
13	Activity	Corresponding activity
14	Subject	Volunteer subjects (1 – 9)

The ten volunteers generate a lot of data (i.e., instances) and the data is communicated to the authorized entity. The hyperparameter tuning of the ORSDDSPNL-SDC technique is considered such that the value of learning rate used is 0.1, the number of hidden layers is 3, the number epochs is 10, the weight is 1.0 and the bias is 0.1 in the ranges of a network trained, 10 cross-validations are used to solve the overfitting/underfitting, imbalance problems. The steepest descent is a first-order iterative optimization algorithm to minimize the loss as quickly as possible.

5. Results and Discussion

In this section, the performances of the proposed ORSDDSPNL-SDC technique and existing HBRSS [1], TCFDL [2] and deep learning and IoT-based data processing frameworks [3] are discussed with different metrics, such as classification accuracy, error rate and classification time and space complexity. The performance in terms of these parameters is analyzed with the help of a table and a graphical representation.

Classification accuracy: It is calculated as the proportion of cloud users who are correctly classified as authorized or unauthorized users to the total cloud users. The measure of accuracy is mathematically expressed as follows:

$$CA = \left(\sum_{i=1}^{n} \frac{CU_{AC}}{CU_i}\right) * 100 \qquad (8)$$

where, CA denotes the classification accuracy and is estimated based on the number of cloud users 'CU_i' and the number of cloud users accurately classified 'CU_{AC}' in the simulation. It is expressed as a percentage (%).

Error rate: It is measured as the ratio of cloud users that are incorrectly classified to the total number of cloud users taken as input. This error rate during the classification is expressed as follows:

$$ER = \left[\sum_{i=1}^{n} \frac{CU_{wC}}{CU_i}\right] * 100 \qquad (9)$$

where, 'ER' denotes the error rate, CU_{wC} denotes the cloud users wrongly classified and CU_i represents the number of cloud users. It is measured in terms of a percentage (%).

Classification time: It is defined as the amount of time consumed by the algorithm for classifying authorized and unauthorized users to perform secure communication. The following is a mathematical definition of the classification time:

$$CT = \sum_{i=1}^{n} CU_i * Time\ [classification] \qquad (10)$$

where, CT indicates the classification time, CU_i represents the number of cloud users and *time* $[classification]$ denotes the time consumed in classification. It is quantified in milliseconds (ms).

Storage overhead: It is defined as how much memory the algorithm uses to classify authorized and unauthorized users when performing secure communication. The total amount of space consumed is calculated as follows:

$$SO = \sum_{i=1}^{n} CU_i * MEM\ [classification] \qquad (11)$$

where, SO indicates storage overhead, CU_i denotes the number of cloud users and MEM denotes the memory space consumed for classification. It is measured in terms of Megabytes (MB).

Figure 4 compares the classification accuracy of three different methods; namely, ORSDDSPNL-SDC, existing HBRSS [1], TCFDL [2] and Deep learning and IoT-based data processing frameworks [3] to several cloud users. Based on the classification, user authentication is performed for secure data access from the server. This significant enhancement of the ORSDDSPNL-SDC is verified through statistical estimation. As illustrated in Figure 4, increasing the number of cloud users causes a slow decrease in classification accuracy. This is because by increasing the number of cloud users, cloud users being classified accurately for secure data communication also gets deteriorated. This in turn causes a small decreasing trend in classification accuracy also. In the first iteration, we'll use 10000 cloud users to run our experiments. The accuracy of classifying authorized or unauthorized users using ORSDDSPNL-SDC is obtained as 99.65%. On the other hand, the accuracy of identifying authorized or unauthorized users by applying [1] and [2] is 98.1% and 98.8%, respectively. Likewise, different performance results are observed for each method. Overall, the observed results show that using ORSDDSPNL-SDC improves classification accuracy by 6% when compared to [1], 4% when compared to [2] and 5% when compared to [3].

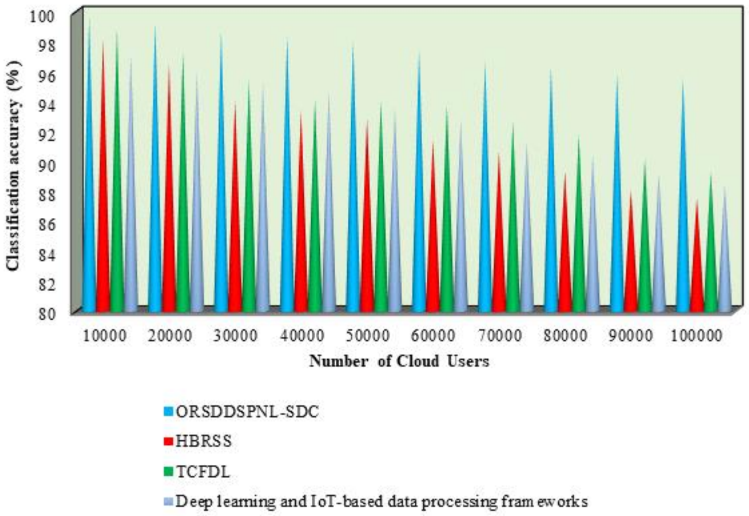

Figure 4. Comparison according to classification accuracy *versus* the number of cloud users.

Figure 4 depicts a comparison of classification accuracy for 100000 different cloud users. The figure shows that the number of cloud users is input in the x-axis and the overall performance results of classification accuracy are received in the y axis. According to Figure 4, the proposed ORSDDSPNL-SDC technique provides superior performance of classification accuracy as compared to HBRSS [1], and TCFDL [2] and deep learning and IoT-based data processing frameworks [3]. The existing TCFDL was employed to investigate security. But, maximum accuracy is achieved as a difficult issue in AI-enabled IoT. Contrary to existing works, orthogonal regressed steepest descent multilayer deep perceptive neural learning is utilized to enhance accuracy. This is owing to the cloud server authenticating the user ID with the already stored ID of that particular user at the time of registration. The verification is performed using the orthogonal regressed steepest descent multilayer deep perceptive neural learning. The orthogonal regression is applied to find the authorized users by using Szymkiewicz–Simpson coefficient. The Szymkiewicz–Simpson coefficient matches the user ID. The max-out activation function analyzes the similarity coefficient values and returns the classification result. Finally, the authorized user or unauthorized user classification results are getting more accurate at the output layer than in the existing methods.

The experimental results of the error rate for various categorization outcomes using three methods; namely, ORSDDSPNL-SDC, HBRSS [1] and TCFDL [2] and deep learning and IoT-based data processing frameworks [3] are shown in Figure 5. The observed results indicate that the overall performance of ORSDDSPNL-SDC technique is found to be minimized when compared to the other existing approaches; HBRSS [1] and TCFDL [2] and deep learning and IoT-based data processing frameworks [3]. This is proved through statistical examination. Let us consider 10000 users for calculating the error rate. By applying the ORSDDSPNL-SDC technique, the observed error rate is 0.35%. By considering [1], [2] and deep learning and IoT-based data processing frameworks [3], the observed error rates are 1.9%, 1.2% and 1.2%, respectively. Totally ten diverse results are observed for each method. The observed results of the ORSDDSPNL-SDC technique are compared to the existing methods. When compared to the existing methods, the obtained comparison result shows that the average error rate using the ORSDDSPNL-SDC technique is reduced by 71%, 64% and 66%, respectively. Figure 5 depicts the graphical representation of the ORSDDSPNL-SDC technique as well as of the three methods.

Figure 5 presents the result comparison of error rate versus the number of cloud users using different methods. However, the experiment is carried out with the number of 100000 cloud users to compute the error rate during the classification. From this result, the error rate using ORSDDSPNL-SDC technique was said to be reduced upon comparison with [1], [2] and deep learning and IoT-based data processing frameworks [3]. The reason behind the minimum error rate using ORSDDSPNL-SDC was owing to the application of the steepest descent function. After classification, the steepest descent function is applied for updating the weight between the layers of the multilayer deep perceptive neural learning classifier. As a result, the error rate of cloud user classification is minimized.

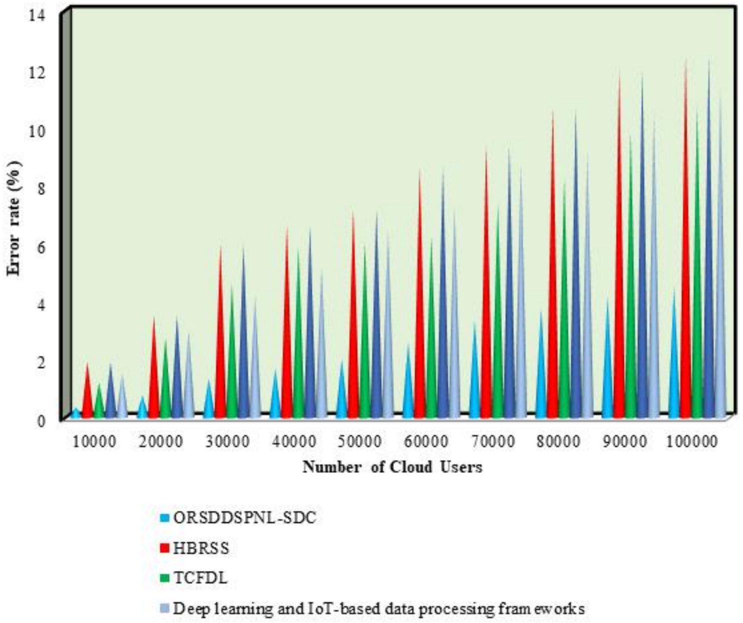

Figure 5. Comparison according to error rate *versus* number of cloud users.

Figure 6 illustrates the classification time using different methods; namely, ORSDDSPNL-SDC, HBRSS [1], TCFDL [2] and deep learning and IoT-based data processing frameworks [3]. From the observed results, the ORSDDSPNL-SDC technique decreases the classification time when compared to existing techniques. For example, with 10000 cloud users, the classification time using the ORSDDSPNL-SDC was observed at $3960ms$, whereas the classification times using [1] and [2] are $6100ms$ and $5500ms$. Similarly, different performance outcomes are obtained with each technique. The overall observed ORSDDSPNL-SDC results are compared to the results of existing methods. Ultimately, the mean of compared data reveals that the classification time using ORSDDSPNL-SDC technique is increased by 31% when compared to [1], 23% when compared to [2] and 13% when compared to [3].

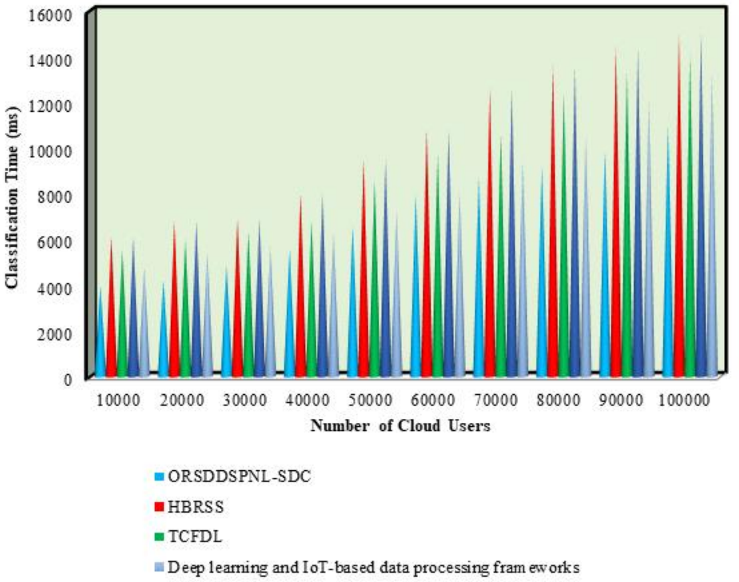

Figure 6. Comparison according to classification time *versus* the number of cloud users.

Figure 6 reveals the graphical representation of classification time for the different techniques for 100000 users. The graphical illustration implies that an increase in cloud users will lengthen the time required for classification. According to the evaluation's data, using the ORSDDSPNL-SDC reduces classification time compared to the existing [1], [2] and deep learning and IoT-based data processing frameworks [3], respectively. In traditional IoT systems, sensed data was often uploaded to the cloud that was examined with CFD. But, the data transmission causes a high classification time. The reason behind this improvement is to perform the cloud user registration. In the ORSDDSPNL-SDC technique, the cloud users first register their details to the server for storing the data collected from the IoT device. After the registration, the cloud server generates the user ID and password for further processing. When the users access their data, the server identifies the users as authorized or unauthorized for secure data communication. Initially, the users' details are registered and then the server collects the data from the registered user. Then, classification is performed using deep learning with the help of the Szymkiewicz–Simpson coefficient by matching the ID of the user. Then, the max-out activation function returns the authorized and unauthorized user classification results. This process minimizes the time consumption of cloud user classification.

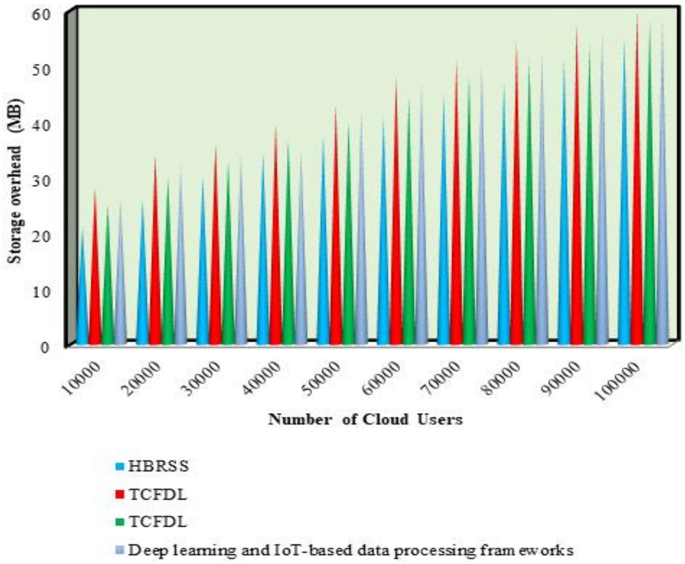

Figure 7. Comparison according to storage overhead *versus* the number of cloud users.

The performance analysis of the storage overhead using the different methods ORSDDSPNL-SDC, HBRSS [1], TCFDL [2] and deep learning and IoT-based data processing frameworks [3] is shown in Figure 7. The observed results indicate that the performance of memory consumption of the ORSDDSPNL-SDC technique is relatively minimum compared to the existing methods. With the consideration of 10000 users, the memory consumption to perform classification was found to be 21 MB. However, the memory consumption of existing [1] and [2] was found to be 28$MB$and 25 MB. The observed results indicate that the ORSDDSPNL-SDC technique reduces memory consumption. The overall results of memory consumption are compared to the previous findings after getting the ten results. In comparison to [1], [2] and [3], the comparison analysis shows a 15%,17% and 11% reduction in the approved user classification's memory use, respectively. Cloud storage, as well as cloud services, offer stronger computing power as well as distributed computing ability for IoT users at a lesser cost. But, the security problems of the cloud limit the growth of cloud computing and storage. The reason behind reducing memory consumption is applying orthogonal regressed steepest descent multilayer deep perceptive neural classifier. The proposed classifier accurately classifies the user as authorized or unauthorized with lesser memory consumption.

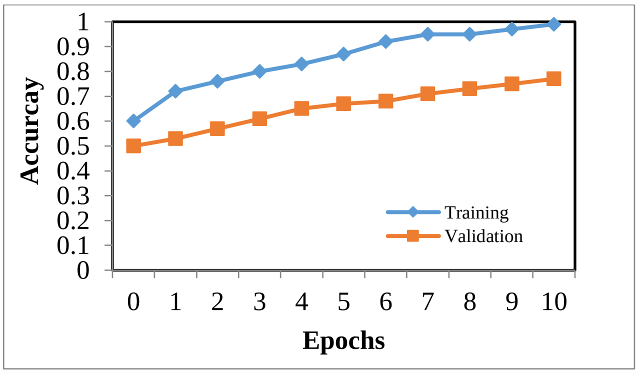

Figure 8. Training and testing accuracy *vs.* epochs.

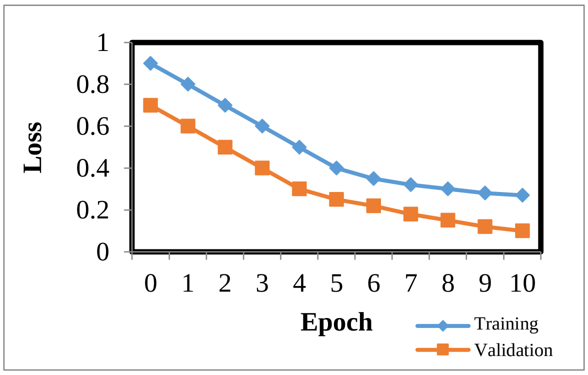

Figure 9. Training and testing loss *vs.* epochs.

In classification, data is used for divided into two parts; one for training and the other for testing. The training's accuracy will be higher than the testing's accuracy. In addition, the training's loss will be smaller than the testing's loss. The accuracy and loss are estimated in terms of percentage (%).

Figure 8 and Figure 9 show the impact of accuracy and loss along with the epoch using the Mobile HEALTH dataset. An epoch is taken in the horizontal direction, while accuracy and loss are observed at the vertical axis. As shown in the graphical chart, there are two various colors of lines such that blue and brown indicating the two parts; namely, training and validation, respectively.

6. CONCLUSION

While the Internet of Things (IoT) strengthens the viability of consumers in cloud environments, a lack of security practices raises the potential risks of protecting sensitive user data. One essential capability of big data communication is the use of IoT devices to secure data transmission. Therefore, a novel technique called ORSDDSPNL-SDC is introduced for guaranteeing secure communication by identifying the authorized cloud users to access the services from the cloud. The cloud users register their details to the server for accessing the various services from the cloud. The server creates a new ID and password for each user in the cloud. The data is then sent to the cloud server for storage by the cloud user. After that, the orthogonal regressed steepest descent multilayer deep perceptive neural learning is employed in ORSDDSPNL-SDC to analyze the receiver ID with the registered ID based on Szymkiewicz–Simpson coefficient. The Maxout activation function correctly classifies the user as authorized or unauthorized based on the similarity coefficient result. To minimize the error, the steepest descent function is then employed, increasing the classification accuracy of the users. In this way, ORSDDSPNL-SDC performs secured communication with authorized cloud users. The comprehensive experimental assessment is carried out using different performance metrics, such as classification accuracy, error rate, classification time and storage overhead *versus* the number of cloud users. The quantitative analysis confirms that the ORSDDSPNL-SDC technique has achieved higher accuracy of user classification with lesser time consumption as well as storage overhead when compared to other conventional methods.

REFERENCES

[1] H. Xie, Z. Zhang, Q. Zhang, S. Wei and C. Hu, "HBRSS: Providing High-secure Data Communication and Manipulation in Insecure Cloud Environments," Computer Comm., vol. 174, pp. 1-12, 2021.

[2] Q. Zhang, H. Zhong, W. Shi and L. Liu, "A Trusted and Collaborative Framework for Deep Learning in IoT," Computer Networks, vol. 193, pp. 1-10, DOI: 10.1016/j.comnet.2021.108055, 2021.

[3] Y. Li, Y. Zuo, H. Song and Z. Lv, "Deep Learning in Security of Internet of Things," IEEE Internet of Things Journal, vol. 9, no. 22, pp. 22133–22146, 2022.

[4] S. Zargar, A. Shahidinejad and M.Ghobaei-Arani, "A Lightweight Authentication Protocol for IoT-based Cloud Environment," Int. J. of Communication System, vol. 34, no.11, pp. 1-17, 2021.

[5] J. Shen, D. Liu, Q. Liu, X. Sun and Y. Zhang, "Secure Authentication in Cloud Big Data with Hierarchical Attribute Authorization Structure," IEEE Transactions on Big Data, vol. 7, no. 4, pp. 668 – 677, 2021.

[6] U. Iqbal, A. Tandon, S. Gupta, A. R. Yadav, R. Neware and F. W. Gelana, "A Novel Secure Authentication Protocol for IoT and Cloud Servers," Wireless Communications and Mobile Computing, vol. 2022, pp. 1-17, DOI: 10.1155/2022/7707543, 2022.

[7] A. Karati, R. Amin, S. K. H. Islam and K. R. Choo, "Provably Secure and Lightweight Identity-based Authenticated Data Sharing Protocol for Cyber-physical Cloud Environment," IEEE Transactions on Cloud Computing, vol. 9, no. 1, pp. 318 – 330, 2021.

[8] M. Wazid, J. Singh, A. K. Das, S. Shetty, M. K. Khan and J. J. P. C. Rodrigues, "ASCP-IoMT: AI-enabled Lightweight Secure Communication Protocol for Internet of Medical Things," IEEE Access, vol. 10, pp. 57990 – 58004, 2022.

[9] M. A. Jan, F. Khan, S. Mastorakis, M. Adil, A. Akbar and N.Stergiou, "LightIoT: Lightweight and Secure Communication for Energy-efficient IoT in Health Informatics," IEEE Transactions on Green Communications and Networking, vol. 5, no. 3, pp. 1202 – 1211, 2021.

[10] X. Lu, L. Yin, C. Li, C. Wang, F. Fang, C. Zhu and Z. Tian, "A Lightweight Privacy-preserving Communication Protocol for Heterogeneous IoT Environment," IEEE Access, vol. 8, pp. 67192–67204, DOI: 10.1109/ACCESS.2020.2978525, 2020.

[11] Z. Guan, W. Yang, L. Zhu, L. Wu and R. Wang, "Achieving Adaptively Secure Data Access Control with Privacy Protection for Lightweight IoT Devices," Science China Information Sciences, vol. 64, pp. 1-14, DOI: 10.1007/s11432-020-2957-5, 2021.

[12] F. M. Awaysheh, M. N. Aladwan, M. Alazab , S. Alawadi, J. C. Cabaleiro and T. F. Pena, "Security by Design for Big Data Frameworks Over Cloud Computing," IEEE Transactions on Engineering Management, vol. 69, no. 6, pp. 3676–3693, 2022.

[13] L. Atiewi, A. Al-Rahayfe, M. Almiani, S. Yussof, O. Alfandi, A. Abugabah and Y. Jararweh, "Scalable and Secure Big Data IoT System Based on Multifactor Authentication and Lightweight Cryptography," IEEE Access , vol. 8, pp. 113498 – 113511, DOI: 10.1109/ACCESS.2020.3002815, 2020.

[14] K. Thilagam, A. Beno, M. V. Lakshmi et al., "Secure IoT Healthcare Architecture with Deep Learning-based Access Control System," Journal of Nanomaterials, vol. 2022, pp. 1-8, DOI: 10.1155/2022/2638613, 2022.

[15] R. Li, H. Asaeda and J. Wu, "DCAuth: Data-centric Authentication for Secure In-network Big-data Retrieval," IEEE Transactions on Network Science and Engineering, vol. 7, no. 1, pp.15 – 27, 2020.

[16] U. Narayanan, V. Paul and S. Joseph, "A Novel System Architecture for Secure Authentication and Data Sharing in Cloud Enabled Big Data Environment," J. of King Saud Uni.-Computer and Information Sciences, vol. 34, no. 6, pp. 3121-3135, DOI: 10.1016/j.jksuci.2020.05.005, 2022.

[17] O. Kwabena, Z. Qin and T. Zhuang "MSCryptoNet: Multi-scheme Privacy-preserving Deep Learning in Cloud Computing," IEEE Access, vol. 7, pp. 29344–29354, DOI: 10.1109/access.2019.2901219, 2019.

[18] M. I. Ahmed and G. Kannan, "Secure End to End Communications and Data Analytics in IoT Integrated Application Using IBM Watson IoT Platform," Wireless Personal Communications, vol. 120, pp.153–168, DOI: 10.1007/s11277-021-08439-7, 2021.

[19] T. Maitra, S. Singh, R. Saurabh and D. Giri, "Analysis and Enhancement of Secure Three-factor User Authentication Using Chebyshev Chaotic Map," J. of Inf. Security and Appli., vol.61, pp.1-13, 2021.

[20] J. Sun, H. Xiong, X. Liu, Y. Zhang, X. Nie and R. H. Deng, "Lightweight and Privacy-aware Fine-grained Access Control for IoT-oriented Smart Health," IEEE IoT J., vol. 7, no. 7, pp. 6566- 6575, 2020.

[21] L. Zhang, Y. Shi, Y. Chang and C. Lin, "Hierarchical Fuzzy Neural Networks with Privacy Preservation for Heterogeneous Big Data," IEEE Transactions on Fuzzy Systems, vol. 29, no. 1, pp. 46–58, 2021.

[22] S. Banerjee, S. Roy, V. Odelu et al., "Multi-authority CP-ABE-based User Access Control Scheme with Constant-size Key and Ciphertext for IoT Deployment," J. of Information Security and Applications, vol. 53, pp. 1-22, DOI: 10.1016/j.jisa.2020.102503, 2020.

[23] S. Xu, Y. Li, R. H. Deng, Y. Zhang, X. Luo and X. Liu "Lightweight and Expressive Fine-grained Access Control for Healthcare Internet-of-Things," IEEE Transactions on Cloud Computing, vol. 10, no. 1, pp. 474–490, 2022.

[24] P. Abirami and S. V.Bhanu, "Enhancing Cloud Security Using Crypto-deep Neural Network for Privacy Preservation in Trusted Environment," Soft Computing, vol. 24, pp. 18927–18936, DOI: 10.1007/s00500-020-05122-0, 2020.

[25] S.Atiewi, A. Al-Rahayfeh, MuderAlmiani et al., "Scalable and Secure Big Data IoT System Based on Multifactor Authentication and Lightweight Cryptography," IEEE Access, vol. 8, pp. 113498–113511, DOI: 10.1109/ACCESS.2020.3002815, 2020.

ملخص البحث:

مع ازدياد انتشار أجهزة إنترنت الأشياء، يتمّ توليد كمياتٍ هائلةٍ من البيانات كُلَّ يوم. وتحتاج البيانات التي يتمّ جمعها الى تحديثٍ مستمرٍ في جهاز الخادِم السّحابي. إلى جانب ذلك، فإنّ نقل البيانات في بيئة السَّحابة بواسطة الإنترنت يواجه العديد من التّهديدات. ومع ذلك فإن قضية الأمْن الخاصّ بالبيانات تنقصها الفعالية في مجال الاتّصال بالبيانات الضّخمة. لذا، فإنّ هذه الورقة تهدف الى اقتراح تقنية مبتكرة للحفاظ على أمْن البيانات الضّخمة؛ إذ تتكون التّقنية المقترحة من ثلاث مراحل هي: التّسجيل، والتّحقُّق من هوية المستخدم، والنّقل الآمن للبيانات. وتتميز التّقنية المقترحة بدقّةٍ أعلى وزمنٍ أقل مقارنة بالتّقنيات التي تمّ استخدامها في دراساتٍ سابقة. حيث يقوم المستخدم بإدخال رمزٍ تعريفي ويقوم النّظام بمقارنته برقمٍ تعريفي مخزّن في جهاز الخادِم، مما يؤدّي الى تصنيف الستخدمين الى مستخدمين مخوّلين وآخرين غير مخوَّلين، بناءً على عددٍ من العوامل، ومنها: دقة التّصنيف، والزمن المستغرَق في التّصنيف، ومعدَّل الخطأ. وقد أثبت التّقييم نجاعة التّقنية المستخدَمة في هذا البحث من حيث تحقيق دقّة تصنيفٍ أعلى وزمن تصنيف أقلّ ومعدّل خطأ أقلّ، مقارنةً بالطّرق القائمة التي استُخدمت في دراسات أخرى.