IET Information Security

Research Article

IET Journals
The Institution of Engineering and Technology

Secure data storage and intrusion detection in the cloud using MANN and dual encryption through various attacks

ISSN 1751-8709
Received on 11th March 2017
Revised 12th June 2018
Accepted on 5th October 2018
E-First on 25th February 2019
doi: 10.1049/iet-ifs.2018.5295
www.ietdl.org

J. Anitha Ruth[1] ✉, *H. Sirmathi*[1], *A. Meenakshi*[1]

[1]*SRM University, SRM Nagar, Kattankulathur, Kancheepuram District-603 203, Tamilnadu, India*

✉ *E-mail: anitharuth0875@gmail.com*

Abstract: Nowadays, it is very important to maintain a high level security to ensure safe and trusted communication of information between various organisations. But secured data communication over the Internet and any other network is always under threat of intrusions and misuses. So intrusion detection system (IDS) has become a needful component in terms of computer and network security. In this research, the authors have intended to propose an effective method for text data based IDS and secure data storage. In the proposed preprocessing steps, the input text document is preprocessed and then change to the desired format. Next the resultant output is fed to the IDS. Here user text data is checked; whether the given data is normal or intrusive based on a modified artificial neural network (MANN). Here traditional neural network is modified by means of modified particle swarm optimisation. The final process of the authors' proposed method is to encrypt the file using dual encryption algorithms (RSA and AES). To improve the storage security of the proposed method, steganography techniques are utilised after the dual encryption. Their proposed system is implemented with the help of Cloud simulator in the working platform Java.

1 Introduction

Nowadays, the usage of the Internet and local networks are growing as well as the intrusion events to the computer in [1]. Computer systems can easily vulnerable to attack by anyone because of increased network connectivity. The main purpose of such an attack is to threaten the conventional security mechanisms on the systems and perform surplus operations by the intruder's authorisation. Some of the intruder's operations are reading protected or private data and doing some malicious damage to the system or user files [2]. With the use of building a complex tool, the system security operator can easily find malicious activities when they occur because it will monitor and report activities simultaneously. Intrusion detection systems (IDSs) are important to preserve proper network security [3, 4]. The main work of the IDS is to monitor networked devices and find the anomalous or malicious behaviour in the patterns of movement in the audit stream [5]. This software is mainly used to monitor the events take place in a computer system or network, examine the system events, find suspected intrusion and then raise an alarm. IDS can be divided into two types. They are host-based IDS and network-based IDS.

Host-based sensors are called host-based IDS and network-based sensors are called network-based IDS [6]. Host-based technology is used to analyse that accessed files and executed applications [7]. Host-based IDS collects information to form an individual computer system such as operating system trails, C2 audit logs, and system logs and then it will perform [8, 9]. Host-based intrusion detection is a powerful tool to detect the previous attacks and proper method to overcome their future application. Host-based IDS also use audit logs but they are much more automated and having developed sophisticated and responsive detection methods. Basically, host-based IDS is a monitor system, event and security logs on Windows NT and SYSLOG in UNIX atmosphere. Suppose if any one of these files changed, the IDS will compare the new log entry with attacked signature to check whether there is a match or not [10].

Moreover, network-based IDS is mainly used to detect the unauthorised usage of computers over a network like the Internet [11]. Network-based IDS will collect a huge number of data from the network [9]. Network-based IDS analyses events as packets of information exchange between computers (network traffic) [7]. Network-based IDSs will use raw network packets as the data source. Network-based IDS basically apply promiscuous mode to monitor through network adapter and examine all real-time traffic over the network. It attacks the reorganisation module. There are four basic techniques used to determine an attack signature. They are (i) pattern, expression or byte code matching, (ii) frequency or threshold crossing, (iii) correlation of lesser events and (iv) statistical anomaly detection [10]. When the host-based system finds more inside activity, as well the network-based system will find more incoming network activity. Finally, both systems can respond and/or alert the security officers to the proper manner [7]. A good IDS should be capable to differentiate between normal and abnormal user activities [12]. Security is a key requirement for cloud computing combine as a robust and feasible versatile solution [13]. This security has been divided into several parts and one of the most important parts is ensuring about the user authentication processes and managing accesses when users outsource sensitive data share on public or private cloud servers [14].

2 Related works

In Anil Kumar and Nanda Mohan [15], it was a combination of three techniques comprising two machine-learning paradigms. K-means clustering, fuzzy logics and neural network (NN) techniques deployed to configure an effective IDS. Out of the several problems in the traditional techniques of IDSs, the presence of a high rate of false alerts causes unnecessary interference of human analyst. The human analysts, in turn, perform an intensive analysis repeatedly to distinguish the nature of such alerts and initiate sufficient actions. The approach proposed reveals the advantage of converging K-means–fuzzy–neuro techniques to eliminate the preventable interference of human analyst on such occasions. The technique was tested using a multitude of background knowledge sets in DARPA network traffic datasets. The experimental results render remarkable improvement in reducing the false alarms in addition to increased ability to capture intrusion packets that are no similar to the ones in the training datasets.

Zarrabi and Zarrabi [16] incredibly proposed to limelight the IDS as a service cloud to successfully shield the user network. It

was capable to exploit confirmed qualities in the facilitated the extortion of the desired data from the user network for assessment. The connected model was projected to be extensible by allowing the users to tab into varies categories of the IDSCs concurrently to mix the attributes of varies products for other consistent IDS solution. The various modules of the user network on far-off frameworks could be displayed by an equal framework which established acceptance in the recruitment of the IDS result in vibrant scenarios. By processing of a basic edition of the novel design, the theory was realistic in the local network.

Due to the complexity of the cloud architecture, Mathew and Jose [17] appreciate emphasised the requirement for the deployment architecture of IDS in the cloud. They successfully explained and listed different issues in the cloud infrastructure and successfully engaged the IDSs and its execution in the cloud. In addition, they capably propounded the employment of the included and covered IDS on a cloud which was planned to overcome several attacks. Their original IDS method included the knowledge and performance evaluation to make the security of the cloud. In addition, they envisaged two basic intrusion detection approaches, with the advantage that the fault of one approach was remunerated by the other one. The main objective of this paper was to introduce a novel method which gives the cloud computing system to reach the efficiency of the deployment of the system resources and the vitality of the safety service without any change between them.

At present days, the researchers are more interested in the intrusion detection because it is generally maintaining the security over the network. Here, they represented some of the intrusion detection techniques. An IDS using network profiling and online sequential extreme learning machine (OS-ELM) has proposed by Singh *et al.* [18]. In this technique, alpha profiling is used to decrease the time complexity during irrelevant features were redundant using the collection of filtered, correlation and consistency-based feature selection methods. Beta profiling is used to decrease the size of the training dataset rather than sampling. The estimation of projected technique performance was the standard NSL-KDD 2009 (Network Security Laboratory-Knowledge Discovery and Data mining) dataset. The time and space complexity of their technique is also discussed in this paper. The conclusion of this paper was that the technique was an efficient method for network intrusion method from the achieved results.

The structure of two grain levels of network intrusion detection has proposed by Safaa O. Al-mamory and Firas S. Jassim [19]. The intrusions cannot detect in the normal case. The appropriate IDS level was the coarse-grained to improve IDS performance. When any intrusion was detected by coarse-grained grained IDS, the fine-grained was activating to find probable attack details. For both of these detection levels, the decision tree algorithm was used. The KDD CUP 99 offline dataset and real traffic dataset were used to examine the efficiency of their model. Experimentally, it proves that their model was highly successful in detecting known and unknown attacks, and ′was efficiently modified with packet flow to increase IDS performance.

The feature selection-based hybrid anomaly IDS was using K-means and RBF kernel function has proposed by Ravale *et al.* [20]. One of the main threats to the intrusion detection was the issue of misjudgment, misdetection and lack of real-time response to the attack. Some of the data mining techniques are used for intrusion detection such as clustering, classification and association rule discovery. The projected hybrid method groups data mining techniques like K-means clustering algorithm and RBF kernel function of the support vector machine as a classification module. The main principal of this method was to reduce the number of attributes combined with each data point.

Jiang *et al.* [21] have proposed a basic secure multi-authority ABE against the drawbacks on the decryption algorithm of some existing MA-ABE-based schemes. The attribute keys generating by authorities could be protected with algorithm design and securely issued on normal channels between users and attribute authorities. Then, based on their SMA-ABE, they have proposed a data sharing scheme SDSS-MAC which supports fine-grained access control. Special signature approach was introduced into their scheme to achieve the enforceability and provenance of the outsourced cipher text and the integrity verifiability of the cipher text. Meanwhile, the correctness of decryption transformation on cloud could also be verified. Based on Dolev–Yao model, the SDSS-MAC could withstand multiple collusions such as user-user and user-cloud, and also provides a provably secure way for key distribution and attribute revocation without the assumption of secure communication channels between all involved entities. SDSS-MAC features backward and forward security in context of attribute revocation and could withstand the two vulnerabilities of the revocation security in DAC-MACS that the revoked user has no chance to decrypt any objective cipher text even if he actively eavesdrop to obtain an arbitrary number of non-revoked users' key update keys to collude with other non-revoked users, or directly obtain the cipher text update keys. Finally, the performance simulation have shown that they balance security goals with overhead efficiency and the overall storage, computation, and communication overheads of the SDSS-MAC were superior to that of Ruj and relatively the same as that of DAC-MACS and DCP-ABE.

Hajimirzaei and Navimipour [22] have proposed an IDS using a combination of a multilayer perceptron (MLP) network, and artificial bee colony (ABC) and fuzzy clustering algorithms. Normal and abnormal network traffic packets were identified by the MLP, while the MLP training was done by the ABC algorithm through optimising the values of linkage weights and biases.

Mazini *et al.* [23] have projected a new reliable hybrid method for an anomaly network-based IDS (A-NIDS) using ABC and AdaBoost algorithms in order to gain a high detection rate (DR) with low false positive rate. Here ABC algorithm was utilised for feature selection and AdaBoost was used to evaluate and classify the features.

Pandeeswari and Kumar [24] have proposed an anomaly detection system at the hypervisor layer named Hypervisor Detector that uses a hybrid algorithm which was a mixture of fuzzy C-means clustering algorithm and artificial NN (ANN) to improve the accuracy of the detection system.

Chiba *et al.* [25] have proposed an optimal approach to build an effective A-NIDS based on back propagation NN (BPNN) using back propagation learning algorithm, and employed a novel architecture for that network. Their approach consists firstly of generation of all possible combinations of most relevant values of the parameters included in construction of such classifier, or influencing its performance in anomaly detection, like feature selection, data normalisation, architecture of NN and activation function. Secondly, they have built 48 IDSs corresponding to those combinations.

Iqbal *et al.* in [26], cloud-based attacks and vulnerabilities were collected and classify with respect to their cloud models. They have also presented a taxonomy of cloud security attacks and potential mitigation strategies with the aim of providing an in-depth understanding of security requirements in the cloud environment. They also highlight the importance of intrusion detection and prevention as a service.

3 Proposed method

Cloud computing has generated significant interest in both academia and industry, but it's still an evolving paradigm. One of the most critical concerns of cloud computing is data security. In order to improve the storage security, here we have intended to propose an efficient approach for providing very high secure storage data to the cloud system. In this work, I have planned to detect intrusions using modified ANN (MANN) over cloud data. Here the traditional NN is modified with the help of the optimisation algorithm. The proposed technique used a modified particle swarm optimisation (MPSO) algorithm for weight updation. After verifying the intrusion of the storage system the user wants to store the data to the cloud. To improve the storage security, our proposed method is to encrypt the file by using cryptography method. In our proposed method, a dual encryption algorithm is used for the encryption. When the combination of two algorithms provides more secure than using only one algorithm.

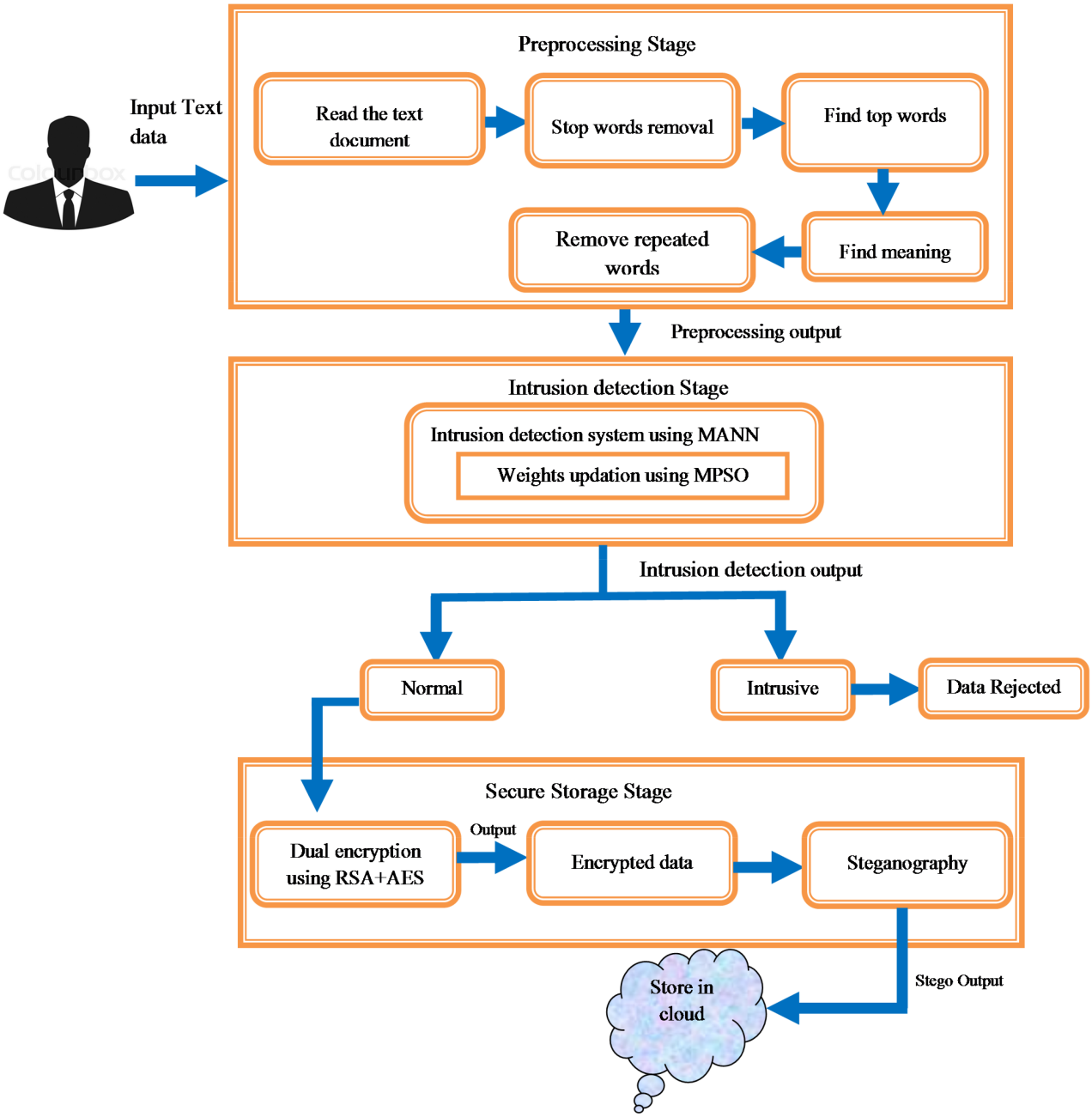


Fig. 1 *Overall block diagram of the proposed method*

And the output of dual encryption technique is more complex, so the hackers are not able to view or modify the data and hence this dual encryption algorithm is more secure. Here RSA and advanced encryption standard (AES) algorithms are used to encrypt the document with high security. To enhance the storage safety of the novel technique, the steganography methods are employed immediately after the encryption approach. And it provides additional security to the data. Finally, the data is stored in the cloud with more secure in our proposed method. In order to verify the security of the proposed system, here we propose some attacks on our proposed method. By means of different security attacks such as Man in Middle (MiM) attack, Denial of service attack, the security of the suggested method will be examined. The overall process of the proposed technique is shown in Fig. 1 and the detail process is illustrated in the further section. In the recommended technique, at first, the input data is preprocessed. After preprocessing, IDS using MANN is utilised in the proposed method. After that encryption process can be performed using a dual encryption algorithm. To improve storage security, steganography methods are employed immediately after the dual encryption. Finally, the encrypted input data is stored in the cloud with more secure. The step by step process is illustrated in the further section.

The proposed method has three main processes namely,

(i) preprocessing,
(ii) intrusion detection,
(iii) secure storage.

3.1 Preprocessing

Initially, the input data is fed to the preprocessing stage, here we are considering input data as a text document. At first, it reads the input text document and then removes the stop words from the text document. Next, find the top words from the text document. After finding the top words, next, we find the meaning of that top word. Finally, we are having both top words and the meaning for each top word. Finally, we remove the repeated words from the document, our proposed technique having the unique words from the input text data. And then the unique words are used for the further process of our proposed technique.

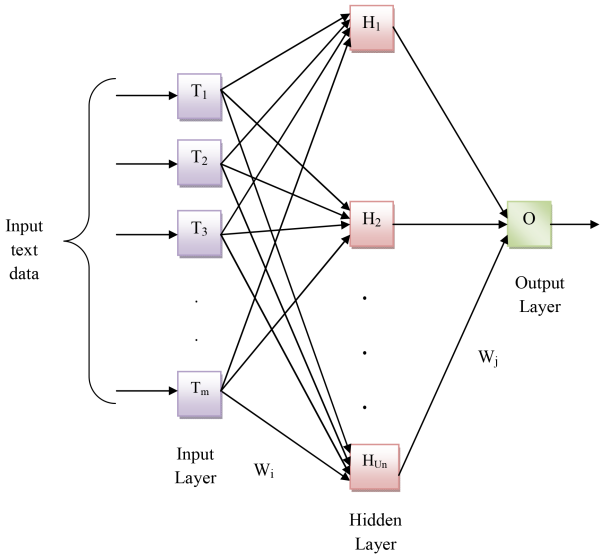


Fig. 2 *ANN structure of the proposed method*

3.2 Intrusion detection using MANN

In our proposed technique use the MANN for intrusion detection. Here the traditional NNs are modified by means of MPSO algorithm. MPSO algorithm is employed to optimise the weights in NN. The main objective of the MANN is categorising the input text data into normal or intrusive. For training purpose back propagation algorithm is used in our suggested technique. In ANN consists of a series of nodes (neurons) which have multiple connections with other nodes. Each connection has a weight associated with it which can be varied in strength, in analogy with neurobiology synapses. The principal with it which an NN operates is relatively simple. ANN consists of three layers such as an input layer, a hidden layer and an output layer. Each neuron in the input layer holds a value so that the input layer holds the input vector. Each of these neurons connects to every neuron in the next layer of neurons. ANN structure consists of an input layer, an output layer and hidden layers between these two layers. The number of these layers is dependent on the problem we are trying to solve, that is basically on the user. The overall structure of ANN is shown in Fig. 2.

MANNs function steps

(1) Fix loads for every neuron's except the neurons in the input layer.

(2) Develop the NN with the input text data as the input units, HU_a hidden units and O as the output unit.

(3) The computation of the proposed bias function for the input layer is

$$X = \beta + \sum_{n=0}^{H_U - 1} \begin{matrix} w_{(n)}T_1(n) + w_{(n)}T_2(n) + \\ w_{(n)}T_3(n) + \cdots + w_{(n)}T_m(n) \end{matrix} \quad (1)$$

In our proposed MANN, the weights are optimised with the help of the MPSO algorithm. The step by step procedure of PSO is illustrated in the below section.

3.2.1 Modified PSO: PSO is a population-based optimisation algorithm modelled. The algorithm of PSO is initialised with a group of arbitrary particles and next searches for optima by revising generations. Each of the particles is flown through the search space enclosing its position altered based on its distance from its own personal best position and the distance from the best particle of the swarm. The presentation of each particle, i.e. how close the particles is from the global optimum, is calculated by means of a fitness function which relies on the optimisation problem. According to PSO, there are two dissimilar kinds of versions are employed. The first is 'individual best' and the second is 'global best'. It is the individual best selection algorithm by evaluating each individual position of the particle to its own best position. It is the global best selection algorithm, which obtains the global knowledge by making the movement of the particles comprises the position of the best particle from the whole swarm. Here the traditional PSO algorithm is modified for better performance. In the proposed MPSO consider the worst cases also for finding the optimal weights. Each particle i, flies in an n-dimensional search space.

cp_i – It symbolises the current position of the ith particle in the search space.

lb_i – It points out the location of the best solution of the ith particle in the search space.

cv_i – It points out the direction for which the particle i will travel (the current velocity).

The step-by-step procedure of MPSO is described in below.

MPSO algorithm steps:

Step 1: Initialise a population of i weight with each weight's position cp_i and velocity cv_i on a problem space of dimension n.

Step 2: Calculate the fitness function for each weight using the following equation:

$$\text{fitness} = \min \sum_{i=1}^{n} \text{MSE} \quad (2)$$

Step 3: Make comparison among the weight's fitness value, cp_{fitness} and weight's *pbest* fitness value lb_{fitness}. If the current fitness value of weight is better than the weight's *pbest* fitness value, then set the *pbest* value into the current position

$$pbest = \begin{cases} lb_i = cp_i \\ lb_{\text{fitness}} = cp_{\text{fitness}} \end{cases} \quad \text{if } cp_{\text{fitness}} > lb_{\text{fitness}} \quad (3)$$

Step 4: Inspect all of the weight's *pbest* fitness value, lb_{fitness} with a value of *gbest*. If the current value, *pbest* is better than the *gbest* value means, then set the *gbest* value into the current weight's array index and value

$$gbest = best\,(pbest) \quad (4)$$

Step 5: Revise the velocity and position of the weights specified as

$$\begin{aligned} cv_i^{\text{new}} = cv_i + \phi_1 \cdot r_1 \cdot (pbest_i - cp_i) \\ + \phi_2 \cdot r_2 \cdot (gbest_i - cp_i) \end{aligned} \quad (5)$$

$$cp_i^{\text{new}} = cp_i + cv_i^{\text{new}} \quad (6)$$

where i is the weight, ϕ_1, ϕ_2 are the learning rates governing the weight towards its best position; r_1, r_2 are the random numbers that are uniformly distributed in the range [0, 1]; cv_i indicates the current velocity; cv_i^{new} indicates the new velocity; cp_i represent the current position; cp_i^{new} represents the new position; $pbest_i$ specifies the current best position; $gbest_i$ specifies the global best position.

Step 6: Repeat step 2, until a better fitness or the maximum number of iterations are met.

Based on the above procedure, we select the optimal weights and then these optimal weights are preceded for further steps in ANN,

(4) The activation function for the output layer is estimated as

$$\text{Active}(X) = \frac{1}{1 + e^{-X}} \quad (7)$$

(5) Recognise the learning error as offered beneath

$$\text{Output}(O) = \text{LE} = \frac{1}{2}\sum_{n=0}^{H_U-1}(D_n - A_{n'})^2 \quad (8)$$

where LE is the learning error rate; D_n is the desired outputs; A_n is the actual outputs.

In the MANN the error should be in minimum value then only the trained ANN is well trained for performing the testing phase. The suggested technique assigns the dual threshold value which satisfies the minimum criteria. Then the threshold value and the result of the NN (O) is compared if the output of MANN is in between the threshold value means then the data is normal otherwise the input text data is intrusive data. After classifying the input text data is normal or intrusive, the normal input text data is carried for the further process.

3.3 Secure storage using dual encryption

For secure storage, the suggested technique uses the dual encryption technique. In the paper, the dual encryption process is done to improve the security level. Here we are using RSA and AES algorithms for dual encryption. Encryption is the process of translating plain text into cipher text. Decryption is the process of converting cipher text back to plaintext. The detail explanation of dual encryption technique is illustrated in the below section.

3.3.1 RSA algorithm for encryption: In the current investigation, the RSA algorithm is elegantly employed to encrypt the text file to furnish security in order that only the appropriate user is competent to access it. In the following section, the RSA algorithm is discussed in detail. The RSA technique proceeds through three fundamental phases as shown below:

1. key generation,
2. encryption,
3. decryption.

The most extensively employed Public Key technique is known as the RSA. The RSA represents fundamentally an asymmetric encryption/decryption technique. The Public key distributed to all the clients through which they are competent to encrypt the text file and private key which is employed for the purpose of decryption is kept confidential and is not disclosed to all. It is invariably dependent on the exponentiation in a restricted field over integers modulo a prime number. The pseudo code for the RSA encryption algorithm is described below.

Pseudo code for RSA encryption:

1. Key generation

   Input: Select the two prime numbers p_n1 and p_n2.

   Output: Public key $\{i, x\}$ and private key $\{d, x\}$.

   Procedure:

   - Establish $x = p_{n1} * p_{n2}$
   - Calculate $\phi(x) = (p_{n1} - 1)*(p_{n2} - 1)$
   - Select an integer (i)

     $\gcd(\phi(x), i) = 1$ and $1 < i < \phi(x)$
   - Calculate $d = e^{-1}(\bmod(\phi(x))$
2. Encryption

   Input: Plaintext message T.

   Output: Cipher text C.

   Procedure:

   - Find the plaintext T, $T < x$.
   - Calculate cipher text

$$C \leftarrow T^i \bmod x$$

3. Decryption

   Input: Cipher text C.

   Output: Plaintext message T.

   Procedure:

   - Find the cipher text C.
   - Estimate plaintext

$$T \leftarrow C^d \bmod x$$

The overall steps include in the proposed RSA algorithm is given below:

Steps

1. The cloud service provider has to furnish or hand on the Public Key (i, x) to the user who intends to store the text file with him.
2. The text file is encrypted and the consequential cipher text (data) C is represented as: $C \leftarrow T^i \bmod x$
3. Then the corresponding cipher text or encrypted text file is now stockpiled with the Cloud service provider.

In order to improve the security level, the proposed method uses another one encryption algorithm after the RSA encryption. Here AES algorithm is used to encrypt the encrypted file from RSA once again. AES is a symmetric block cipher algorithm. The overall explanation of the AES algorithm is shown in the below section.

3.3.2 AES algorithm: AES is a block cipher with a block length of 128 bits [27]. It has three different key lengths: 128, 192, or 256 bits. We propose AES with 512 bit key length. The encryption process consists of ten rounds of processing for 512-bit keys. Except for the last round in each case, all other rounds are identical. The 4×4 matrix of bytes made from 512-bit input block is referred to as the state array. The new algorithm (AES-512) uses input block size and key size of 512-bits which makes it more resistant to cryptanalysis with tolerated area increase. AES-512 will be suitable for applications with high security and throughput requirements and with less area constrains. The different transformation operates on the intermediate results, known as a state; the state is basically in the form of a rectangular array of bytes. Before any round-based processing for encryption can begin, input state is XOR with the first four words of the schedule.

A state of the proposed work is represented as given at the bottom of the page.

A key value of the proposed work is represented as given at the top of the next page.

Encryption

For encryption, each round consists of the following four steps:

- sub bytes
- shift rows
- mix columns
- add round key

E0,0	*E0,1*	*E0,2*	*E0,3*
E1,0	*E1,1*	*E1,2*	*E1,3*
E2,0	*E2,1*	*E2,2*	*E2,3*
E3,0	*E3,1*	*E3,2*	*E3,3*

K0,0	*K0,1*	*K0,2*	*K0,3*
K1,0	*K1,1*	*K1,2*	*K1,3*
K2,0	*K2,1*	*K2,2*	*K2,3*
K3,0	*K3,1*	*K3,2*	*K3,3*

*Sub-bytes operation***:** The sub-bytes operation is a non-linear byte substitution, operates on each byte of the state independently. The substitution table (S-box) is invertible and it is constructed using two transformations.

(i) Take the multiplicative inverse in Rijindael's finite field.
(ii) Apply the affine transformation, they have been documented in the Rijindeal documentation.

We utilise the pre-calculation when the S-box is independent of any input. Then, we substitute each byte of the state in the s-box whose index corresponds to the value in the state $S(i, j) = \text{S-box}[S(i, j)]$.

Shift row operation: In shift row operation, every row of the state is cyclically shifted to the left which depends on the row index.

(i) First row to 0 positions to the left.
(ii) Second row to 1 position to the left.
(iii) Third row to 2 positions to the left.
(iv) Fourth row to 3 positions to the left.

Mix-column operation: The mix-columns transformation operates on the state column-by-column, treating each column as a four-term polynomial. The purpose of this step is to provide diffusion of the bits over multiple rounds. This is achieved by performing multiplication one column at a time. Each value in the column is multiplied against every row value of a standard matrix.

Add round key: In add round key, we apply a round key to the state by bitwise XOR. The round key can be derived from the cipher key using key schedule.

Simply, it can be written as $C_{ij} = E_{ij} \text{XOR } K_{ij}$.

Decryption: In decryption mode, the operations are in reverse order compared to their order in encryption mode. Thus it starts with an initial round, followed by nine iterations of an inverse normal round and ends with an add round key. An inverse normal round consists of the following operations in this order: Add Round Key, Inv Mix Columns, Inv Shift Rows, and Inv Sub Bytes. From that process, we encrypt the text file in an effective manner and then the dual encrypted text file is fed to the steganography process to improve the security level of our proposed techniques.

3.3.3 Steganography: The dual encrypted data emerges as the input for the steganography approach. To hide the encrypted data, our method first gets the user key and then finds which position of the encrypted data is selected to hide. Based on the user key, we have to select the position and replace the data with their corresponding information. After this process, we got the stego data. Then we have to store the stego data into cloud with high secure. In order to verify the security of the proposed technique, various security attacks are used. The security attacks used in our work are illustrated in the below section.

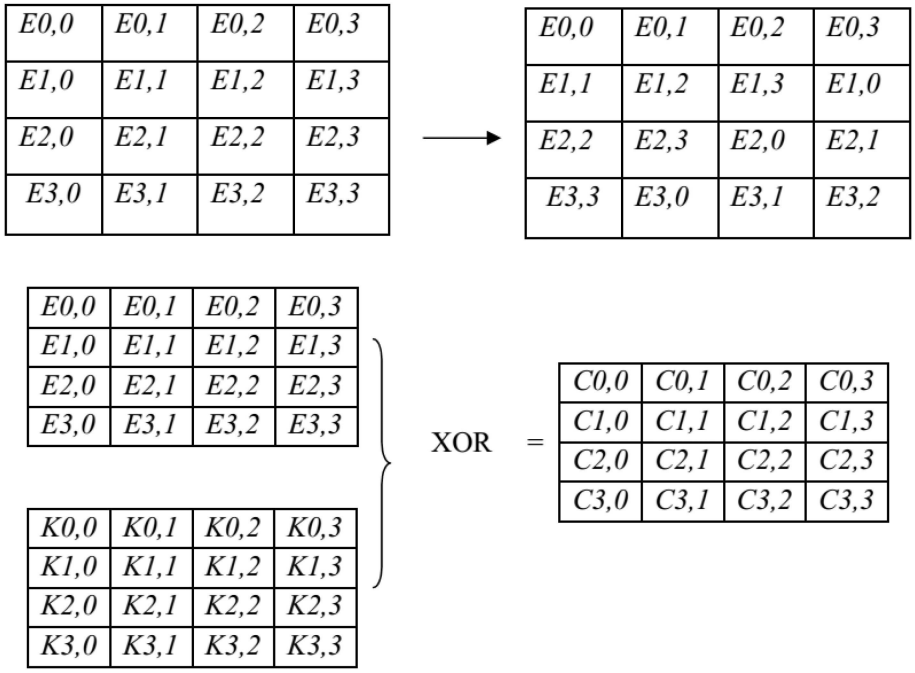
3.4 Various security attacks

The suggested method employs different security attacks for the security purpose in a video sequence such as Denial of service (*DoS*) attack and MiM Attack. It is elucidated underneath

DoS attack: DoS attacks have turned out to be a most significant hazard to current computer networks. DoS or distributed DDoS attack is an effort to make a machine or network resource connected to its future users. DoS attacks were instigated from distributed attacking hosts. DoS attack is instigated in two phases. Initially, an attacker constructs an attack network, which is distributed and contains thousands of compromised computers. After that, the attacking hosts flood a great volume of traffic towards victims either under the command of the attacker or routinely.

MiM Attack: Inside cryptography the attacker secretly relays and possibly alters the communication between two parties, in man-in-the-middle attack who believes they are directly communicating with each other. A man-in-the-middle attack; without either outside party knowing until it is too late, it allows a malicious actor to intercept, send and receive data meant for someone else, or not meant to be sent at all.

In this paper, the execution time and memory usage are analysed which is minimised for proving the quality of the proposed method. From the result analysis, the proposed method



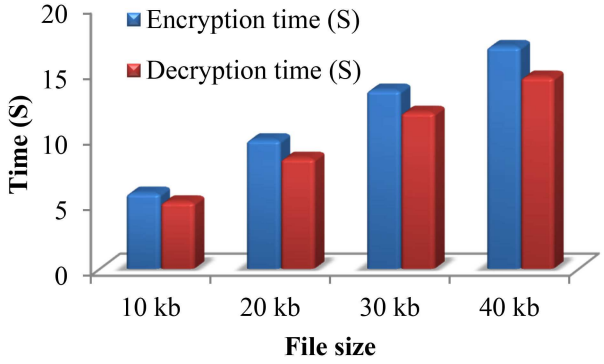


Fig. 3 *Encryption and decryption time by varying the file size*

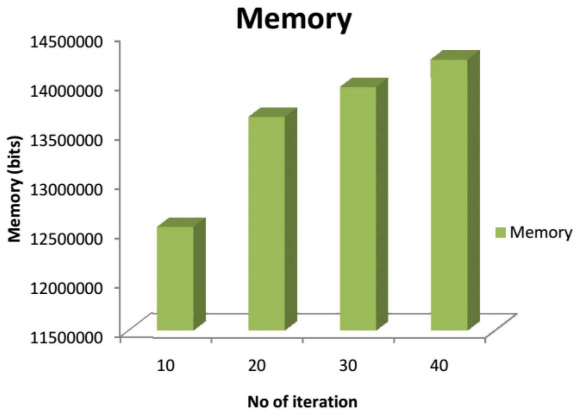


Fig. 4 *Memory value for the proposed technique*

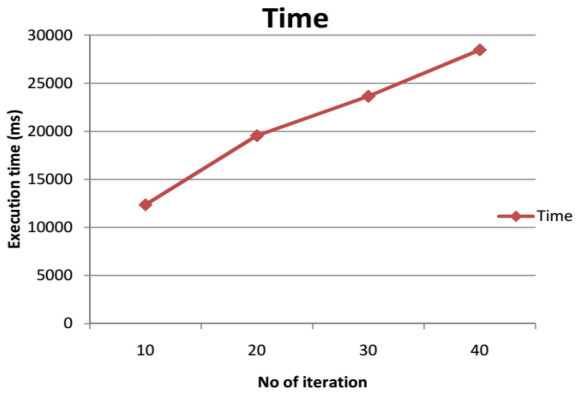


Fig. 5 *Execution time for the proposed technique*

proves their effectiveness analysed with the factors and compared with the existing technique.

4 Results and discussion

This section gives the detailed view of the result that is obtained by our proposed method of secure data storage and intrusion detection in the cloud using MANN and dual encryption with steganography which is performed in the working platform of JAVA with Cloud Sim. The dual encryption algorithm is used for providing secured storage in this method, here RSA and AES algorithms are used for dual encryption. In order to improve the security in storage, steganography is applied after dual encryption. The dataset used in our proposed technique is news20 dataset. This data set consists of 20,000 messages taken from 20 newsgroups. Here one thousand Usenet articles were taken. The datasets are taken from UCI machine learning repository. The dataset is available from https://archive.ics.uci.edu/ml/datasets/Twenty+Newsgroups. The experimental result and the performance of the proposed method are given below in detail.

Table 1 Encryption and decryption time for various file size

File size, kb	Encryption time, s	Decryption time, s
10	5.796	5.123
20	9.864	8.457
30	13.654	11.965
40	16.985	14.652

Table 2 Memory value and execution time of the proposed method

No of iterations	Memory (bits)	Time, ms
10	12,544,752	12,364
20	13,655,423	19,554
30	13,956,542	23,645
40	14,236,274	28,469

4.1 Performance analysis

The performance analysis of our proposed technique is shown in the below section. Table 1 shows the various file sizes and the corresponding encryption and decryption time. In our method, we take the file size as 10, 20, 30 and 40 kb. To encrypt the file contain 10 kb it takes 5.796 s for dual encryption and hence if the file size varies the time consumption to encrypt the file also vary.

To encrypt the 10 kb file our method takes 5.796 s for encryption and 5.123 s for decryption. Varying the file size like 20, 30 and 40 kb encryption time and decryption time is also varying. Here 9.864 s to take to encrypt the 20 kb file and 8.457 s obtain to decrypt the same file size. The graphical representation of the proposed encryption and decryption time by varying the file size is shown in Fig. 3.

In our proposed technique, Table 2 shows the overall memory value and execution time of the proposed method. In Table 2, we vary the number of iteration and evaluate the memory value and execution time.

Figs. 4 and 5 show the graph value for the number of iterations with memory value and execution time. It is plotted in the below section.

The overall memory value of the proposed method achieves 13,598,247.75 bits by varying the number of iterations. The overall execution time of the proposed methods achieves 21,008 ms. Fig. 5 shows the execution time for the proposed method by varying the number of iterations.

Fig. 6 shows the fitness value of the proposed technique. The text with the lowest error rate is chosen as the best fitness value using the MPSO. Here the fitness value gets decreases as the number of iterations gets increases. This gradual decrease in fitness value finally achieves the best fitness value with a very low error rate.

The overall classification accuracy of the proposed MANNs-based back propagation algorithm is tabulated in Table 3. Here the proposed MANN achieves 91.25% of the accuracy value. It is tabulated in below:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

where TP (true positive) is the proportion of normal data that are correctly diagnosed as normal; FP (false positive) is the proportion of intrusive data that are wrongly diagnosed as normal; FN (false negative) is the proportion of normal data that are wrongly diagnosed as intrusive; TN (true negative) is the proportion of intrusive data that are correctly diagnosed as intrusive.

4.2 Effectiveness of the proposed technique

In this section, the effectiveness of the proposed technique is compared with the other existing technique. The detail explanation is illustrated in the further section.

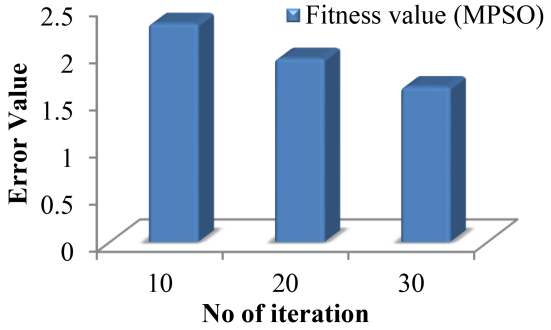


Fig. 6 *Fitness value for the proposed technique*

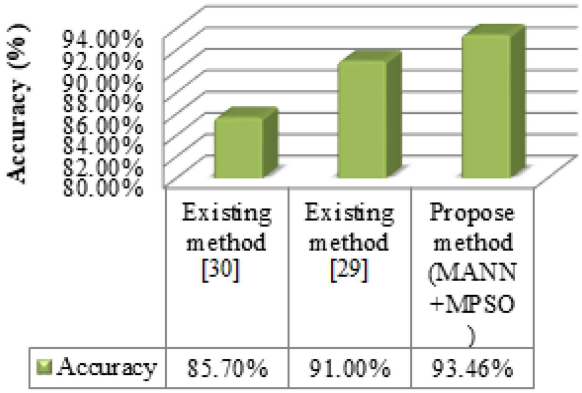


Fig. 7 *Comparison graph for accuracy value*

4.2.1 Comparison analysis for intrusion detection: The accuracy value plays the major factor in IDS. It is important for the method to provide the high accuracy value in order to serve as the best method and Fig. 7 shows the comparative analysis for accuracy value using the existing intrusion detection method. Here we are considering existing IDS as traditional NN and existing [28], intrusion detection using genetic algorithm [29] for comparison.

From the graph, it is clear that the accuracy value for the existing method [29] is 85.7% and for the existing method [28] attains 91% while for the proposed method achieves 93.46%. Since the accuracy value for the proposed method is very high it seems to be better than the existing methods.

4.2.2 Comparative analysis for various attacks: In order to verify the security of the proposed technique, various security attacks are used. Here we are using MiM the attack and DoS attack. For an encryption method to be much effective the impacts of these attacks on the data should be low so that it provides more security and access to data only to the authorised person. Our proposed method serves its best irrespective of all attacks on comparing to the existing methods. Table 4 shows the comparative analysis of proposed and existing methods for various attacks like MIM and DoS attack.

The attack percentage is greater for the existing methods and is lower for the proposed methods. Thus from the above table is clear that the proposed method offers better protection to data irrespective of various attacks compared to existing methods.

4.2.3 Comparison analysis using key breaking time: The key breaking time is an essential one to ensure the duration taken for the hackers to hack the key and get access to the secured data. As the duration for key breaking gets decreases the protection of the data gets increases. Our proposed method has higher key breaking time when compared to the existing methods. Hence the proposed method assures high security. Table 5 shows the key breaking time comparison for the existing and proposed method.

From the above table, it is clear that that the duration taken for key breaking in the proposed method is more than the duration taken for key breaking in the existing method. To break the key value for 10 kb file size the proposed technique tries 128 times but the existing method tries 120 times. For 20 kb file size the suggested technique tries 132 times to break the key value the existing method tries 123 times, which is the minimum number of times when compared to our implemented technique. Likewise, we are evaluating the key braking time for the proposed technique is 112 and 136 times but the existing method tries 95 and 129 times to break the key value fo4 30 and 40 kb. Thus the proposed method serves the best security. From all the above results it is clear that the proposed method has high intrusion detection accuracy and high security than the existing methods.

For evaluating the proposed performance, here we are considering the KDD dataset and the performance is compared with recent research work. Here we are considering the existing work as fuzzy C-means and ANN [24] and hybrid approach [23]. The results are tabulated in Table 6.

And also, here we are comparing the performance of the proposed technique by means of kappa statics, mean absolute error and root mean square error value. The results are tabulated in Fig. 8.

The value of the average data rate for encryption is low compared to that of DES for both encryption and decryption. Here, the developed hybrid technique will consume less memory for encryption and decryption operation. The average data rate indicates the amount of information or data encrypted or decrypted per second. The results are tabulated in Table 7.

From the results, it is clear that our recommended technique attains better performance when compared to the other methods.

Table 3 Accuracy value of proposed method

Classifier	Accuracy value for testing
MANN (ANN + MPSO)	93.46%

Table 4 Comparison for attack percentage using MIM and DoS attacks

File size, kb	MIM		DoS attack	
	Proposed method (RSA + AES), %	Existing method (RSA), %	Proposed method (RSA + AES), %	Existing method (RSA), %
10	7.3	11.25	8.2	10.92
20	8.2	10.9	9.6	10.1
30	10.2	12.75	10.3	12.2
40	11.6	13.82	10.8	12.45

Table 5 Key breaking time comparison for proposed and existing methods

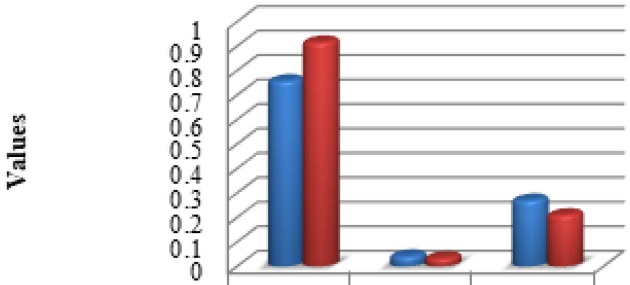
File size, kb	Proposed method (RSA + AES)	Existing method (RSA)
10	128	120
20	132	123
30	112	95
40	136	129

Table 6 Comparison of intrusion DR

Classifier	Intrusion DR, %
MANN (ANN + MPSO)	93.46
FC-ANN [24]	91.25
hybrid approach [30]	93.29

5 Conclusion

Text data-based secure storage and IDS was proposed in this paper. By means of the Cloud Sim, the proposed technique was performed. At first, the input text document is preprocessed and then check the user text data is normal or intrusive with the help of MANN, in which the weights are optimised by means of MPSO.



	Kappa statics	MAE	RMSE
FC-ANN [25]	0.758	0.042	0.27
Proposed Approach	0.916	0.034	0.21

Fig. 8 *Performance comparison*

Table 7 Data rate comparison of proposed method

Performance criteria	Hybrid (DES-RSA) [31]	Proposed (AES-RSA)
average data rate of encryption, kb/s	2718.14	1725.32
average data rate of decryption, kb/s	3827.01	1951.98

Then the encryption process is carried out by means of dual encryption technique. Here RSA and AES algorithms are used to perform the dual encryption. In order to improve the secure storage data in the cloud we used steganography method after the encryption. The performance of the method was accessed, analysed and compared with other modern methods. In order to verify the security of the proposed system use various security attacks. Based on the results, we obtain that the proposed technique achieves better security and high classification rate when compared with other existing techniques. The use of dual encryption and steganography model provided more complex to the proposed model and the encrypted data is hard to reveal. So that the proposed structure is more secure. In future, the researcher uses various encryption methods and intrusion detection techniques to achieve higher excellence in performance.

6 References

[1] Yao, J.T., Zhao, S.L., Saxton, L.V.: 'A study on fuzzy intrusion detection'. Data Mining, Intrusion Detection, Information Assurance and Data Networks Security, 2005, Vol. 5812, pp. 23–31

[2] Naidu, N., Dharaskar, R.V.: 'An effective approach to network intrusion detection system using genetic algorithm', *Int. J. Comput. Appl.*, 2010, **1**, (3), pp. 26–32

[3] Allen, J., Christie, A., Fithen, W.: 'State of the practice of intrusion detection technologies', Technical Report, CMU/SEI-99-TR-028, 2000

[4] Dasarathy, B.V.: 'Intrusion detection', *Inf. Fusion*, 2003, **4**, (4), pp. 243–245

[5] Campos, M.M., Milenova, B.L.: 'Creation and deployment of data mining-based intrusion detection systems in oracle database l0 g'. Proc. Fourth Int. Conf. on Machine Learning and Applications, 2005, p. 8

[6] Sarmah, A.: 'Intrusion detection systems: definition, need and challenges', White Paper from SANS Institute, 2001

[7] Kozushko, H.: 'Intrusion detection: host-based and network-based intrusion detection systems', White Paper from Independent Study, September 11, 2003

[8] Cha, B., Park, K., Seo, J.: 'Neural networks techniques for host anomaly intrusion detection using fixed pattern transformation'. ICCSA 2005, 2005, (LNCS, 3481), pp. 254–263

[9] Adetunmbi, A.O., Falaki, S.O., Adewale, O.S., *et al.*: 'Network intrusion detection based on rough set and k-nearest neighbour', *Int. J. Comput. ICT Res.*, 2008, **2**, (1), pp. 60–66

[10] 'Network- Vs. Host-Based Intrusion Detection: A Guide to Intrusion Detection', White Paper from ISS Internet Security Systems, Atlanta, GA, October 2, 1998

[11] Farid, D.M., Rahman, M.Z.: 'Anomaly network intrusion detection based on improved self adaptive Bayesian algorithm', *J. Comput.*, 2010, **5**, (1), pp. 23–31

[12] Zainal, A., Maarof, M.A., Shamsuddin, S.M.: 'Research issues in adaptive intrusion detection'. Proc. of the 2nd Postgraduate Annual Research Seminar (PARS'06), Faculty of Computer Science & Information Systems, Universiti Teknologi Malaysia, 2006, pp. 24–25

[13] Sahu, V., Dubey, B.: 'An analysis of current security issues and solutions for cloud computing', *Int. J. Res. Appl. Sci. Eng. Technol. (Ijras Et)*, 2014, **2**, pp. 365–372

[14] Moghaddam, F.F., Moghaddam, S.G., Rouzbeh, S., *et al.*: 'A scalable and efficient user authentication scheme for cloud computing environments'. Proc. of IEEE Region of Symp., 2014, pp. 508–513

[15] Anil Kumar, K.S., Nanda Mohan, V.: 'Novel anomaly intrusion detection using neuro-fuzzy inference system', *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, 2008, **8**, (8), pp. 6–11

[16] Zarrabi, A., Zarrabi, A.: 'Internet intrusion detection system service in a cloud', *Int. J. Comput. Sci.*, 2012, **9**, (5), pp. 308–315

[17] Mathew, S., Jose, A.P.: 'Securing cloud from attacks based on intrusion detection system', *Int. J. Adv. Res. Comput. Commun. Eng.'*, 2012, **1**, (10), pp. 753–759

[18] Singh, R., Kumar, H., Singla, R.K.: 'An intrusion detection system using network traffic profiling and online sequential extreme learning machine', *Expert Syst. Appl.*, 2015, **42**, (22), pp. 8609–8624

[19] Singh, R., Kumar, H., Singla, R.K.: 'On the designing of two grains levels network intrusion detection system', *J. Karbala Int. J. Mod. Sci.*, 2015, **1**, (1), pp. 15–25

[20] Ravale, U., Marathe, N., Padiya, P.: 'Feature selection based hybrid anomaly intrusion detection system using K means and RBF kernel function', *Procedia Comput. Sci.*, 2015, **45**, pp. 428–435

[21] Jiang, R., Wu, X., Bhargava, B.: 'SDSS-MAC: secure data sharing scheme in multi-authority cloud storage systems', *J. Comput. Secur.*, 2016, **62**, pp. 193–212

[22] Hajimirzaei, B., Navimipour, N.J.: 'Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm', *ICT Express*, 2018, pp. 1–6

[23] Mazini, M., Shirazi, B., Mahdavi, I.: 'Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and AdaBoost algorithms', *J. King Saud Univ. Comput. Inf. Sci.*, 2018, pp. 1–13

[24] Pandeeswari, N., Kumar, G.: 'Anomaly detection system in cloud environment using fuzzy clustering based ANN', *Mobile Netw. Appl.*, 2016, **21**, (3), pp. 494–505

[25] Chiba, Z., Abghour, N., Moussaid, K., *et al.*: 'A novel architecture combined with optimal parameters for back propagation neural networks applied to anomaly network intrusion detection', *Comput. Secur.*, 2018, **75**, pp. 36–58

[26] Iqbal, S., Kiah, M.L.M., Dhaghighi, B., *et al.*: 'On cloud security attacks: a taxonomy and intrusion detection and prevention as a service', *J. Netw. Comput. Appl.*, 2016, **74**, pp. 98–120

[27] Pitchaiah, M., Daniel, P.: 'Implementation of advanced encryption standard algorithm', *Int. J. Sci. Eng. Res.*, 2012, **3**, (3), pp. 1–6

[28] Anitha Ruth, J., Sirmathi, H., Meenakshi, A.: 'Steganography based secure data storage and intrusion detection for cloud computing using signcryption and artificial neural network', *Res. J. Appl. Sci., Eng. Technol.*, 2016, **13**, (5), pp. 354–364

[29] Hoque, M.S., Mukit, M., Bikas, M., *et al.*: 'An implementation of intrusion detection system using genetic algorithm', *Int. J. Netw. Secur.Appl. (IJNSA)*, 2012, **4**, (2), pp. 109–120

[30] Guo, C., Ping, Y., Liu, N., *et al.*: 'A two-level hybrid approach for intrusion detection', *Neuro Comput.*, 2016, **214**, pp. 391–400

[31] Adedeji Kazeem, B., Akinlolu, P.: 'A new hybrid data encryption and decryption technique to enhance data security in communication networks: algorithm development', *Int. J. Sci. Eng. Res.*, 2014, **5**, (10), pp. 804–811