International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 6 (2020), pp. 1441-1447


# An Optimized Approach-Based Machine Learning to Mitigate DDoS Attack in Cloud Computing

**Ahmed Saeed Alzahrani**

*Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia.*

**Abstract**

With the emerging growth of cloud computing technology and on-demand services, users can access cloud services and software freely and applications based on the "pay-as-you go" concept. This innovation reduced service costs and made them cheaper with high reliability. One of the most significant characteristics of the cloud concept is on-demand services. One can access the applications of cloud computing at any time at a much lower cost. In addition to providing cloud users with much-needed services, the cloud also gets rid of security concerns which are not tolerated by the cloud. One of the most security problems in the cloud environment is Distributed Denial of Service (DDoS) attack that are responsible for overloading the cloud servers. This paper highlights a prevention technique (CS-ANN) which detect the DDoS attack and makes the server side more sensitive by integrating a Cuckoo Search (CS) approach with the Artificial Neural Network (ANN) approach. The cloud user features, along with the attacker features, are optimized using CS as a nature-inspired approach. Later on, these optimized features are passed to the ANN structure. The trained features are stored in the database and used during testing process to match the test features with the trained features and hence provide results in terms of attacker and normal cloud users. The test results of CS-ANN show a True Positive Rate (TPR), False Positive Rate (FPR) and detection accuracy of 0.99, 0.0105 and 0.9865% respectively. The proposed approach outperforms in contrast to the other two state-of-the-art techniques.

**Keywords:** Cloud Computing, Distributed Denial of Service, Cuckoo Search, Artificial Neural Network.

## 1. INTRODUCTION

Cloud computing is one of the most renowned services in today's technological world. In these services two elements, namely service providers and service users, hold the most important position. Cloud not only offers large storage space but is also bestowed with enormous computing power to support numerous web-based services. Salient features and services offered by cloud computing environment are illustrated in Figure 1.

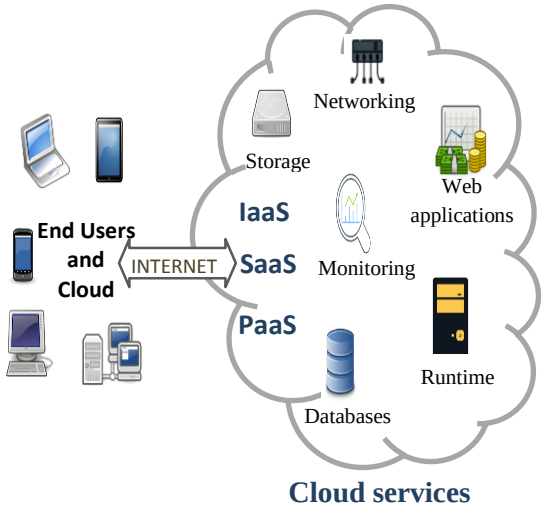


**Figure 1:** Cloud Computing Architecture

Cloud Service Providers (CSPs) offer Virtual Machines (VMs) that support a large number of applications for multiuser sharing. VM utilization in cloud computing has been very advantageous to deliver a flexible, cost effective, expandable, interoperable and consistent interface to a large section of end users [1]. Popular services provided by cloud computing are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) [2] and Software-as-a-Service (SaaS) [3]. However, the cloud environment is highly susceptible to various types of intrusions. Among them, Distributed Denial-of-Service (DDoS) attack is the most frequently observed threat that significantly challenges the quality of service in cloud computing environment. Commonly five types of DDoS attacks - namely SYN flooded attack, Network Type Protocol (NTP) amplification, Ping-of-Death (PoD), User Datagram Protocol (UDP) flood, Hyper Text Transfer Protocol (HTTP) flood and Zero-day attacks - have been observed. Kaspersky lab reports have shown that UDP and Transmission Control Protocol (TCP) flooding has decreased from 31.1% and 8.4% to 8.9% and 3.3% within a period of 3-6 months (Figure 2). However, the SYN (SYNchronization) request packet, HTTP and Internet Control Message Protocol (ICMP) flooding have increased by 25.8%, 1% and 0.5% according to the fourth quarter report of 2018 and first quarter report of 2019 [4] [5].

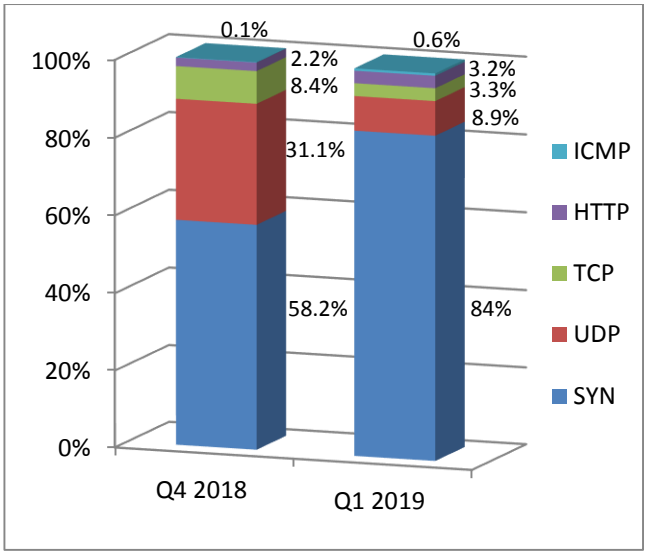


**Figure 2:** Changing trends in the type of DDoS attacks

The majority of the existing techniques were developed to address the issue while taking advantage of Neural Networks (NN), machine learning approaches, bioinspired algorithms, and so on. However, they have resulted in overheads to offer defence against a DDoS attack without meeting the designed level of accuracy and alarm rate [6]. It is therefore highly recommended to design and deploy strategies to successfully defend the cloud computing environment. It has also been observed that existing research lacks consistent measures to detect vivid patterns adopted by DDoS attack [7]. In this context, the author has proposed an improved strategy to offer a secure cloud computing environment.

## 2. LITERATURE REVIEW

Security has been the most important aspect of cloud computing services. In this regard, [8] had conducted a comprehensive survey to summarize the published detection and preventive measures to lessen the instances of DDoS attack. They concluded that multi-layered defence strategies, along with highly accurate auto-scaling decisions, are the necessities of current defence measures in addition to guiding for cyber security. [9] had presented a generalized design of denial of service and DDoS attacks in cloud computing environment along with the available defensive mechanisms. They had discussed rising challenges along with developed tools and devices to tackle the situation. [10] discussed the salient features of Software Defined Networking (SDN) and proposed a flow-table based sharing technique in order to secure an SDN based cloud computing platform from DDoS attack. Simulations had shown that the designed approach had significantly increased the resistance against DDoS attack. [11] proposed a statistical approach as multilayer attack detecting strategy that took advantage of attack variation with the help of a chi square test. In the process, they enlisted the sources of DDoS attack to filter-out the attack traffic in order to secure the cloud environment. Experimental evaluation was performed in terms of throughput, packet count, burst rate, packet length and time consumed round the trip. A multiple attribute-based mechanism was implemented by [12] to decrease the incidence of DDoS attack. The authors had proposed two payment mechanisms in addition to attack detection strategies. Here, legitimate users were charged based on critical value. However, malicious users were charged based on differential pricing. Extreme simulations had shown that the proposed design outperformed by significantly decreasing the maliciousness of the cloud environment. [13] had postulated an anomaly-based detection technique for securing cloud computing environment against Distributed Denial-of-Service attack. The feasibility of the anomaly-based design was evaluated using simulation tests that demonstrated the effectiveness of the proposed design, reflected by low false negative detections (<2%) with zero false positive detections. [14] had proposed a parallel and multi-stage security mechanism (PMSSM) that consists of intrusion identification and authentication along with encryption strategies. This multi-staging security measures significantly increases the probability of intrusion detection to enhance the data security of the cloud environment. [15] addressed security concerns by authenticating the data reliability by involving conditional proxy re-encryption along with authorized third-party auditors into the design. In the process, metadata is generated that is stored along with the data file to aid with auditing to address remote data storage issues. The proposed design not only lowered the computational cost of metadata generation but also governs the number of times an audit request could be addressed. An Artificial Immune System was employed by [16] to lessen the DDoS attacks pertaining to the cloud environment. The technique mimics the human biological resistance mechanism to identify the attacks. Experimental evaluation was performed using a KDD cup 99 dataset that exhibited a low false alarm rate along with DDoS attack accuracy of 96.56%.

## 3. PROPOSED SECURITY ALGORITHM

In this research, a CS-ANN security algorithm against DDoS attack has been presented using a natural-inspired Cuckoo search approach with ANN as an artificial intelligence approach. The designed model is shown in Figure 3.

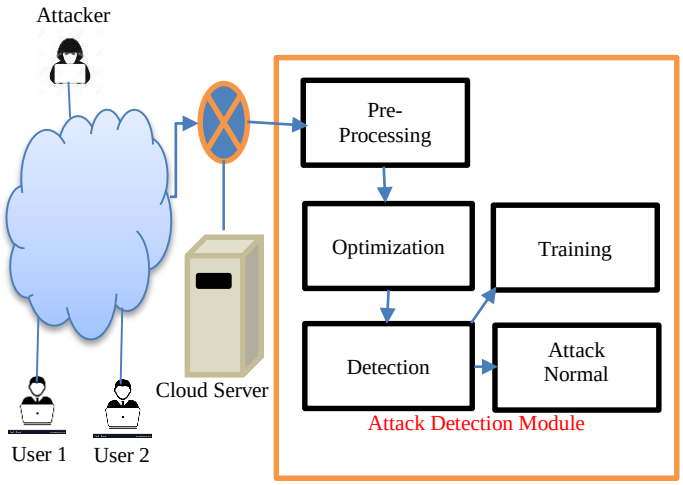


**Figure 3:** The CS-ANN Proposed Model

The users access the services from the cloud server through the router, which is again connected to the detection module. Using this module, the system is able to observe the data traffic flow through the router. The entire process mainly goes into two phases training and validation, which are described below.

### 3.1 Pre-processing

The foremost step of the work is to train the system as per the desired samples or the properties of the users as well as the DDoS attacker. Before training, the pre-processing of the data has been performed. The purpose of pre-processing is to transform the incoming data traffic into a meaningful form, which can be later used by the classifier. Initially, the relevant features from the raw data are extracted and the features with symbolic values are converted into numeric values having values lying between [0,1].

### 3.2 Feature optimization (CS)

Now, the extracted features are optimized using a Cuckoo Search (CS) algorithm. CS inspired by the cuckoo bird, who lay eggs in others' nests to be hatched and looked after. This scheme is used to determine optimal function. The bird lays its eggs in the other birds' nests, which increases the adult cuckoo's generation, but sometimes the eggs are distinguished by the other bird and then these eggs are removed from the nest. If the eggs are not removed, then that area is considered as optimized area by the cuckoo [17].

The cuckoo, which survives after laying the eggs in the nests of other birds can be moved by a distance of maximum gain, which is distance travelled by the cuckoo to lay eggs. In a similar way, the process of laying eggs continues until maximum gain value is achieved.

CS is one of the fastest algorithms and is used to solve optimization problems by selecting appropriate features of both normal and attacker users in terms of CPU utilization, RAM, network and system utilization. Initially, a set of users as well as attacker features are selected on a random basis and are represented in the form of array, known as habitat, as represented by equation (1).

$$Habitat = [x_1, x, x_3, \dots\dots\dots\dots x_{nvar}] \quad (1)$$

After initialization, the user and attacker properties are calculated using equation (2)

$$Gain\ value = f_p(habitat f_p(x_1, x_2, x_3, \dots\dots\dots\dots x)) \text{profitvalue} = \mathrm{f_p}(\text{habitat} = \mathrm{f_p}(\mathrm{y_1}, \mathrm{y_2}, \mathrm{y_3}, \dots\dots\dots\dots \mathrm{y_{nvar}})) \quad (2)$$

Maximum gain value has been obtained using equation (2) and the eggs that have been generated by the cuckoo contain lower and upper limit values. The range of laying eggs in the host of others is being calculated by using equation (3)

$$R_{L\_egg} = \frac{\beta\ (I_{L\_eggs})}{Tot_{L\_eggs}(a_{high} \times a_{low})} \quad (3)$$

The concept of laying egg and selecting nest based on the distance is represented by Figure 4.

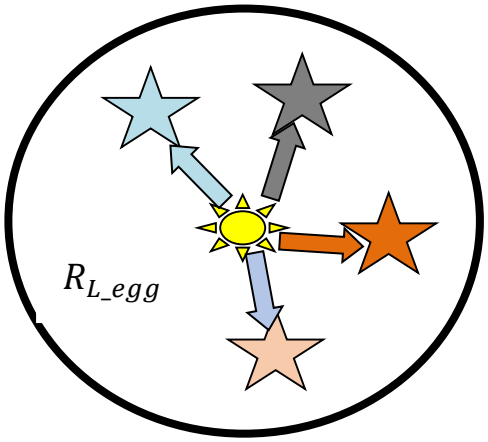


**Figure 4:** Egg laying range [18]

The initial cuckoo bird is represented by a yellow star and the range covered by this cuckoo bird is denoted by distinct stars represented by different colours. It is believed that a single egg is laid by each individual cuckoo and then the gain is determined as per the equation (2). If the cuckoos' number exceeds the highest possible number, the low-earning cookies are emptied so that one can reach the maximum number of cookies. For this process, K means the clustering process has been used. In this way a cuckoo with maximum value has been chosen and used to resolve the problem. Based on the above-mentioned process, the CS algorithm is written as below.

**Algorithm: Cuckoo Search Algorithm**

1. Start.
2. Initialize cuckoo habitat randomly.
3. Defined number of eggs for each cuckoo.
4. Defined range covered by each cuckoo as calculated by equation (3).
5. Let cuckoos lays eggs in the defined range.
6. Remove those eggs that are identified by the host bird.
7. Let eggs survive and be contributed to in the chick's generation.
8. Calculate the habitat of newly produced cuckoos.
9. Decide lower and higher limit and remove/kill those eggs that alive in the poorest habitats.
10. Find best cluster to choose best habitat.
11. Allow newly grown cuckoos to move towards the best habitat.
12. If desired properties of users are obtained then stop, otherwise; repeat the above process.

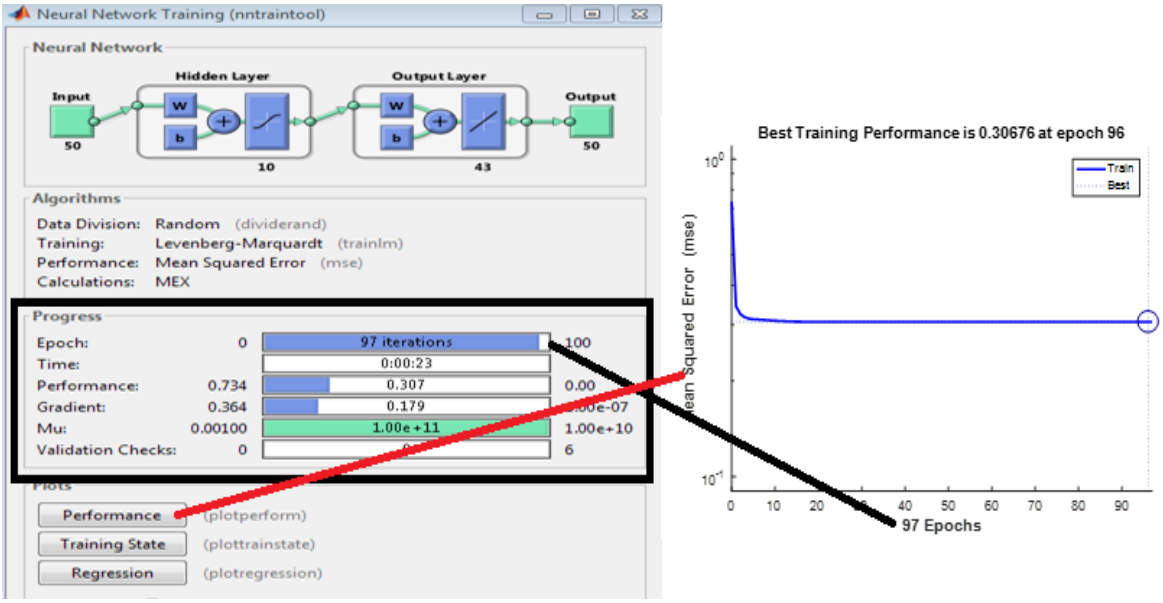


**Figure 5:** Trained ANN structure with performance in terms of MSE

### 3.3 Training using ANN

Based on the optimized user/attacker features, ANN can be used for training. ANN is defined as a computational system inspired by the biological human brain. ANN consists of several single units, artificial neurons, connections among coefficients, also termed as weights, that involve the neural structure. The feature of a biological neural network is attributed to its structure along with functions and the fundamental units of the network, known as a neuron or a nerve cell. The neural network is composed of hidden layers including input as well as output layers. This neural network technique is most suitable for a human programmer to solve difficulties from complex or large problems [19]. Figure 5 shows the trained structure with obtained Mean Square Error (MSE) value.

A total of 50 input user features are passed to the ANN structure, which is being modified in the hidden layer using 10 neurons. After modification, there are 43 features have appeared at the output as normal and the remaining seven are considered to be attacker node's properties. In this way, the system is trained and obtained as normal, as well as the attacker features being stored into the database into two different categories. Later, during the testing process, these have been used to match the upcoming data features with the sorted features and hence detect the presence of a DDoS attack. The designed algorithm is written as the following:

**Algorithm: Artificial Neural Network**

**Input:** Pass optimized features as training data (T) and initialize Neurons (N)

**Output:** Attacker/Normal Cloud User

**Training:**

**1 Initialize the ANN parameters:**

– Number of Neurons (N).

– Number of Epochs (E).

– Performance parameters of training: MSE.

– Techniques: Levenberg Marquardt.

– Data Division: Random.

**2 for i = 1 → T.**

**3 if T belongs to Normal cloud user properties.**

**4** Group (1) = features of training data traffic as per the normal user.

**5 else.**

**6** Group (2) = Properties of training data related to the attacker user.

**7 end.**

**8 end.**

**9** Initialized the ANN using training data and group. **10** Net = Newff $(T, G, N)$.

**11** Set the training parameters according to the requirements and train the system.

**12** Net = Train (net, training data, group).

**Testing:**

**13** Test data = Optimized properties of the present cloud user.

**14** Affected server = simulate (net, test data).

**15 Returns;** normal or attacker cloud user.

**16 end.**

## 4. EXPERIMENTAL RESULTS AND DISCUSSION

The experiments have been conducted using Intel processor core i3@3.60GHz with 4 GB of RAM and a Windows 10 operating system has been used to perform the experiments. The experiments were conducted in MATLAB. The effectiveness of DDoS detection in cloud network should have a high detection accuracy, including high True Positive Rate (TPR) value and low False Positive Rate (FPR) value. So, the performance of the proposed approach has been analysed based on the following parameters:

$$TPR = \frac{\sum Correctly\ Identified\ features}{\sum Positive\ Instances} \quad (4)$$

$$FPR = \frac{\sum Incorrectly\ Identified\ features}{\sum Positive\ Instances} \quad (5)$$

$$Detection\ Accuracy = \frac{Correctly\ Identified}{Correctly\ identified + Incorrectly\ rejected} \quad (6)$$

**Table 1:** True Positive Rate (TPR)

| Number of iterations | Using CS | Using CS-ANN |
|---|---|---|
| 20 | 0.852 | 0.985 |
| 40 | 0.862 | 0.987 |
| 60 | 0.871 | 0.989 |
| 80 | 0.879 | 0.990 |
| 100 | 0.883 | 0.995 |
| 120 | 0.881 | 0.994 |
| 140 | 0.886 | 0.989 |

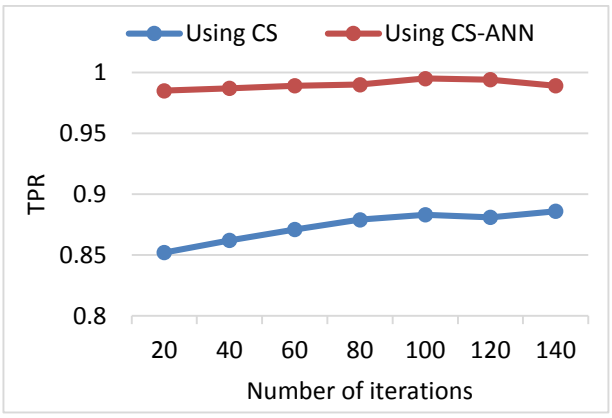


**Figure 6:** True Positive Rate (TPR)

Figure 6 shows that the TPR has been analysed for the proposed model and examined in two scenarios based on Table 1. First, it was looked at when only an optimization scheme is used for detection of DDoS attacker cloud users, and secondly when CS-ANN is used. The analysed values for both CS and CS-ANN are examined with a number of iterations, varied from 20 to 140, and increases in steps of 20. The results show that, when using CS-ANN, the identification of the attacker and normal user have been performed with high accuracy. This is because the CS algorithm accurately distinguished the cloud user features and both the normal and attacker user. Therefore, these highly optimized features help to train the cloud network with minimum MSE. The average TPR examined for the proposed work is 0.99.

**Table 2:** False Positive Rate (FPR)

| Number of iterations | Using CS | Using CS-ANN |
|---|---|---|
| 20 | 0.0135 | 0.0025 |
| 40 | 0.0175 | 0.0036 |
| 60 | 0.0196 | 0.0045 |
| 80 | 0.0215 | 0.013 |
| 100 | 0.0261 | 0.015 |
| 120 | 0.0276 | 0.017 |
| 140 | 0.0289 | 0.018 |

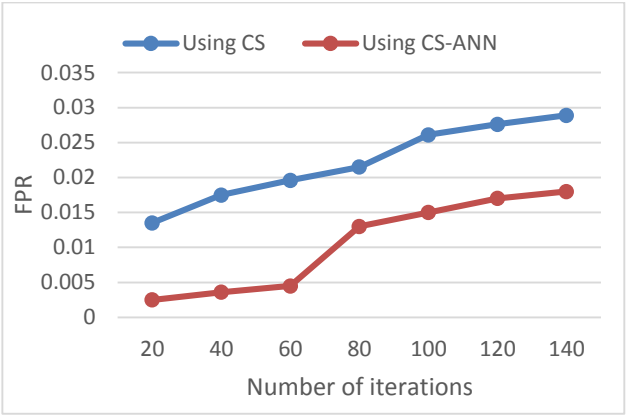


**Figure 7:** False Positive Rate (FPR)

Figure 7 shows the graphical representation of the FPR values analysed for the CS individually, and when CS-ANN approach according to the values shown in Table 2. The graph shows the results with number of iterations varied from 20 to 140 and increases in steps of 20. By performing this operation, the probability of measuring FPR with a high precision rate increases. In addition, the average rate of FPR observed for the proposed work is 0.0105, which is very small and indicates that for the detection of DDoS attack less false alarms are generated.

**Table 3:** Detection accuracy

| Number of iterations | Using CS | Using CS-ANN |
|---|---|---|
| 20 | 0.966 | 0.992 |
| 40 | 0.961 | 0.990 |
| 60 | 0.954 | 0.989 |
| 80 | 0.953 | 0.988 |
| 100 | 0.942 | 0.985 |
| 120 | 0.940 | 0.982 |
| 140 | 0.932 | 0.980 |

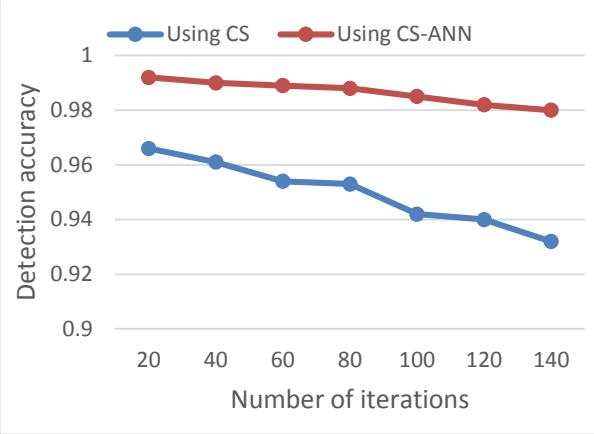


**Figure 8:** Detection accuracy

To examine the overall accuracy of the proposed approach, detection accuracy has been calculated. Figure 8 shows that the detection accuracy examined for the proposed model, using CS and CS-ANN, is 0.9497% and 0.9865% respectively. Thus, there is an improvement of 3.87% when using the CS-ANN approach in contrast to the individual CS technique. Moreover, to see the efficiency of the CS-ANN approach, a comparative analysis has been performed with the existing approaches, as shown in Table 4.

**Table 4:** Comparative analysis

| Technique used | TPR | FPR | Detection accuracy |
|---|---|---|---|
| CS-ANN | 0.99 | 0.0105 | 0.9865 |
| Amjad et al. (2019) [20] | 0.982 | 0.013 | - |
| Zareapoor et al. (2018) [21] | - | 0.07 | 0.97 |

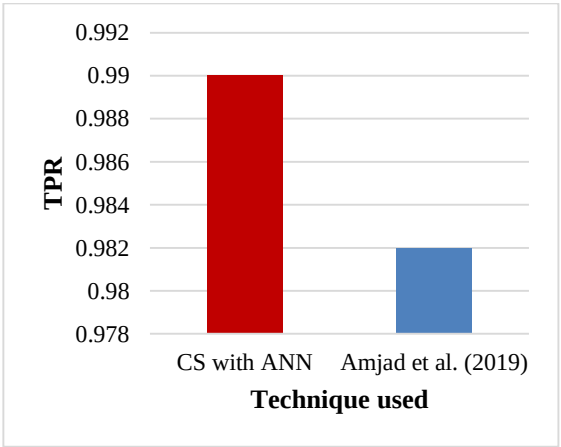


**Figure 9:** Comparison of TPR

The comparison of TPR to the proposed approach within the existing work in [20] has been performed as shown in Figure 9. [20] has used Naïve Bayes as a classification approach to detect DDoS attack on a cloud computer. Whereas, in the proposed approach, ANN is used as a classifier, which is being trained on the basis of optimized features and hence the TPR has been determined as high in contrast to the existing work. The percentage increment in the TPR of 0.82 has been observed compared to the [20] work.

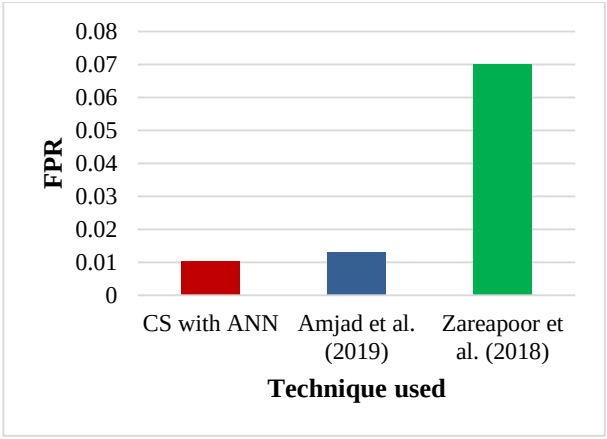


**Figure 10:** Comparison of FPR

Figure 10 represents the comparison of FPR in the proposed approach with the existing two techniques and approaches by [20] and [21] respectively. Above all, the proposed work detects a DDoS attack with minimum FPR and the percentage reductions in the FPR in contrast to [20] and [21] are 19.23% and 85% respectively.

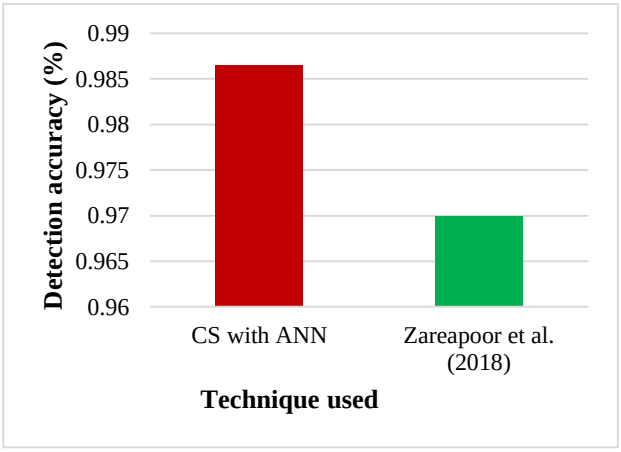


**Figure 11:** Detection accuracy

Figure 11 examined that the high detection rate has been observed as compared to [21], which is about 0.9865%. Therefore, there is an enhancement of 1.7% that has been observed, which is due to the proper selection of cloud user features and then training the ANN structure with high accuracy, which in results in increasing the detection rate.

## 5. CONCLUSION

Cloud computing is a widespread technology and works as an on-demand service basis in order to provide several features to cloud users a on pay-per-use basis. However, there are few security issues. DDoS attack is one of the most commonly noticed attacks on the cloud that affects the availability of cloud services. An effort has been made to solve this problem by detecting a DDoS attack before it enters the cloud server.

Therefore, an optimized-based machine-learning approach for the detection of DDoS attack has been presented. This paper

uses a prevention technique (CS-ANN) which detect the DDoS attack by integrating a Cuckoo Search (CS) approach with Artificial Neural Network (ANN) approach. The cloud user features, along with the attacker features, are optimized using CS as a nature-inspired approach. These are then passed to the ANN structure. The trained features are stored in the database and used during the testing process to match the test features with the trained features and hence provide results in terms of the attacker and normal cloud users. The test results of CS-ANN show a True Positive Rate (TPR), False Positive Rate (FPR) and detection accuracy of 0.99, 0.0105 and 0.9865% respectively. The designed hybrid model performed well and demonstrated better results in contrast to the two state-of-the-art techniques.

## REFERENCES

[1] Ramachandra, A. C., & Bhattacharya, S. (2020). Literature Survey on Log-Based Anomaly Detection Framework in Cloud. In *Computational Intelligence in Pattern Recognition* (pp. 143-153). Springer, Singapore.

[2] Mondal, H. S., Hasan, M. T., Hossain, M. B., Rahaman, M. E., & Hasan, R. (2017, December). Enhancing secure cloud computing environment by Detecting DDoS attack using fuzzy logic. In *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)* (pp. 1-4). IEEE.

[3] Biswas, R., & Wu, J. (2018, December). Filter assignment policy against distributed denial-of-service attack. In *2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)* (pp. 537-544). IEEE.

[4] SecureList, Kaspersky Lab Report, Distribution of DDoS attacks by type, Q4 2018, Available Online at https://securelist.com/ddos-attacks-in-q4-2018/89565/

[5] SecureList, Kaspersky Lab Report, Distribution of DDoS attacks by type, Q1 2019, Available Online at https://securelist.com/ddos-report-q1-2019/90792/

[6] Verma, P., Tapaswi, S., & Godfrey, W. W. (2019). An Adaptive Threshold-Based Attribute Selection to Classify Requests Under DDoS Attack in Cloud-Based Systems. *Arabian Journal for Science and Engineering*, 1-22.

[7] Kumar, S. B., Mukherjee, K., & Dwivedi, R. K. (2020). Secured Cloud System Using Deep Learning. In *Computational Intelligence in Pattern Recognition* (pp. 503-510). Springer, Singapore.

[8] Somani, G., Gaur, M. S., Sanghi, D., Conti, M., & Buyya, R. (2017). DDoS attacks in cloud computing: Issues, taxonomy, and future directions. *Computer Communications*, *107*, 30-48.

[9] Gupta, B. B., & Badve, O. P. (2017). Taxonomy of DoS and DDoS attacks and desirable defense mechanism in a cloud computing environment. *Neural Computing and Applications*, *28*(12), 3655-3682.

[10] Bhushan, K., & Gupta, B. B. (2019). Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment. *Journal of Ambient Intelligence and Humanized Computing*, *10*(5), 1985-1997.

[11] Devi, B. K., & Subbulakshmi, T. (2019). Cloud-based DDoS attack detection and defence system using statistical approach. *International Journal of Information and Computer Security*, *11*(4-5), 447-475.

[12] Dahiya, A., & Gupta, B. B. (2020). Multi Attribute Auction Based Incentivized Solution Against DDoS Attacks. *Computers & Security*, 101763.

[13] Hezavehi, S. M., & Rahmani, R. (2020). An anomaly-based framework for mitigating effects of DDoS attacks using a third party auditor in cloud computing environments. *Cluster Computing*, 1-19.

[14] Goyal, R., Manoov, R., Sevugan, P., & Swarnalatha, P. (2020). Securing the Data in Cloud Environment Using Parallel and Multistage Security Mechanism. In *Soft Computing for Problem Solving* (pp. 941-949). Springer, Singapore.

[15] Salim, A., Tiwari, R. K., & Tripathi, S. (2020). An Efficient Public Auditing Scheme for Cloud Storage with Secure Access Control and Resistance Against DOS Attack by Iniquitous TPA. *Wireless Personal Communications*, 1-26.

[16] Prathyusha, D. J., & Kannayaram, G. (2020). A cognitive mechanism for mitigating DDoS attacks using the artificial immune system in a cloud environment. *Evolutionary Intelligence*, 1-12.

[17] Mareli, M., & Twala, B. (2018). An adaptive Cuckoo search algorithm for optimisation. *Applied computing and informatics*, *14*(2), 107-115.

[18] Yakhchi, M., Ghafari, S. M., Yakhchi, S., Fazeli, M., & Patooghi, A. (2015, May). Proposing a load balancing method based on Cuckoo Optimization Algorithm for energy management in mobile cloud computing infrastructures. In 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO) (pp. 1-5). IEEE.

[19] Alzahrani, S., & Hong, L. (2018, July). Detection of distributed denial of service (DDoS) attacks using artificial intelligence on cloud. In *2018 IEEE World Congress on Services (SERVICES)* (pp. 35-36). IEEE.

[20] Amjad, A., Alyas, T., Farooq, U., & Tariq, M. (2019). Detection and mitigation of DDoS attack in cloud computing using machine learning algorithm. *EAI Endorsed Transactions on Scalable Information Systems*, *6*(23).

[21] Zareapoor, M., Shamsolmoali, P., & Alam, M. A. (2018). Advance DDOS detection and mitigation technique for securing cloud. *International Journal of Computational Science and Engineering*, *16*(3), 303-310.