International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

# Cloud Based DDOS Attack Detection Using Machine Learning Architectures: Understanding the Potential for Scientific Applications

**Manjunath C R [1], Ketan Rathor [2], Dr Nandini Kulkarni [3], Dr. Prashant Pandurang Patil [4], Dr. Manoj S. Patil [5], Jasdeep Singh [6]**

**Submitted:** 19/08/2022 **Accepted:** 18/11/2022

***Abstract:*** Cloud computing technology has become a crucial component of IT services utilized in daily living in this era of technology. Website hosting services are gradually migrating to cloud in this regard. This increases the value of cloud-based websites while also creating new risks for those services. A severe threat of this nature is DDoS attack. This research propose novel technique in cloud based DDOS attacks using machine learning architectures. here the input has been collected based on cloud module and it has been processed for dimensionality reduction and noise removal. Then this data feature has been extracted and classified using ResNet-101 based KELM. The experimental analysis has been carried out in terms of data delivery ratio, transmission rate, validation accuracy, training accuracy, end-end delay. the proposed technique attained data delivery ratio of 92%, transaction rate of 82%, validation accuracy of 89%, training accuracy of 96%, end-end delay of 56%

***Keywords:*** *cloud computing, DDoS attack, machine learning, dimensionality reduction, data delivery ratio*

## 1. Introduction

A form of Internet-based computing known as "cloud computing" offers a shared pool of resources, including memory, processing power, network bandwidth, and user applications. These resources can be quickly and inexpensively made available to end users over the Internet [1] upon request. Three methods—Software-as-a-service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service—are used to group cloud computing services. (IaaS). It can also be used as a hybrid cloud or as a private, public, or community cloud. One of the main problems this technology is now facing is a lack of businesses eager to fully adopt the cloud. An aggressive attack called a DDoS can seriously disrupt cloud servers. The phrase "Denial of Service" (DoS) was first used by Gligor in relation to operating systems, claims [2], but it has subsequently gained widespread usage. A distributed denial of service (DDoS) assault is a coordinated DoS attack that uses multiple computers to target a single victim. Main goal of this research is to apply ML to identify DDoS attacks. Literature has addressed the issue of attack detection using machine learning methods before. While anomaly detection can identify new and unknown assaults, signature detection approaches are successful against established attacks (zero-day). Attack-generated data flow has an erratic condition. Due to this, DDoS assaults are simple to launch, complex to detect and trace, and so on. DDoS assaults have also emerged as one of the main dangers to network security [3].

[1] *Associate Professor,*
*Department of Compute Science and Engineering, Jain(Deemed-to-be University), Bangalore, India.*
*cr.manjunath@jainuniversity.ac.in*
[2] *GyanSys Inc Carmel,*
*myemail.ketan@gmail.com*
[3] *Symbiosis School of Planning Architecture and Design, Symbiosis International Deemed University.*
*Symbiosis School of Planning Architecture and Design, Nagpur.*
*deputydirector@sspad.edu.in*
[4] *Associate Professor,*
*Bharati Vidyapeeth (Deemed to be University), Y. M. Institute of Management Karad.*
*Prashant.patil@bharatividyapeeth.edu*
[5] *Research Consultant,*
*Research and Development, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences, Wardha, Maharashtra, India.*
*mpatil98dent@gmail.com*
[6] *Assistant Professor,*
*Department of CSE, RIMT University, Mandi, Gobind garh, Punjab, India.*
*jasdeepsingh@rimt.ac.in*

## 2. Related Works

There are various machine learning methodologies accessible. The application of ML in CC is covered in certain review publications [4]. The authors of [5] use self-adaptive evolutionary extreme learning to find DDoS attacks. The technique contains two key components, namely the automatic recognition of the neurons in hidden layer as well as identification of optimum crossover operator. Experimental findings are used to test the suggested strategy and indicate enhanced accuracy. DDoS assaults are detected in Software Defined Networks using a different method provided in [6]. (SDN). For real-time DDoS attack detection, the authors used DNN. According to experimental findings, this approach detects DDoS attacks more accurately, faster, and with less resource use. In [7], authors compared various machine learning techniques for identifying DDoS assaults. According to experimental findings, RF more accurately identifies DDoS attacks. In a different study, scientists choose the most pertinent characteristics for DDoS attack

detection using [8] correlation, data gain, and relief feature selection technique. To detect attacks, they employed a deep belief network model. The CICIDS 2017 assaults dataset serves as the basis for the experiments. The suggested method shows 99.37% accuracy in recognising attacks of the usual class and 96.67% accuracy in detecting DDoS attacks [9].Work [10] created a methodology that identifies DDoS assaults in SDN in order to identify all DDoS attacks.

## 3. System Model

This section discuss novel technique in cloud based DDOS attacks using machine learning architectures. here the input has been collected based on cloud module and it has been processed for dimensionality reduction and noise removal. Then this data feature has been extracted and classified using ResNet-101 based KELM. The proposed architecture is represented in figure-1.

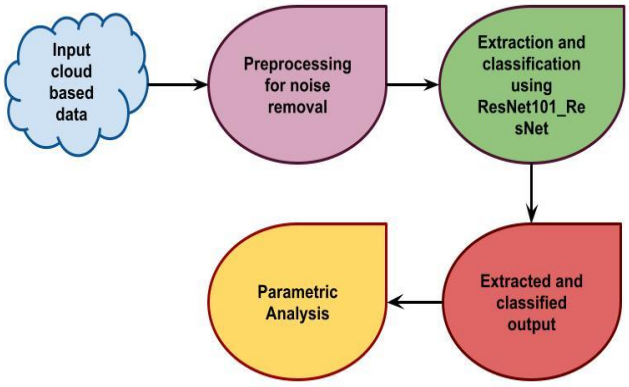


**Figure 1.** Overall proposed architecture

Dataset preparation. This step involves storing normal and flooding attack traffics in a SQL database that has a single table with following fields: the number of packets, the number of bytes, the number of packets A-B, the number of bytes B-A, the number of packets B-A, the number of bytes B-A, the duration, the bps A-B, and the bps B-A.

**ResNet-101 based KELM in feature extraction and classification:**

Comparatively speaking to other CNN architectures like VGGnet, ResNets are simpler to train. A 50,101-layer ResNet, for instance, is 8 times deeper than a VGGnet yet remains less complex and trains more quickly. Additionally, 3.6 billion multiply-add operations are found in a 34-layer ResNet, compared to 19.6 billion in a 19-layer VGGnet. Extremely deep networks are known to exhibit overfitting and accuracy saturation. As a result, ResNets performs tasks like picture segmentation, localisation, and detection with remarkable results. Remaining units make up residual networks. You can express each residual unit as eq. (1):

$$y_i = h(x_i) + F(x_i, w_i) \quad (1)$$

where xi and yi are inputs and outputs of i-th layer, wi is weight matrix, and F is a residual function and a ReLU function. Identity mapping function h is provided by eq. (2):

$$h(x_i) = x_i \quad (2)$$

Residual function F is as as eq. (3):

$$F(x_i, w_i) = w_i \cdot \sigma\left(B(w_i') \cdot \sigma\big(B(x_i)\big)\right) \quad (3)$$

where "" denotes convolution, B(xi) is the batch normalisation, and (x) = max (x, 0). The branching of gradient propagation routes is the fundamental concept of residual learning. Highway networks and residual networks both have elements like residual blocks and shortcut links in common. Instead of considering ResNets as an extremely complex architecture, consider it as an ensemble of numerous pathways. These network paths in the ResNets, however, do not all have same length. Through all of the residual units, there is only one passage. Additionally, none of these signal routes transmit gradient, which explains why ResNets can be optimised and trained more quickly.

Skip links between deep layers are the foundation of the ResNet. These connections that skip layers of non-linear transformation can skip one or more of those layers. ELM is a powerful learning technique for feed-forward neural networks with a single hidden layer (SLFNs). Low computational cost is achieved through the random generation of the hidden layer weights and biases in ELM. Model of a single hidden layer NN with L hidden nodes can be stated as, taking into account n different training samples with xi ∈ RM, yi ∈ RC, and{xi, yi}  n i=1 from C classes by eq. (4).

$$\sum_{j=1}^{L} \beta hh\big(w_j \cdot \mathbf{x}_i + e_j\big) = \mathbf{y}_i, \quad i = 1, \ldots, n \quad (4)$$

By rearranging compact form of model (3) utilizing the matrix notation Hβ = Y,where $\beta = [\boldsymbol{\beta}_1^T \cdots \boldsymbol{\beta}_L^T]^T \in \mathbb{R}^{L\times C}, \mathbf{Y} = [\mathbf{y}_1^T \cdots \mathbf{y}_n^T]^T \in \mathbb{R}^{n\times C}$ and H is usually referred to as hidden layer output matrix by eq. (5):

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \vdots \\ \mathbf{h}(\mathbf{x}_n) \end{bmatrix} = \begin{bmatrix} h(\mathbf{w}_1 \cdot \mathbf{x}_1 + e_1) & \cdots & h(\mathbf{w}_L \cdot \mathbf{x}_1 + e_L) \\ \vdots & \ddots & \vdots \\ h(\mathbf{w}_1 \cdot \mathbf{x}_n + e_1) & \cdots & h(\mathbf{w}_L \cdot \mathbf{x}_n + e_L) \end{bmatrix} \quad (5)$$

β is evaluated by a least squares solution by eq. (6).

$$\beta' = \mathbf{H}'\mathbf{Y} \quad (6)$$

where H† stands for H's pseudoinverse matrix. The ELM's output function by eq. (7)

$$\mathbf{f}_L(\mathbf{x}_i) = \mathbf{h}(\mathbf{x}_i)\beta = \mathbf{h}(\mathbf{x}_i)\mathbf{H}^T \left(\frac{1}{\rho} + \mathbf{H}\mathbf{H}^T\right)^{-1} \mathbf{Y}$$

$$\mathbf{f}_L(\mathbf{x}_i) = \left(\frac{1}{\rho} + \mathbf{H}\mathbf{H}^T\right)^{-1} \mathbf{Y} \quad (7)$$

Index of output nodes with highest value determines the label of a test sample.

## 4. Performance Analysis

We put our detection system's prototype into practise in actual cloud environments. The servers run on a 64 GB memory, Intel Xeon E5-2690 processor. The other three have an Intel Xeon CPU E3-1230 V2 processor and 32 GB of RAM. Ubuntu 14.04 is running on all cloud servers and VMs that have been launched on them. Using the OpenStack cloud infrastructure, each virtual machine that is launched on server runs on a single 3.2GHz VM.

Datasets: Datasets for CICIDS 2017 and CICDDoS 2019 were taken from corresponding websites. 3.1 million records of traffic flow make up the CICIDS 2017 dataset. This data set includes traffic flow log files for 5 days, from Monday through Friday. On Friday night's network traffic log file, we performed our experiments. The class label is one of 79 features and 225,711 instances in this log file. The DrDoS NTP file is chosen as one from the CICDDoS 2019 dataset. The file cleans 84 input features and has 1,209,961 instances.

**Table-1** Comparative analysis between proposed and existing technique

| Parameters | SDN | DNN_DoS | CDDoS_MLA |
|---|---|---|---|
| **Data Delivery ratio** | 86 | 88 | 92 |
| **Transaction rate** | 77 | 79 | 82 |
| **Validation accuracy** | 83 | 85 | 89 |
| **Training accuracy** | 92 | 95 | 96 |
| **End-End delay** | 51 | 53 | 56 |

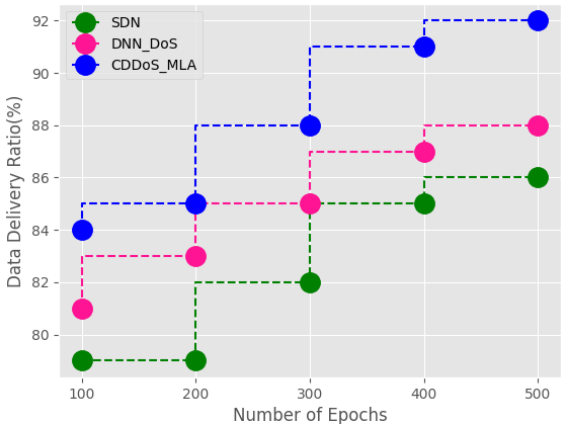


**Figure- 2** Comparison of data delivery ratio

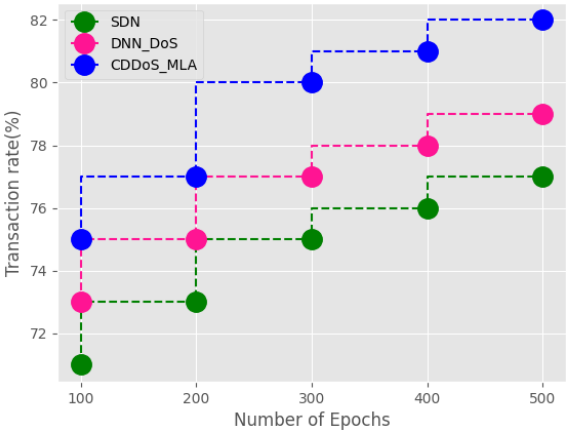


**Figure- 3** Comparison of transaction rate

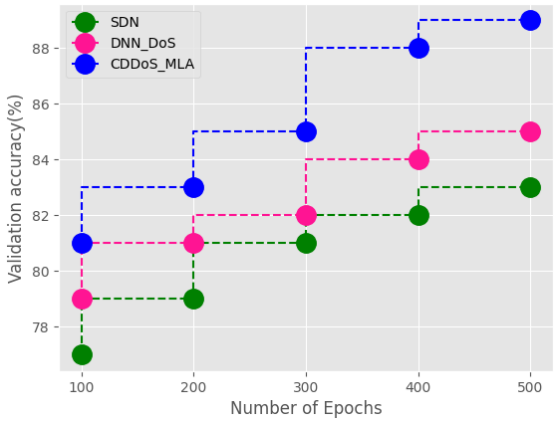


**Figure- 4** Comparison of validation accuracy

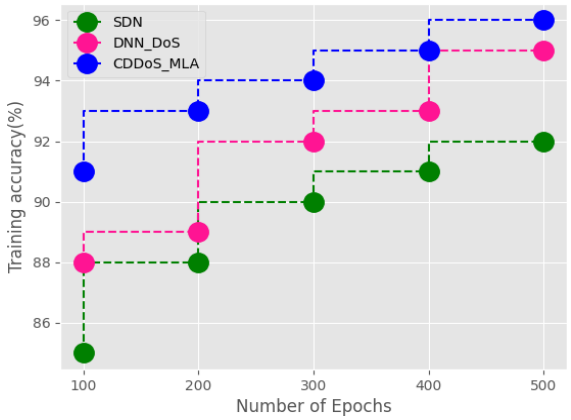


**Figure- 5** Comparison of training accuracy

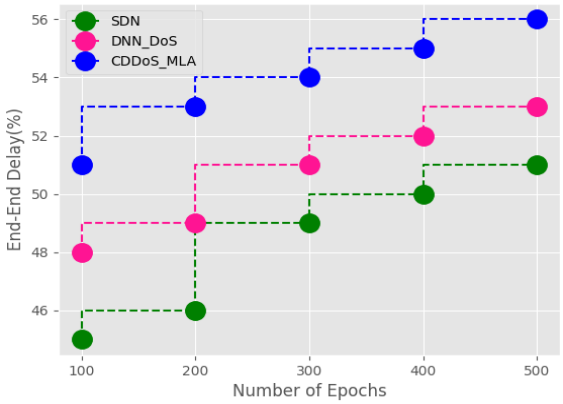


**Figure- 6** Comparison of end- end delay

The above table-1 shows comparative analysis between proposed and existing technique based on DDOS attack detection. Here the parametric analysis has been carried out in terms of data delivery ratio, transaction rate, validation accuracy and training accuracy, end-end delay. the proposed technique attained data delivery ratio of 92%, transaction rate of 82%, validation accuracy of 89%, training accuracy of 96%, end-end delay of 56% as shown in figure 2-6.

## 5. Conclusion

This research propose novel technique in cloud based DDOS attacks using machine learning architectures. The processed data has been extracted and classified using ResNet-101 based KELM. ResNets' residual units are not stacked together like the convolutional layers in a typical CNN are. Instead, quick connections are added between each convolutional layer's input and output. Complexity of residual networks is reduced when identity mappings are used as shortcut connections, making deep networks easier to train. the proposed technique attained data delivery ratio of 92%, transaction rate of 82%, validation accuracy of 89%, training accuracy of 96%, end-end delay of 56%.

## Reference

[1] Ahuja, N., Singal, G., & Mukhopadhyay, D. (2021, January). DLSDN: Deep learning for DDOS attack detection in software defined networking. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 683-688). IEEE.

[2] Kachavimath, A. V., & Narayan, D. G. (2021). A deep learning-

based framework for distributed denial-of-service attacks detection in cloud environment. In Advances in computing and network communications (pp. 605-618). Springer, Singapore.

[3] Tonkal, Ö., Polat, H., Başaran, E., Cömert, Z., & Kocaoğlu, R. (2021). Machine learning approach equipped with neighbourhood component analysis for DDoS attack detection in software-defined networking. Electronics, 10(11), 1227.

[4] Yungaicela-Naula, N. M., Vargas-Rosales, C., & Perez-Diaz, J. A. (2021). SDN-based architecture for transport and application layer DDoS attack detection by using machine and deep learning. IEEE Access, 9, 108495-108512.

[5] Gaurav, A., Gupta, B. B., & Panigrahi, P. K. (2022). A novel approach for DDoS attacks detection in COVID-19 scenario for small entrepreneurs. Technological Forecasting and Social Change, 177, 121554.

[6] Alduailij, M., Khan, Q. W., Tahir, M., Sardaraz, M., Alduailij, M., & Malik, F. (2022). Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method. Symmetry, 14(6), 1095.

[7] Sudar, K. M., Beulah, M., Deepalakshmi, P., Nagaraj, P., & Chinnasamy, P. (2021, January). Detection of Distributed Denial of Service Attacks in SDN using Machine learning techniques. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-5). IEEE.

[8] Dinh, P. T., & Park, M. (2021). BDF-SDN: A big data framework for DDoS attack detection in large-scale SDN-based cloud. In 2021 IEEE Conference on Dependable and Secure Computing (DSC) (pp. 1-8). IEEE.

[9] Awan, M. J., Farooq, U., Babar, H. M. A., Yasin, A., Nobanee, H., Hussain, M., & Zain, A. M. (2021). Real-time DDoS attack detection system using big data approach. Sustainability, 13(19), 10743.

[10] Gupta, B. B., Gaurav, A., & Peraković, D. (2021, October). A big data and deep learning based approach for ddos detection in cloud computing environment. In 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE) (pp. 287-290). IEEE.