

# A Novel Analysis of Network Traffic in Distributed Denial of Service (DDoS) Attack to Improve Accuracy Using Extreme Learning Machine Algorithm over Regression Algorithm

E. Dharanidharan<sup>a</sup> and V. Parthipan<sup>b,1</sup>

<sup>a</sup>Research Scholar, Department of CSE, Saveetha School of Engineering,

<sup>b,1</sup> Associate Professor, Department of CSE, Saveetha School of Engineering,

<sup>a,b</sup>Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu, India

**Abstract.** The research work aims to analyze novel optimized network traffic using Distributed Denial of Service attack dataset and get accuracy using Regression Algorithm (RA) and compare it with Extreme Learning Machine Algorithm (ELM). Considering two groups as Regression Algorithm and Extreme Learning Machine Algorithm. For each framework take N=10 from the dataset to perform both iterations on each framework value. sample size is calculated using Gpower software. ccuracy of network traffic DDoS attack dataset using Regression algorithm is 84.6% and Extreme Learning Machine algorithm is 90.7% with significance value of 0.03 ( $p < 0.05$ ). Based on analysis, the Extreme Learning Machine algorithm provides better accuracy over the Regression Algorithm.

**Keywords:** Extreme Learning Machine Algorithm, Regression Algorithm, Accuracy, DDoS, Novel optimized Network Traffic.

## 1. Introduction

The research is all about analysis of service in secure document storage in cloud computing using RA. Implementation of novel secure document analysis with use of RA has succeeded in developing secure document storage in cloud computing [1]. In today's world, the cloud computing worldview has arrived at a high level. Numerous individuals currently like to use this element for calculation, handling, and capacity of their information. Building and keeping up particular server farms are very expensive. So normally information stockpiling is kept with an outsider cloud specialist co-op that deals with this expense. In such a scenario, maintaining the privacy of user's data from unauthorized users is not an easy task [4]. Improvised ELM technique exhibits potential

---

<sup>1</sup>V.Parthipan, Associate Professor Associate Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences,parthipan@saveetha.com.

improvements in terms of classification accuracy and number of features. Therefore this is applied in research. ELM is a promising technique to acknowledge adaptable, fine-grained admittance control answers for detecting information in the cloud [30]. Application is to predict android malware using extreme learning machines with different kernel functions in securities based on the platform of privacy [12][13]. Extreme learning machine model is used to evaluate the college English classroom teaching based on particle swarm [28].

Nearly 102 papers were published in accordance with DDoS using RA. This paper utilizes different machine learning techniques for recognition of service and productively to ensure offered administrations from web workers are accessible.[11]

To ensure offered administrations from web workers are accessible. Outcome from this approach recommends that 97.1% of DDoS assaults are effectively recognized by ELM[17] [18] DDoS attacks make challenges to provide services of data resources to web clients [6], Denial of Service (DoS) attacks are intended primarily to disrupt computing systems in a network. Fundamentally, these attacks are initiated from one machine with the problem of targeting a server system through an attack. [2][9]. RA refers to a method of studying the relationship between independent variable and dependent variable. The application of research is for prevention and securable of data that has been detected using and especially in pattern examination, promoting, time series forecast and even medication reaction demonstrating[19][21]. It has been used in various fields and applications because of better generalization ability, robustness, and controllability, and fast learning rate [15][16]. It first analyzes theory and algorithm ideas of ELM, then tracking describes the latest progress of ELM in recent years, including model and specific applications of ELM, finally points out the research and development prospects of ELM in future [7] [8]. Our wide portfolio in research has translated into publications in numerous interdisciplinary projects.[10][23].

Based on the literature review, ELM techniques are used to separate DDoS attack instances from beginning instances and therefore, experimental results show that forward and backward data bytes of a dataset are observed more similarly for DDoS attacks compared to bytes for beginning attempts [24]. Network-level DoS attacks disable connectivity of valid users to access network resources, and application-level DoS novel based analysis of network traffic disrupt services from server resources temporarily or indefinitely[27]. DDoS attacks prevent authentic clients from accessing information from online servers [3] [26]. Optimization problems arising in learning parameters of an ELM model are often solved analytically, leading to a closed-form involving only matrix operation and inversion. ELM will address batch learning, sequential attacks, and threats. To improve and increase the feasibility detection of network traffic and its prevention methods has been supported using ELM technique and RA in cloud computing.

## **2. Material and Methods**

In this research work, implementation, execution, error detection, and rectification was done at OOAD Laboratory, Saveetha School of Engineering. The study does not require any ethical permission. Number of groups required in the research is two, Group 1 is RA compared with Group 2 of ELM algorithms. Samples were taken from the algorithm and iterated with 10 samples to get the accuracy. Sample size was

calculated by using the clinical.com website by keeping G power 80% threshold 0.05%, CI 95%. Mean and Standard deviation has been taken from a previous literature for size calculation [27]. Two different groups were taken for analysis through RA (N=10) and ELM (N=10) [20].

Denial of service is attacking a server from the other end while the user is requesting access to the server for information. Denial of service attack is rejected and makes server inaccessible to intended users. Have taken a dataset that contains the attacked IP and TCP with that accuracy and data of the attacker can be acquired by using a neural network on Protocol of IP address. Accuracy is a percentage of calculating the current forecast for test data and training data is given in Equation 1.

$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{FP} + \text{FN} + \text{TP}) (1)$$

Where

(TN+TP) is the current prediction

(TN+FP+FN+TP) is the total number of predictions.

### 2.1. Regression Algorithm

Regression algorithms fall under the group of Supervised Machine Learning calculations which is a subset of AI calculations. One of the fundamental highlights of managed learning calculations is that model conditions and connections between the objective yield and info highlights to anticipate incentive for new information. A regression algorithm anticipates yield estimates dependent on input highlights from the information taken care of in the framework. Go-to philosophy is calculation constructs a model on highlights of preparing information and utilizing the model to anticipate incentive for new information.

### 2.2. Pseudo Code for Regression Algorithm

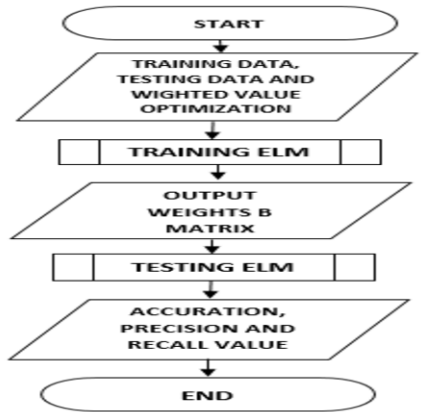
```
private double []
doRegression(boolean [] selectedAttributes) throws Exception {
    if (b_Debug) {
        for (int i = 0; i < selectedAttributes.length; i++) {
            for (int i = 0; i < selectedAttributes.length; i++)
            {
                if (selectedAttributes[i]) {
                    numAttributes++;
                    if (numAttributes > 0)
                    {
                        for (int i = 0; i < m_TransformedData.numInstances(); i++) {
                            Instance inst = m_TransformedData.instance(i);
                            double sqrt_weight = Math.sqrt(inst.weight())
                            for (int j = 0; j < m_TransformedData.numAttributes(); j++)
```

### 2.3. Extreme Learning Machine Algorithm

Extreme Learning Machine formula was projected by Huang et al. in 2006. This algorithm relies on a single-layer feedforward neural network (SLFN). ELM

additionally solves the classification and regression problems. Consequently, ELM will be wont to notice anomalies in novel optimized network traffic. It's widely employed in several kinds of research. ELM incorporates a high computation rate and doesn't need reiterative parameter calibration for coaching. In ELM model, input file training is meted out at intervals of seconds or minutes, even for big data sets and sophisticated applications, which cannot be achieved by victimization conventional methods.

Figure 1. Represents the structure of ELM framework which is targeted in infrastructure service. The main characteristics are inputting dataset into process and reaching hidden layer in the traffic layer and calculating precision and then accuracy.



**Figure 1.** Flow diagram for training and testing the data

The data collection of this research topic detection data and no detection data are observed and collected and stored as a dataset. With the help of algorithms and data to find throughput from statistics tools or software. The testing setup of both hardware and software configurations. Hardware configuration of the system/Laptop has an Intel Core i3 5<sup>th</sup> generation processor with 8 GB RAM (Random access Memory) and 64-bit Operating System, x64 - based processor type, 1TB HDD. The software configuration has a Windows OS and the language used here to compile and execute programs in Python language.

#### 2.4. Statistical Analysis

Statistical analysis was done with SPSS statistics by IBM version 21. The independent variable is Flow ID, Protocol, Source Port, Source IP, Destination port, Destination IP and dependent variable is accuracy. Independent sample T test analysis is carried out.

### 3. Results

Table 1. the accuracy analysis of RA and ELM algorithms for various iterations RA and ELM were run at different times in Jupiter with a sample size of 10 and accuracy was measured.

**Table 1.** Comparison accuracy analysis of RA and ELM algorithms for various iterations. Accuracy of RA and ELM is obtained from different network traffic using DDoS attack dataset.

| S.No | Regression Algorithm<br>(Accuracy) | Extreme Learning<br>Machine (Accuracy) |
|------|------------------------------------|--|
| 1    | 87.00                              | 91.00                                  |
| 2    | 86.00                              | 93.00                                  |
| 3    | 80.00                              | 84.00                                  |
| 4    | 88.00                              | 94.00                                  |
| 5    | 80.00                              | 86.00                                  |
| 6    | 86.00                              | 91.20                                  |
| 7    | 84.00                              | 89.00                                  |
| 8    | 81.00                              | 87.50                                  |
| 9    | 83.00                              | 88.90                                  |
| 10   | 87.00                              | 90.10                                  |

Table 2. Group Statistics Results depicts RA has a mean accuracy of 84.2%, std. deviation is 3.04777, whereas for ELM has mean accuracy as 90.7%, std.deviation is 2.39863.

**Table 2.** Group statistical analysis provides the accuracy of 90.77% and 84.2% for ELM and RA respectively. Standard Error Mean of RA is (0.96379) and ELM is (0.75851).

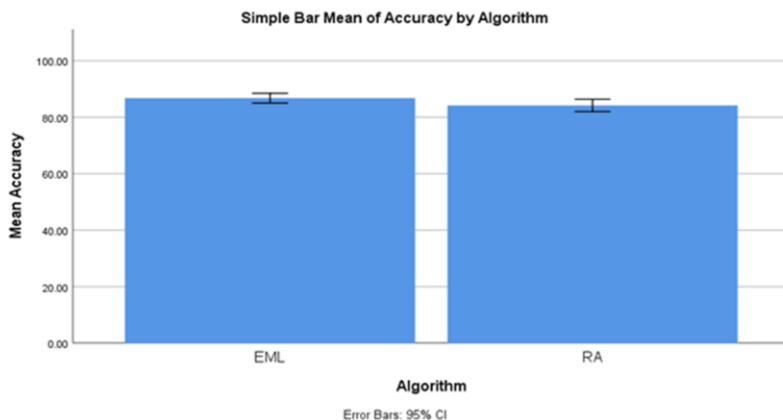
| Group Statistics |     |    |         |                |                 |
|------------------|-----|----|---------|----------------|-----------------|
| Algorithm        |     | N  | Mean    | Std. Deviation | Std. Error Mean |
| Accuracy         | RA  | 10 | 84.2000 | 3.04777        | 0.96379         |
|                  | ELM | 10 | 90.7700 | 2.39863        | 0.75851         |

Table 3.Independent samples T-test provides the mean difference of -2.57 and Std Error Difference is 1.22647. The significance value is 0.036 ( $p<0.05$ ) which reveals that both groups are statistically significant.

**Table 3.** Independent sample T-test t is performed for the two groups for significance and standard error determination. P value is 0.0367 ( $p<0.05$ ) and it is considered to be statistically significant.

|          |                                | Levene's Test for Equality of Variances |        |       |        |                     |                            |                                 |   |         |
|----------|--------------------------------|---|--------|-------|--------|---------------------|----------------------------|---------------------------------|---|---------|
|          |                                | F                                       | Sig.   | t     | df     | Mean Differ<br>ence | Sig.<br>(2-<br>taile<br>d) | Std.<br>Error<br>Differen<br>ce | 95% Confidence<br>Interval of the<br>Difference |         |
|          |                                |   |        |       |        |                     |                            |                                 | Lower   | Upper   |
| Accuracy | Equal variances<br>assumed     | .00                                     | 0.0367 | -3.84 | 18     | -2.57               | .001                       | 1.22647                         | -5.1467   | 0.00672 |
|          | Equal variances<br>not assumed |   |        | -3.84 | 17.058 | -2.57               | .001                       | 1.22647                         | -5.1979   | 0.01789 |

Figure 2. represents the error bar graph using SPSS tools, statistical comparison of mean accuracy of the network traffic using DDoS with the RA and ELM. In that comparison the RA has less accuracy (84.2%) to analyze network traffic compared to ELM (90.7%).



**Figure 2.** Comparison between mean accuracy of network traffic using DDoS attack dataset in RA and ELM algorithm. Mean accuracy RA is greater than ELM and the standard deviation of RA seems to be better than ELM. X-Axis: RA vs. ELM frameworks. Y-Axis: mean throughput of detection  $\pm 1$  SD. Mean accuracy for RA is 84.2% and ELM is 90.7%. ELM performs better than RA.

#### 4. Discussion

The research work observed that ELM proved to be better than Regression Algorithm in terms of accuracy. There is a statistical difference between two groups, 0.036 ( $p < 0.05$ ).

In another research, the article ELM is used to predict HIV drug resistance in method of target protein sequence-level [5]. They have proposed an integration of remote sensing and Mexican water quality monitoring systems using Extreme Learning Machine algorithm [1] [17]. This scheme has declared that the neural network is used to analyze malware traffic classification, a conventional method of representational learning. There is a convincing need to refresh IDS arrangements to address flow circumstance and dynamic examination is required to guarantee that these techniques don't become old [22]. Almost every research on analyzing network traffic is done with use of neural networks and machine learning, this however results in high time consumption in analyzing network traffic [14]. An answer for above issues is to change algorithms with the Extreme Learning Machine algorithm which is deep learning and is used for analysis of a small amount of data in short time consumption [25].

The limitation of the proposed system ELM is that it is not suitable for large training samples and is difficult to solve multi-case problems. In comparison to deep neural networks, extreme learning machines have less accuracy. This research work by extending some more features in Extreme learning machines can be modified and iterations can be changed with some added items. The proposed system can analyze network traffic with high accuracy and less time consumption.

## 5. Conclusion

Attacks on networks are one of causes of abnormal phenomena determined within network instrumentality operation, as well as traffic transmission over network. Anomaly detection is feasible by suggesting those machine learning ways to improve dependability of prevailing systems. During this paper, a technique for anomalies detection in network traffic supported ELM has been proposed. Compared ELM with the RA and in that RA has lesser accuracy for analyzing network traffic (84.2) and ELM has higher accuracy (90.7). Hence, ELM will analyze network traffic using DDoS dataset with high accuracy.

## References

- [1] Arias-Rodriguez, Leonardo F., Zheng Duan, José de Jesús Díaz-Torres, Mónica Basilio Hazas, Jingshui Huang, Bapitha Udhaya Kumar, Ye Tuo, and Markus Disse. 2021. "Integration of Remote Sensing and Mexican Water Quality Monitoring System Using an Extreme Learning Machine." *Sensors* 21 (12). <https://doi.org/10.3390/s21124118>.
- [2] Atre, Saloni, and Mayank Namdev. "Attribute-Based Homomorphic Encryption Based Integrity Auditing for Secure Outsourced Storage in Cloud." *SMART MOVES JOURNAL IJOSCIENCE*.2018 <https://doi.org/10.24113/ijoscience.v4i8.150>.
- [3] Bakker, Jarrod N., "Intelligent Traffic Classification for Detecting DDoS Attacks Using SDN/OpenFlow:" A Thesis Submitted to the Victoria University of Wellington in Fulfilment of the Requirements for the Degree of Master of Engineering in Network Engineering,2017.
- [4] Blaze, Matt, Gerrit Bleumer, and Martin Strauss. "Divertible Protocols and Atomic Proxy Cryptography." *Lecture Notes in Computer Science*. <https://doi.org/10.1007/bfb0054122> ,1998
- [5] Cai, Qihang, Rongao Yuan, Jian He, Menglong Li, and Yanzhi Guo. "Predicting HIV Drug Resistance Using Weighted Machine Learning Method at Target Protein Sequence-Level." *Molecular Diversity*, July. <https://doi.org/10.1007/s11030-021-10262-y>, 2021
- [6] Dittrich, David, Peter Reiher, and Sven Dietrich."Internet Denial of Service: Attack and Defense Mechanisms." Pearson Education,2004.
- [7] Ghosh, Shaona. "Online Machine Learning for Networked Data." *AI Matters*.2016. <https://doi.org/10.1145/2911172.2911179>.
- [8] Jain, Ravindra Kumar, Sridhar Prem Kumar, and W. S. Manjula."Comparison of Intrusion Effects on Maxillary Incisors among Mini Implant Anchorage, J-Hook Headgear and Utility Arch." *Journal of Clinical and Diagnostic Research: JCDR*,2014
- [9] Johnson, Jayapriya, Ganesh Lakshmanan, Biruntha M, Vidhyavathi R M, Kohila Kalimuthu, and Durairaj Sekar. "Computational Identification of MiRNA-7110 from Pulmonary Arterial Hypertension (PAH) ESTs: A New microRNA That Links Diabetes and PAH." *Hypertension Research: Official Journal of the Japanese Society of Hypertension*,2020.
- [10] Kannan, Roghith, and M. S. Thenmozhi. "Morphometric Study of Styloid Process and Its Clinical Importance on Eagle's Syndrome." *Journal of Advanced Pharmaceutical Technology & Research*,2016.
- [11] Keerthana, B., and M. S. Thenmozhi."Occurrence of Foramen of Huschke and Its Clinical Significance." *Journal of Advanced Pharmaceutical Technology & Research*, 2016.
- [12] Krishna, R. Nivesh, and K. Yuvaraj Babu "Estimation of Stature from Physiognomic Facial Length and Morphological Facial Length." *Journal of Advanced Pharmaceutical Technology & Research*. 2016.
- [13] Kumar, Lov, Chinmay Hota, Arvind Mahindru, and Lalita Bhanu Murthy Neti. "Android Malware Prediction Using Extreme Learning Machine with Different Kernel Functions." *Proceedings of the Asian Internet Engineering Conference on - AINTEC '19*. <https://doi.org/10.1145/3340422.3343639>,2019.
- [14] Liu, Yun-Hui, Ai-Jun Li, and Si-Wei Luo. n.d. "Data Modeling in Machine Learning Based on Information-Theoretic Measures." *Proceedings. International Conference on Machine Learning and Cybernetics*. <https://doi.org/10.1109/icmlc.2002.1167394>.
- [15] Malhotra, Yogesh. n.d. "AI, Machine Learning & Deep Learning Risk Management & Controls: Beyond Deep Learning and Generative Adversarial Networks: Model Risk Management in AI, Machine Learning & Deep Learning." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3193693>.

- [16] Nandhini, J. S. Thaslima, K. Yuvaraj Babu, and Karthik Ganesh Mohanraj. "Size, Shape, Prominence and Localization of Gerdy's Tubercle in Dry Human Tibial Bones." *Journal of Advanced Pharmaceutical Technology & Research*, 2018.
- [17] Polaki, Sonali Patro, Kshira Sagar Sahoo, and Bibhudatta Sahoo. *DDoS Attack Detection Using Machine Learning Techniques in SDN*, 2018.
- [18] Rubika, J., A. Sumathi Felicita, and V. Sivambiga. "Gonial Angle as an Indicator for the Prediction of Growth Pattern." *World Journal of Dentistry*, 2015.
- [19] Sekar, Durairaj, Ganesh Lakshmanan, Panagal Mani, and M. Biruntha. "Methylation-Dependent Circulating microRNA 510 in Preeclampsia Patients." *Hypertension Research: Official Journal of the Japanese Society of Hypertension*, 2019.
- [20] Sharafaldin, Iman, Arash Habibi Lashkari, Saqib Hakak, and Ali A. Ghorbani. "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy." *International Carnahan Conference on Security Technology (ICCST)*, 2019. <https://doi.org/10.1109/ccst.2019.8888419>.
- [21] Soldan, David L. *A Study of the Sequential Regression (SER) Algorithm and Its Applications*, 1980.
- [22] Son, Myungjun, and Kyungyong Lee. "Distributed Matrix Multiplication Performance Estimator for Machine Learning Jobs in Cloud Computing." *IEEE 11th International Conference on Cloud Computing (CLOUD)*, 2018. <https://doi.org/10.1109/cloud.2018.00088>.
- [23] Sriram, Nirisha, Thenmozhi, and Samrithi Yuvaraj "Effects of Mobile Phone Radiation on Brain: A Questionnaire Based Study." *Journal of Advanced Pharmaceutical Technology & Research*, 2015.
- [24] Subashri, A., and M. S. Thenmozhi. 2016. "Occipital Emissary Foramina in Human Adult Skull and Their Clinical Implications." *Journal of Advanced Pharmaceutical Technology & Research* 9 (6): 716.
- [25] Sun, Fuchen, Kar-Ann Toh, Manuel Grana Romay, and Kezhi Mao. "Extreme Learning Machines 2013: Algorithms and Applications", 2014
- [26] Thejeswar, E. P., and M. S. Thenmozhi. "Educational Research-iPad System vs Textbook System." *Journal of Advanced Pharmaceutical Technology & Research*, 2015.
- [27] Venu, Harish, Lingesan Subramani, and V. Dhana Raju. "Emission Reduction in a DI Diesel Engine Using Exhaust Gas Recirculation (EGR) of Palm Biodiesel Blended with TiO<sub>2</sub> Nano Additives." *Renewable Energy*, 2019.
- [28] Wang, Baojian, Jing Wang, and Guoqiang Hu.. "College English Classroom Teaching Evaluation Based on Particle Swarm Optimization – Extreme Learning Machine Model." *International Journal of Emerging Technologies in Learning (iJET)*. 2017 <https://doi.org/10.3991/ijet.v12i05.6782>.
- [29] Website." n.d..A Machine Learning Approach for DDoS (Distributed ... –MDPI) Accessed July 11, 2021.
- [30] Wei, Yan, Haocai Huang, Bin Chen, Bofu Zheng, and Yihong Wang. "Application of Extreme Learning Machine for Predicting Chlorophyll-a Concentration Inartificial Upwelling Processes." *Mathematical Problems in Engineering* 2019 (May).