VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

DDoS Attack Detection Using ML Models

Submitted for the course:

INFORMATION SECURITY ANALYSIS AND AUDIT – CSE 3501

Submitted By:

Swetank Kaushik 19BIT0164

Yash Kumar Patel 19BIT0183

Devansh Chauhan 19BIT0153

Sahil Saxena 19BIT0216

ABSTRACT

Attacks on the web servers and web application is the most common form of attacks that are carried out nowadays. The main reason being is most of the web application or services are vulnerable to attacks and can be easily compromised. One of the popular attacks used is DoS.

DoS stand for denial of services. Most recent websites and web servers are unable to withstand strong attacks like a DOS attack. These private servers don't have protection against simple attacks and are easily compromised. But we can use this type of attack in penetration testing to test the server stress and help to improve the security on the basis of the level of the withstanding of the website under the such type of attacks.

The aim of this paper is to test different web application against the DoS attacks and to also to determine the level to which the servers can protect themselves against malicious attacks. "XerXes" is a simple application layer denial of service (DoS) tool which is used to attack servers directly and can be launched from a single system. It does not rely on a botnet and all connections originate from a single source. Upon execution the tool launches a TCP connection flood to its target causing session table resource exhaustion, effectively crashing the server.

KEYWORDS

Web attacks, DoS, server stress, web security, denial of service, Vulnerability, penetration testing, Xerxes.

INTRODUCTION

DoS attacks happen every day, and most of the time, it is hard to flag the ill-intended traffic from normal traffic. However, you can be better prepared to counter and have measures in place so that you can defend your website against DoS attacks.

A DoS attack can target any component of your network and IT infrastructure. Attackers look for the opportunity to exploit any vulnerabilities in different layers of your network.

The following are some common DoS attacks that we see very often:

➢ Application Layer Attacks

These attacks target your network's application layer by sending HTTP traffic load with malicious intent. When an HTTP request comes to the server, to send a response, the server performs multiple tasks such as load files, querying the database, computing the request, preparing the response, etc. With such a huge amount of traffic, the server gets overloaded, and exhausts infrastructure resources and ultimately goes down. Since it is hard to classify these requests as malicious requests due to their nature being similar to actual users, the application layer DDoS attacks 3 are hard to prevent.

➢ Protocol Attacks

These attacks bring down the service by exhausting intermediate resources like state table capacity, load balancers, firewalls, TCP handshakes, etc. For example, attackers can send a TCP handshake request for connection initialization, the server sends back the response and waits for confirmation from the client. But the client never sends the confirmation, and the server keeps waiting for it, causing the server resources to exhaust. These attacks are also called state-exhaustion attacks.

Related Work

Regression Algorithm for Efficient Detection and Prediction of DDoS Attacks

Gudipudi Dayanandam et al. [2018]

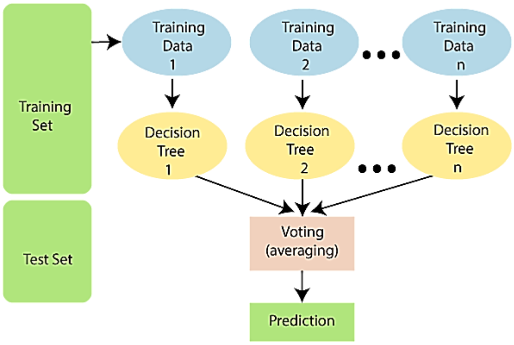
(BASE)

To improve the existing systems used in efficient detection of DDOS attacks.
To use machine learning based technique:

- GBM
- GLM
- Random Forest
- Neural Network

It is quite impossible for humans to manually identify whether a DDoS attack is happening or not so machine algorithm is required for such detections

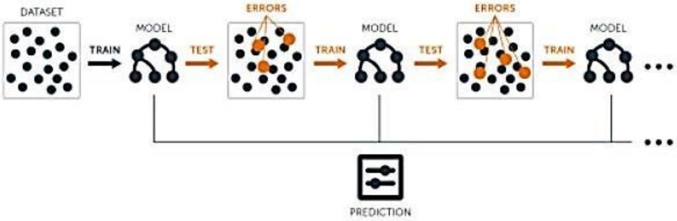
- **A random forest** is a supervised machine learning algorithm that is constructed from decision tree algorithms.



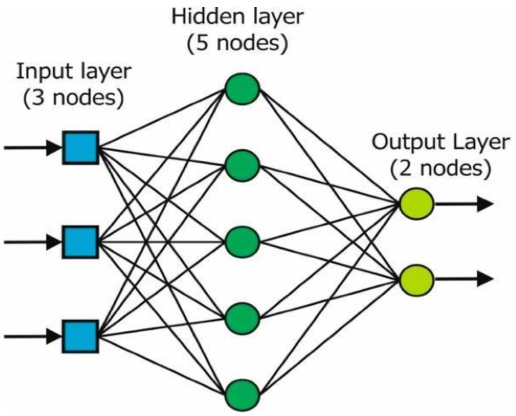
- **Generalized Linear Model** is an advanced statistical modelling technique that encompasses many other models

$$f_Y(\mathbf{y} \mid \boldsymbol{\theta}, \tau) = h(\mathbf{y}, \tau) \exp\left(\frac{\mathbf{b}(\boldsymbol{\theta})^{\mathrm{T}} \mathbf{T}(\mathbf{y}) - A(\boldsymbol{\theta})}{d(\tau)} \right).$$

- **A Gradient Boosting Machine or GBM** combines the predictions from multiple decision trees to generate the final predictions.

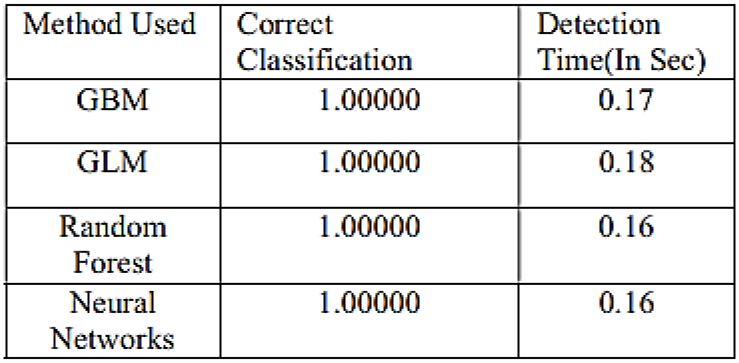


- **A neural network** is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates.



We have observed that the results have higher accuracy because of the techniques used in the research paper.

There is a comparison given between the used models in the research paper and



Method Used	Correct Classification	Detection Time(In Sec)
GBM	1.00000	0.17
GLM	1.00000	0.18
Random Forest	1.00000	0.16
Neural Networks	1.00000	0.16

No comparisons done with any other research paper.

The system can give better result when hybridized with different models used.

DDoS Attack Detection Based on Simple ANN with SMOTE for IoT Environment

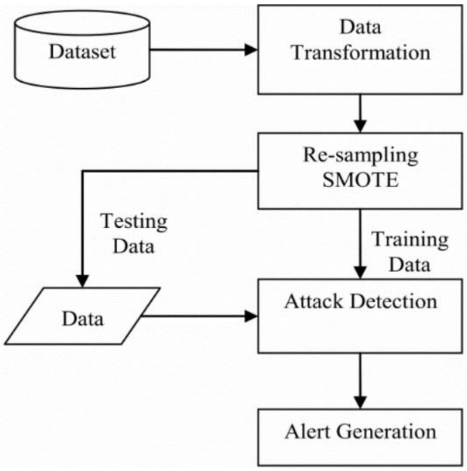
Yan Naung Soe, et al.

[2020]

It was needed to make a productive DDoS detection system, using Artificial Neural Network with SMOTE.

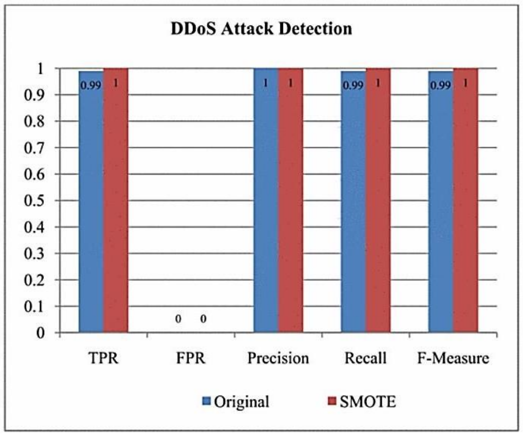
Artificial Neural Network helps in detecting the DDoS attacks and SMOTE helps in re-sampling before applying ANN

To begin, we convert nominal values to numerical values for several features that are provided by nominal data, as the detection model will be based on a neural network, which can only interpret numerical values. Some of the nominal properties are converted to decimal values ranging from 0 to n. We also change the nominal values of two other aspects, such as flags and status.



- Artificial Neural Network, is useful for classification and clustering.
- SMOTE, can be used to balance data between normal and abnormal behaviour.

After applying SMOTE the accuracy increased in comparison with the research where Smote was not applied.



The limitations of the paper can be, that it can not detect other attacks at the very moment.

Machine Learning based DDOS Detection

S.Shanmuga Priya et al.[2020]

The need to create an automated DDoS detecting system, using K Nearest Neighbour, Random Forest and Naïve Bayesian.
These three algorithms require less attribute and low amount of training data.

K Nearest Neighbour:

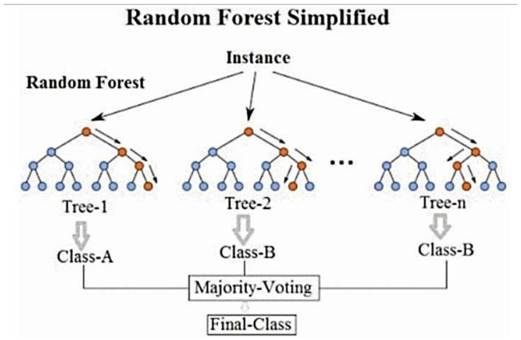
Distance calculation

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Naive Bayes:

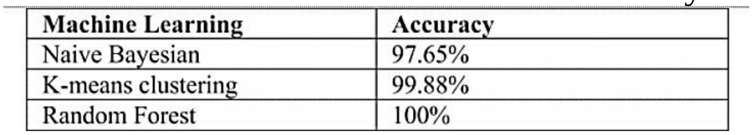
$$P(y|X)=P(X|y)*P(y)/P(X)$$

Random Forest:



The three algorithms require less attribute and low amount of training data. It can run on commodity hardware because it requires less resource power. As a result, our approach can detect DDoS attacks of any form in less time and with greater precision.

The given model is faster and accurate compared to other works that have either used KNN or Random Forest with an accuracy of 98.5%.



Machine Learning	Accuracy
Naive Bayesian	97.65%
K-means clustering	99.88%
Random Forest	100%

Due to tools such as hping3, it may not be possible to detect DDoS created by other DDoS tools.

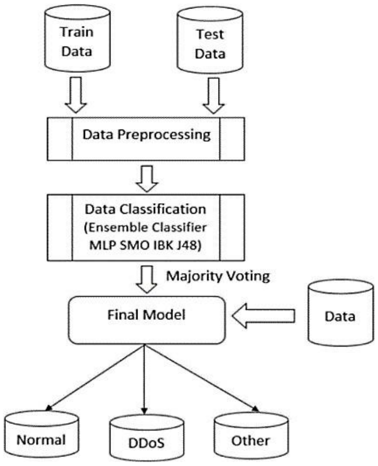
DDoS Intrusion Detection Through Machine Learning Ensemble

Saikat Das et al. [2019]

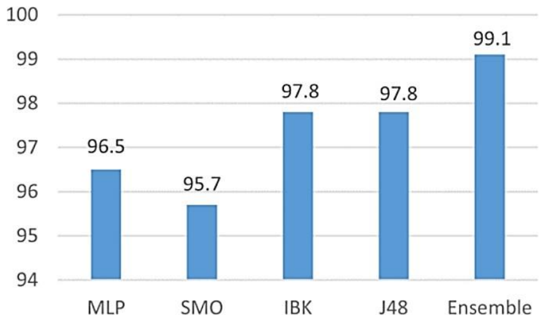
The need to detect existing as well as new types of DDoS attacks, using NIDS that combines different classifiers using ensemble models.

All of these classifiers can target specific intrusions, and provides a more robust defense mechanism.

The classifiers are, MLP (NN), SMO (SVM), IBK (KNN) and J48 (DT-C4.5) that works in parallel.
The four outputs are then combined using majority voting method and then gives the final output.



The paper shows that ensemble model has higher accuracy than the classifier working alone as shown in other papers.



Need for the expansion of method for multiple type of threats.

A DDoS Attack Detection Method Based on Hybrid Heterogeneous Multiclassifier Ensemble Learning

Bin Jia, et al. [2017]

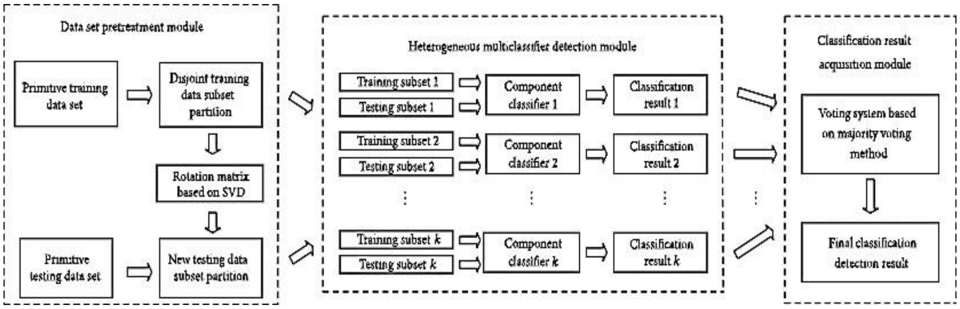
Some traditional methods and techniques have not been able to meet the needs of efficient and exact detection, so the paper propose a DDoS attack detection method based on hybrid heterogeneous multiclassifier ensemble learning and provides a more robust defense mechanism.

The classification learning model are based on Rotation Forest and SVD

The interrelation between error rate of integrated system and correlation of individual classifiers:

$$E = \left(\frac{1 + \rho (N - 1)}{N} \right) \overline{E} + E_{\text{Optimal Bayes}},$$

The majority voting is given by:

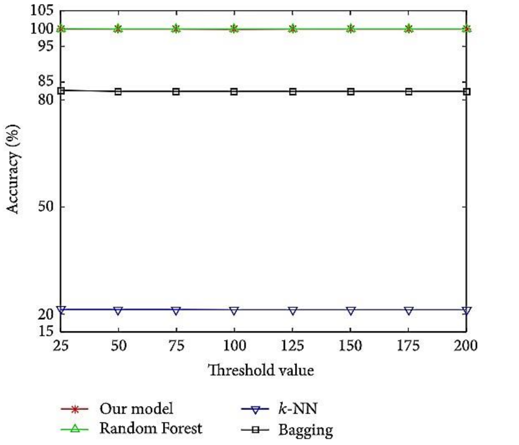

Data set pretreatment module
Primitive training data set
Disjoint training data subset partition
Rotation matrix based on SVD
Primitive testing data set
New testing data subset partition
Heterogeneous multiclassifier detection module
Training subset 1
Testing subset 1
Component classifier 1
Classification result 1
Training subset 2
Testing subset 2
Component classifier 2
Classification result 2
Training subset k
Testing subset k
Component classifier k
Classification result k
Classification result acqusition module
Voting system based on majority voting method
Final classification detection result

$$H(x) = P_i, \quad \text{if} \sum_{i=1}^{t} h_i^j(x) > \frac{1}{2} \sum_{l=1}^{m} \sum_{i=1}^{t} h_i^l(x),$$

The SVD is shown by:

$$C = U \sum V^T,$$

The experimental findings show that the model's accuracy is comparable to that of Random Forest and Bagging, and it is superior to that of KNN.



(a) Accuracy comparison based on SVD

No limitations mentioned in the given paper.

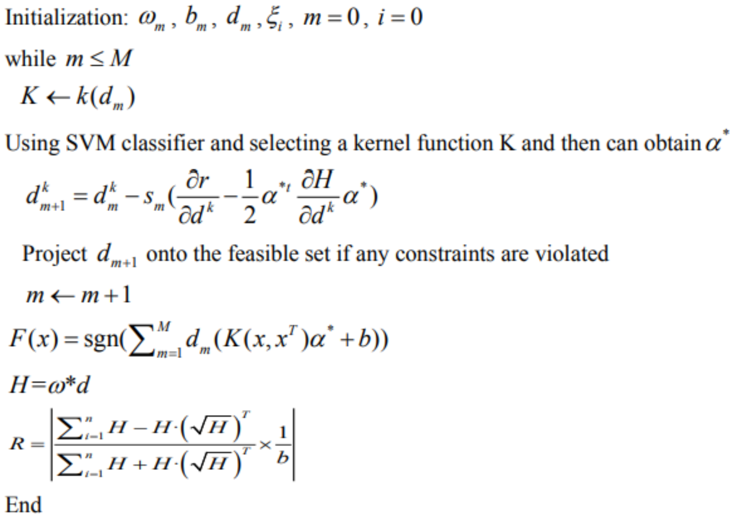
A Novel DDoS Attack Detection Method Using Optimized Generalized Multiple Kernel Learning

Jieren Cheng1 et al. [2020]

The DDoS attack intrusion detection technologies remain gloomy. The network needs to be improved with better methods. To improve the accuracy the author has used the R-GMKL algorithm to increase the accuracy .

R-GMKL algorithm

Input: train-set (x), train-label(y)
Output: ω , b , d , R, F x( )
Processing:



```
Initialization: $\omega_m$ , $b_m$ , $d_m$ , $\xi_i$ , $m = 0$ , $i = 0$
while $m \le M$
  $K \leftarrow k(d_m)$
Using SVM classifier and selecting a kernel function K and then can obtain $\alpha^*$
  $d_{m+1}^k = d_m^k - s_m(\frac{\partial r}{\partial d^k} - \frac{1}{2}\alpha^{*_t}\frac{\partial H}{\partial d^k}\alpha^*)$
  Project $d_{m+1}$ onto the feasible set if any constraints are violated
  $m \leftarrow m+1$
$F(x) = \mathrm{sgn}(\sum_{m=1}^{M} d_m(K(x, x^T)\alpha^* + b))$
$H=\omega*d$
$R = \left| \frac{\sum_{i=1}^{n} H - H\cdot(\sqrt{H})^T}{\sum_{i=1}^{n} H + H\cdot(\sqrt{H})^T} \times \frac{1}{b} \right|$
End
```

The method for kernel function and regularization parameter selection is proposed, which reduces the human-induced error to a certain extent and provides a scientific basis. The experimental results show that the proposed method can detect DDoS attacks effectively and early, and have higher detection rate and lower false negative rate than similar methods.

	Simple MKL	SVM	R-GMKL
DR (%)	78.9%	78.9%	86.2%
ER (%)	17.7%	17.7%	11.6%

It has a higher detection rate and lower error rate. Our algorithm can detect attack flow signatures at an early stage.
Although it does not have the same degree of discrimination as the CDF at the beginning

A New Framework for DDoS Attack Detection and Defense in SDN Environment

Liang Tan et al. [2017]

In order to deal with the single point of failure on SDN controllers caused by DDoS attacks, they propose a framework for detection and defense of DDoS attacks in the SDN environment. They deploy a trigger mechanism of DDoS attack detection on a data plane to screen for abnormal flows in the network using K-Means and KNN to exploit the rate characteristics.

K- Means Pseudo Code

As an advantage of SDN, centralized control also makes the controller in SDN more vulnerable to security threats from DDoS attacks

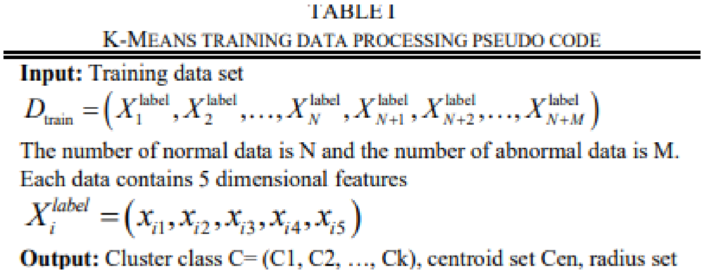
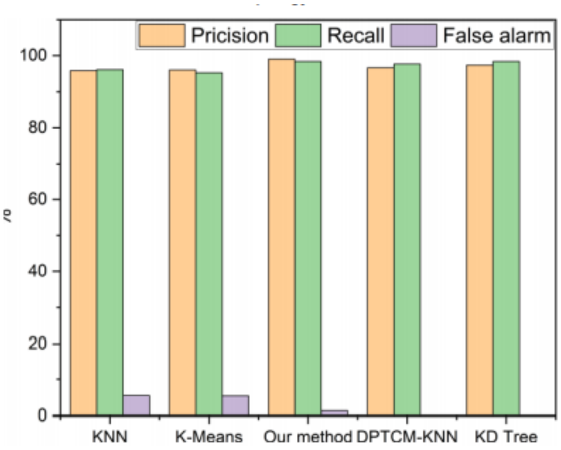
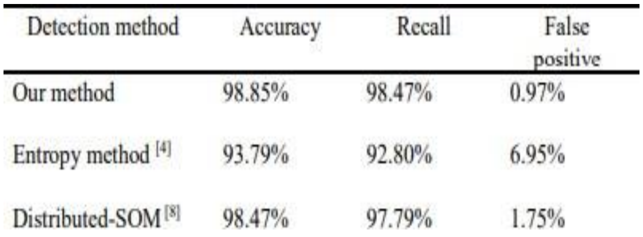


TABLE I

K-MEANS TRAINING DATA PROCESSING PSEUDO CODE

```
Input: Training data set
$D_{train} = (X_1^{label}, X_2^{label}, \dots, X_N^{label}, X_{N+1}^{label}, X_{N+2}^{label}, \dots, X_{N+M}^{label})$
The number of normal data is N and the number of abnormal data is M.
Each data contains 5 dimensional features
$X_i^{label} = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5})$
Output: Cluster class C= (C1, C2, …, Ck), centroid set Cen, radius set
```





Detection method	Accuracy	Recall	False positive
Our method	98.85%	98.47%	0.97%
Entropy method [4]	93.79%	92.80%	6.95%
Distributed-SOM [8]	98.47%	97.79%	1.75%

Comparison between different methods

The average precision of the KNN algorithm and the K-Means algorithm are 95.83% and 95.99%, resp while the average precision of this method is 99.03%, which is higher than the KNN and the KMeans algorithm.

To exploit the technology of streaming computing to reduce the burden of a single controller to ensure the efficiency of DDoS detection and network quality under large-scale network traffic.

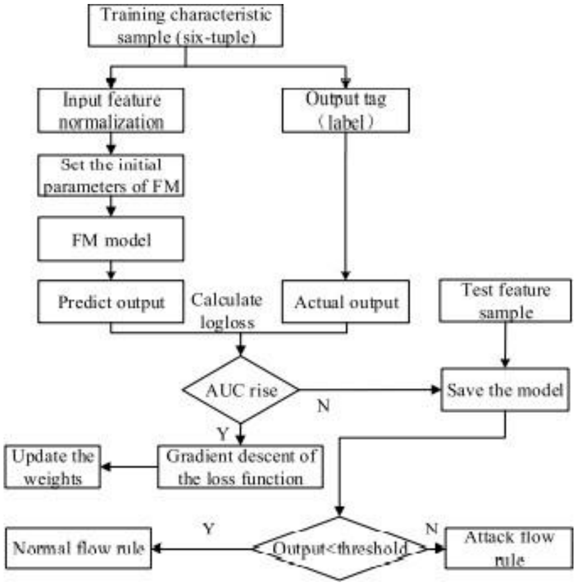
Low-Rate DDoS Attack Detection Based on Factorization Machine in Software Defined Network

WU ZHIJUN et al. [2020]

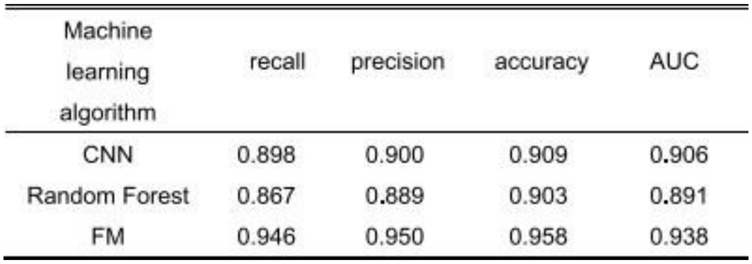
A low-rate DDoS attack against the SDN data layer is highly concealed, and the detection accuracy against this kind of attack is low.
In order to improve the detection accuracy of the low-rate DDoS attack against the SDN data layer, this paper studies the mechanism of such attacks, and then proposes a multi-feature DDoS attack detection method based on Factorization Machine (FM)

FM Algorithm Architecture

It is verified that the DDoS attack detection method based on FM algorithm has higher recall rate, precision rate and AUC value



Performance comparison of different machine learning algorithm



Machine learning algorithm	recall	precision	accuracy	AUC
CNN	0.898	0.900	0.909	0.906
Random Forest	0.867	0.889	0.903	0.891
FM	0.946	0.950	0.958	0.938

They compare and analyze the detection method in this paper with the detection method based on the joint features [17] and SDCC [9]

Need to use more evaluation criteria to compare the proposed scheme with more deep learning methods

A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks

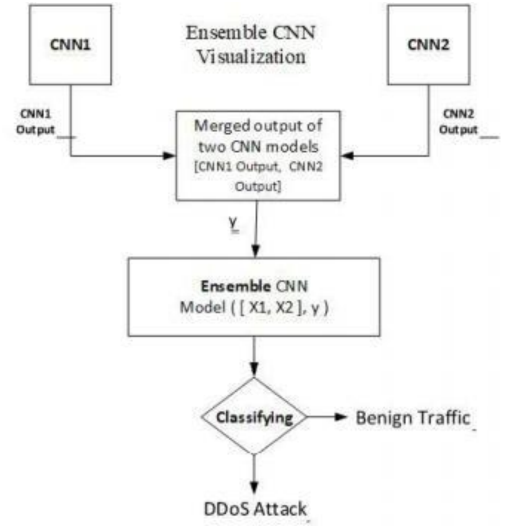
SHAHZEB HAIDER et al. [2020]

DDOS attacks are, perhaps, the most prevalent and exponentially-growing attack, targeting the varied and emerging computational network infrastructures across the globe. This necessitates the design of an efficient and early detection of large-scale sophisticated DDoS attacks

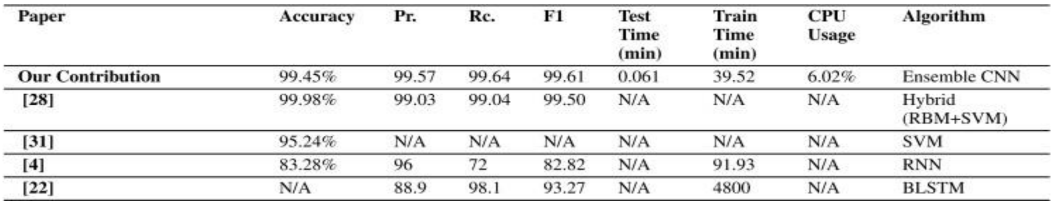
In this work, a deep convolutional neural network (CNN) ensemble framework for efficient DDoS attack detection in SDNs is proposed.

Ensemble CNN Architecture

Ensemble CNN outperforms the other three proposed DL-approaches in almost all the evaluation metrics results.


CNN1
Ensemble CNN Visualization
CNN2
CNN1 Output
Merged output of two CNN models [CNN1 Output, CNN2 Output]
CNN2 Output
y
Ensemble CNN Model ([X1, X2], y)
Classifying
Benign Traffic
DDoS Attack

Comparison of different proposed algorithms.



Paper	Accuracy	Pr.	Rc.	F1	Test Time (min)	Train Time (min)	CPU Usage	Algorithm
Our Contribution	99.45%	99.57	99.64	99.61	0.061	39.52	6.02%	Ensemble CNN
[28]	99.98%	99.03	99.04	99.50	N/A	N/A	N/A	Hybrid (RBM+SVM)
[31]	95.24%	N/A	N/A	N/A	N/A	N/A	N/A	SVM
[4]	83.28%	96	72	82.82	N/A	91.93	N/A	RNN
[22]	N/A	88.9	98.1	93.27	N/A	4800	N/A	BLSTM

It demonstrates the performance of our proposed ensemble CNN approach in efficient DDoS detection in comparison with existing competing approaches presented in.

Training and testing time of Ensemble CNN is almost 3 times longer than the other proposed approaches.

IoT DoS and DDoS Attack Detection using ResNet

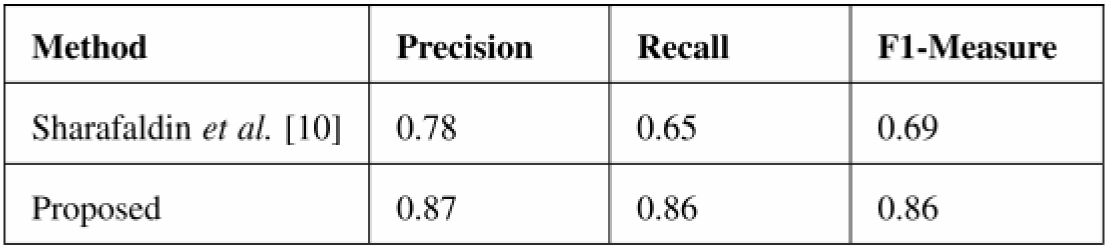
Faisal Abbas et al. [2020]

Traditional security solutions, such as firewalls and intrusion detection systems, are incapable of detecting sophisticated DoS and DDoS attacks. When these solutions are combined with artificial intelligence (AI) techniques, they can become more dependable and successful**.**

Data collecting, data cleaning, data conversion, and attack pattern recognition are the four major processes in the suggested methodology.
The suggested methodology's first step is to collect network traffic statistics.
The next stage is to train and test the CNN model on the preprocessed data in order to see how well it detects DoS and DDoS attack patterns.

Advantages:

Because CNN models are designed to discover patterns in images, they do not perform well when trained on non-image datasets. In order to take advantage of CNN models' potential, we suggested a method for converting a non-image network traffic dataset into a three-channel picture format in this paper.



Method	**Precision**	**Recall**	**F1-Measure**
Sharafaldin *et al.* [10]	0.78	0.65	0.69
Proposed	0.87	0.86	0.86

COMPARISION

In the case of binary classification, the suggested methodology achieved 99.99 percent accuracy in detecting DDoS. In comparison to the state-of-the-art, the proposed methodology achieved an average precision of 87 percent for distinguishing eleven types of DDoS attack patterns, which is 9% higher.

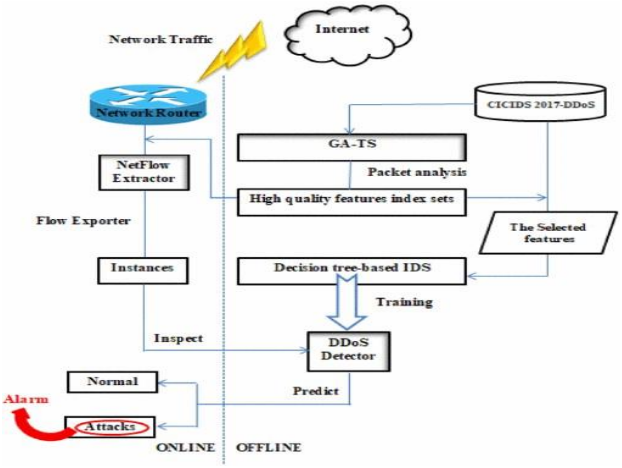
Existing technologies are incapable of detecting sophisticated DoS and DDoS attacks.

A DDoS Attack Detection System: Applying A Hybrid Genetic Algorithm to Optimal Feature Subset Selection

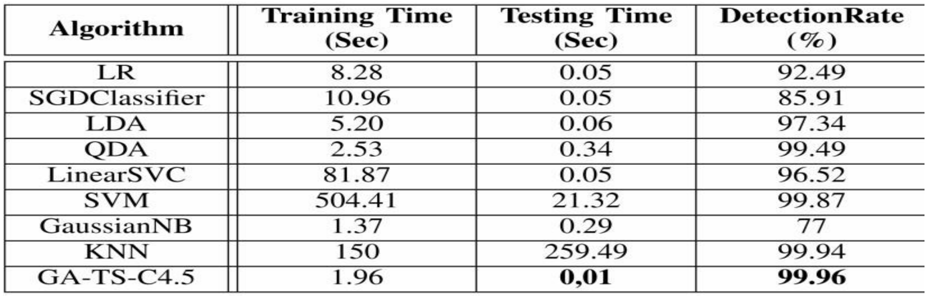
Abid Abbas et al. [2020]

Machine learning based detectors are uncompetitive and they produce false positive. To overcome this problem, we propose a Hybrid Genetic Algorithm.

A hybrid technique is a research approach that combines at least two different research approaches. A hierarchical classification and a flat classification are used in the taxonomy of hybrid metaheuristics.



There are 85 features in network traffic in data streams. The method suggested in this research has the potential to reduce the number of these features to more than 78% of the total number of features and also reduces the cost of computation and training time.



Algorithm	Training Time (Sec)	Testing Time (Sec)	DetectionRate (%)
LR	8.28	0.05	92.49
SGDClassifier	10.96	0.05	85.91
LDA	5.20	0.06	97.34
QDA	2.53	0.34	99.49
LinearSVC	81.87	0.05	96.52
SVM	504.41	21.32	99.87
GaussianNB	1.37	0.29	77
KNN	150	259.49	99.94
GA-TS-C4.5	1.96	**0,01**	**99.96**

Comparison:

In the above table we clearly see that from all the algorithm Genetic algorithm have highest detection rate.

An in-depth analysis of detection and classification is required for different types of DDoS attacks

DoS Attack Detection System using Apache Spark

Heena Mulla et al. [2021]

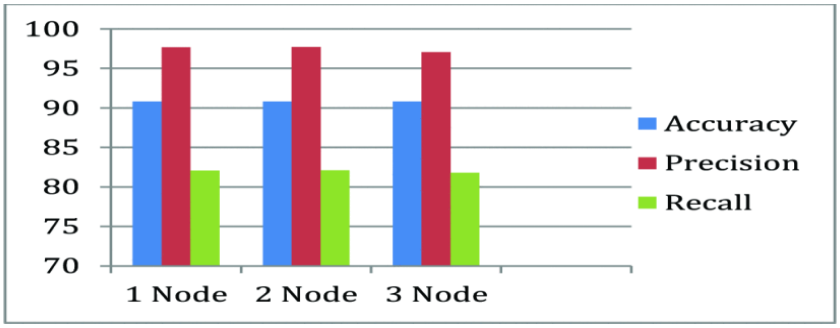
Scalability and performance of these systems are major research issues due to the large volume of network traffic. We use the Apache Spark framework to detect DDoS attacks in this paper.

The steps in the proposed system are outlined in the following sections.

1. The system is trained to detect DDoS attacks using NSL-KDD, which is then stored in HDFS.
2. The classification algorithms are trained to create the models.
3. The future data is tested using the Classified models.
4. The model is subjected to various categorization techniques.
5. The training delay is collected, as well as the accuracy, precision, and recall.

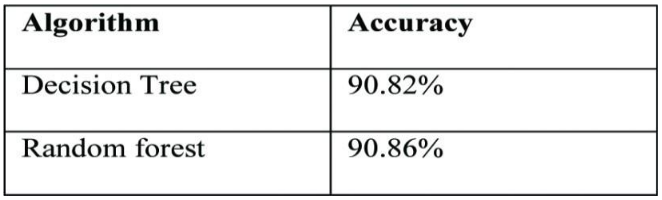
The results show that using Spark technology, the suggested system decreases the time it takes to identify DDoS attacks and considerably improves detection efficiency.

The Figure depicts the Decision tree algorithm's Accuracy, Precision, and Recall for various spark cluster sizes.



Comparison:

Random forests outperform decision trees in terms of pre-processing and training time, and distributed processing enhances performance.



Algorithm	Accuracy
Decision Tree	90.82%
Random forest	90.86%

The result will not accurate if number of nodes is decreasing.

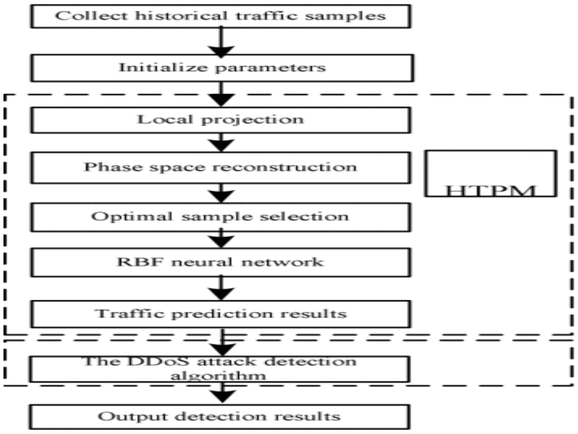
DDoS Attack Detection Algorithm Based on Hybrid Traffic Prediction Model

Yuze MENG et al. [2018]

A DDoS attack detection method based on hybrid traffic prediction model (DADA- HTPM) is suggested in order to detect Distributed Denial of Service (DDoS) attacks correctly and quickly.

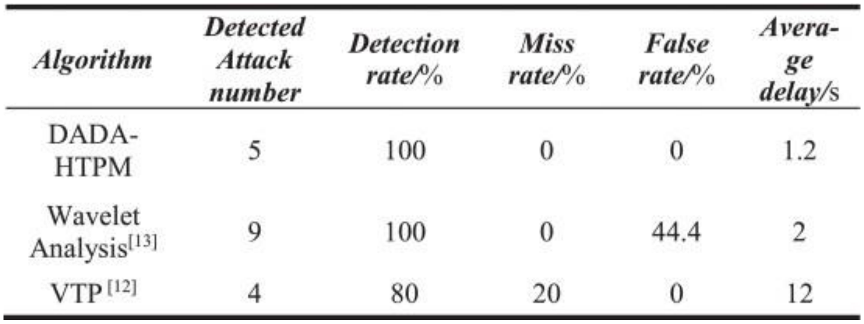
Details:

The hybrid traffic prediction model and the DDoS assault detection algorithm are the two primary components of the DADA-HTPM. Figure shows the DADA-HTPM block diagram.


Collect historical traffic samples
Initialize parameters
Local projection
Phase space reconstruction
HTPM
Optimal sample selection
RBF neural network
Traffic prediction results
The DDoS attack detection algorithm
Output detection results

The DDoS attack detection algorithm given in this research, which is based on a hybrid traffic prediction model, not only has higher traffic prediction accuracy and lesser complexity, but it can also identify DDoS attacks fast and accurately.

Three techniques for detecting DDoS attacks are compared. In our example, there are five DDoS attacks. The results of their detection are displayed in table.



Algorithm	*Detected Attack number*	*Detection rate/%*	*Miss rate/%*	*False rate/%*	*Avera-ge delay/s*
DADA-HTPM	5	100	0	0	1.2
Wavelet Analysis[13]	9	100	0	44.4	2
VTP [12]	4	80	20	0	12

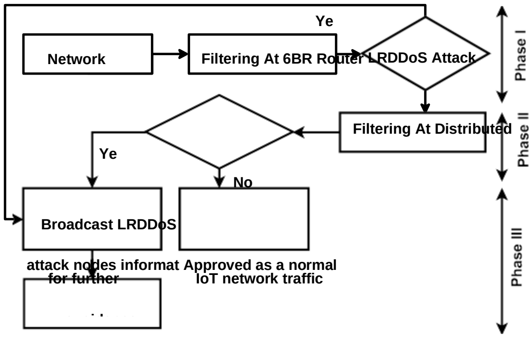
HTPM cannot predict the traffic accurately with high complexity

LORD: Low Rate DDoS Attack Detection and Mitigation Using Lightweight Distributed Packet Inspection Agent in IoT Ecosystem

Pradeepkumar Bhale

In this we innovatively propose lightweight distributed packet inspection agent which detects and alleviates L R DDoS **attack** in the IoT network. The approach is implemented in the Contiki OS, and the experimental outcomes exhibit that the proposed information metrics, and packet flow behavior graph can efficiently identify L R DDoS **attacks** with the minimum response time (1-3.47 seconds).

i i . 2

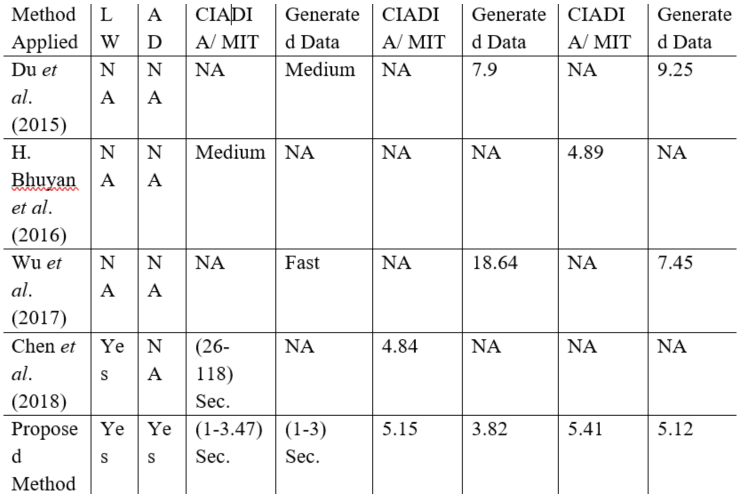


The benefit of this is Generation and analysis of attack and non-attack traffic, Distributed, Adaptability and Low false alarm rate

$$M_{TV(\phi_1,\phi_2)} = \frac{1}{2}\left(\sum_{i=1}^{k} |(\phi_1)_i (A_x) - (\phi_2)_i (A_y)|\right. \quad I_i = \frac{f_i}{\sum_{i=1}^{K} f_i} P(x_i) = \frac{x_i}{\sum_{i=1}^{k} x_i}$$

$$VM(\phi_1,\phi_2) = \chi\left(\sum_{i=1}^{k} (|(\phi_1)_i (A_x) - (\phi_2)_i (A_y)|)^{\mu}\right)^{\min\left(1,\frac{1}{\mu}\right)} \qquad (6)$$

the comparative study of the intended security solution with the existing security methods.



Method Applied	LW	AD	CIADIA/ MIT	Generated Data	CIADIA/ MIT	Generated Data	CIADIA/ MIT	Generated Data
Du *et al.* (2015)	NA	NA	NA	Medium	NA	7.9	NA	9.25
H. Bhuyan *et al.* (2016)	NA	NA	Medium	NA	NA	NA	4.89	NA
Wu *et al.* (2017)	NA	NA	NA	Fast	NA	18.64	NA	7.45
Chen *et al.* (2018)	Yes	NA	(26-118) Sec.	NA	4.84	NA	NA	NA
Proposed Method	Yes	Yes	(1-3.47) Sec.	(1-3) Sec.	5.15	3.82	5.41	5.12

limited study is available on the energy efficient lightweight solution to detect LR DDoS attacks.

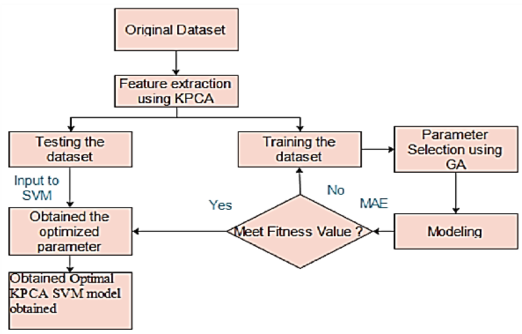
An Evolutionary SVM Model for DDOS Attack Detection in Software Defined Networks

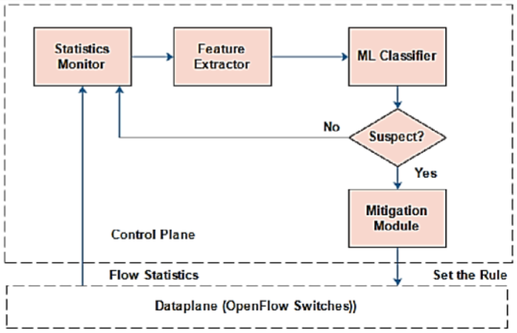
KSHIRA SAGAR

Previous research efforts have shown tremendous improvements in the control layer anomaly detection, but it lacks a detailed analysis
Techniques used were:

- Neural Network
- SVM
- Rate limitation and TRW CB

the goal of this paper is to detect the attack traffic, by taking the centralized control aspect of SDN.

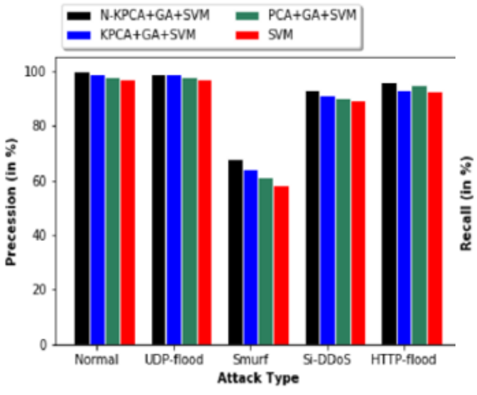

Original Dataset
Feature extraction using KPCA
Testing the dataset
Training the dataset
Parameter Selection using GA
Input to SVM
Yes
No
MAE
Obtained the optimized parameter
Meet Fitness Value ?
Modeling
Obtained Optimal KPCA SVM model obtained



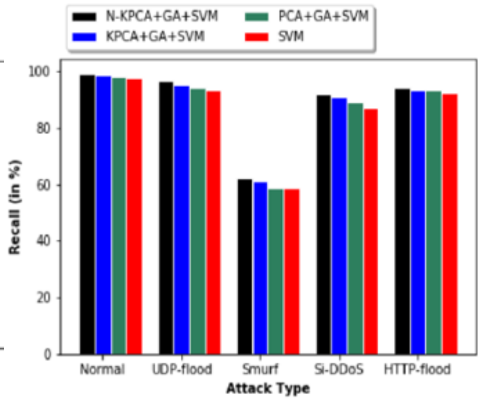
This work utilizes SVM technique as the prime classifier for predicting malicious traffic. The proposed detection approach combines SVM with KPCA and GA. The detection module is run over the controller.

COMPARISION WITH OTHERS

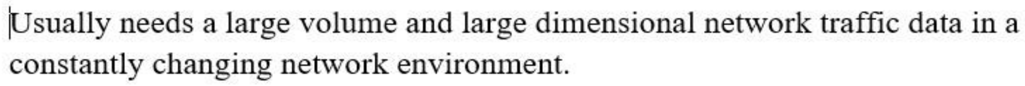
The goal is to monitor how normal traffic and other assault traffic are classified.



(a) Precision results of different models



(b) Recall results of different models



Usually needs a large volume and large dimensional network traffic data in a constantly changing network environment.

DoS/DDoS Attack Detection Using Artificial Neural Networks

Osman Ali et al. [2018]

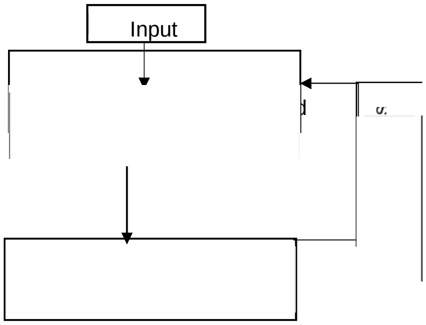
The lack traffic diversity and volumes, while others anonymized packet information and payload which cannot reflect the current trends, or lack of feature set and metadata.

This paper provides a machine learning approach to intrusion detection using Artificial Neural Networks (ANN). The Bayesian Regularization (BR) scaled conjugate gradient (SCG) descent backpropagation algorithms were used in the suggested method.

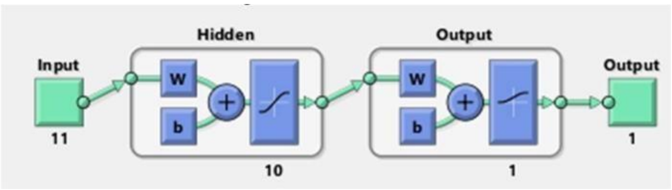
The suggested method successfully detected DoS/DDoS assaults with an accuracy of 99.6% using Bayesian Regularization and 97.7% using scaled conjugate gradient descent, according to the results.

$Accuracy = (TP + TN)/(TP + TN + FP + FN)$

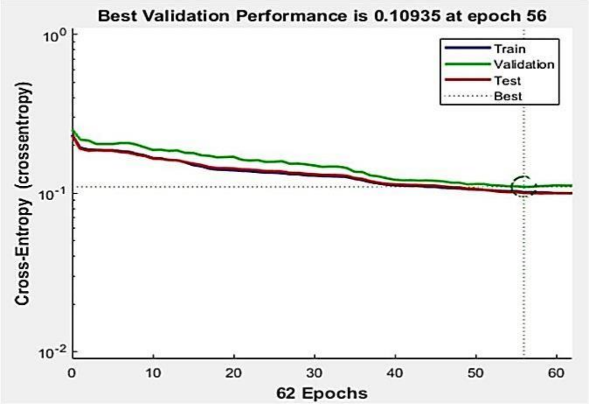
$Detection\ Rate = (TP)/(TP + FN)$

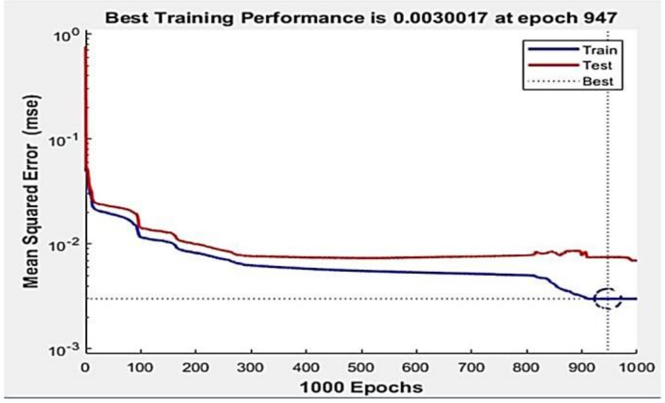


Backpropagation neural network



Comparison



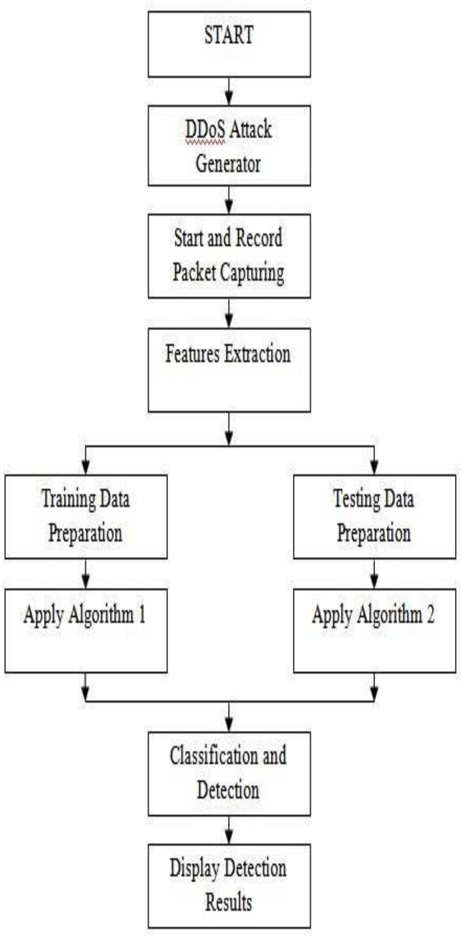


The lack of an up-to-date and credible dataset is also a constraint that prohibits academics from working thoroughly on the outdated dataset.

Web DDOS Attacks Detection Using Multinomial Classifier

Mrs. Shital K et al. [2016]

This method does not work over the scalability and early noise elimination. In this paper classifier based system is proposed in which packets are captures, extraction of important fields those are required for detection and then apply classifier to detection of attack. Multinomial Classifier is the technique employed in this paper. The multinomial Naive Bayes classifier is good for discrete feature classification


START
DDoS Attack Generator
Start and Record Packet Capturing
Features Extraction
Training Data Preparation
Testing Data Preparation
Apply Algorithm 1
Apply Algorithm 2
Classification and Detection
Display Detection Results

This technique is used because it has high accuracy, true positive & false positive rates then other methods

The main Algorithm used are Training Algorithm and Testing Algorithm

Accuracy= (TP+TN / TP + TN + FP + FN) * 100

Advantage

The main advantage of this that it has highly efficient technique of classification with aim of accuracy improvement., it shows the accuracy of approximately in between 85% - 90%

This approach also has the capability of sending out early warning notifications.

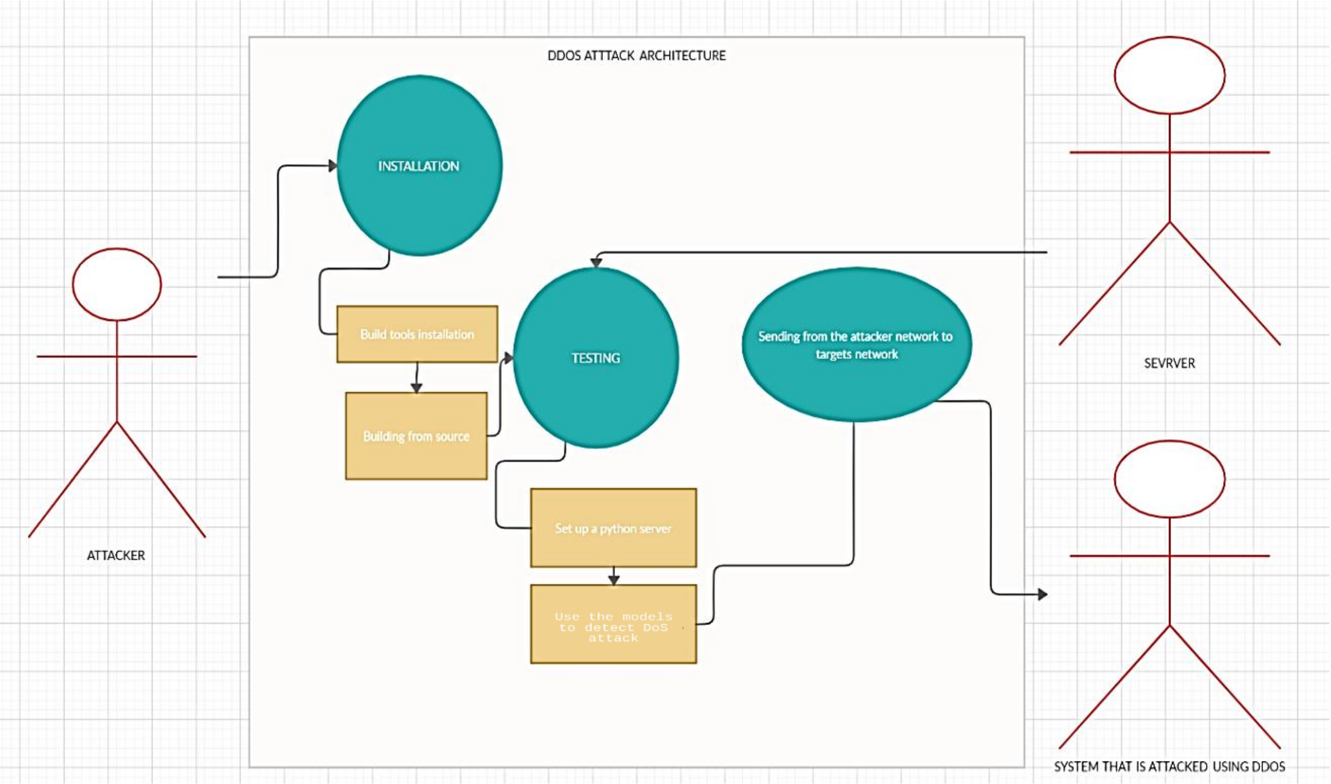
Comparison

The accuracy, true positive, and false positive rates of this technique are compared to those of other existing methods and papers. The existing approach limits and scope of improvement illustrated from the whole study and analysis are the results of this work.

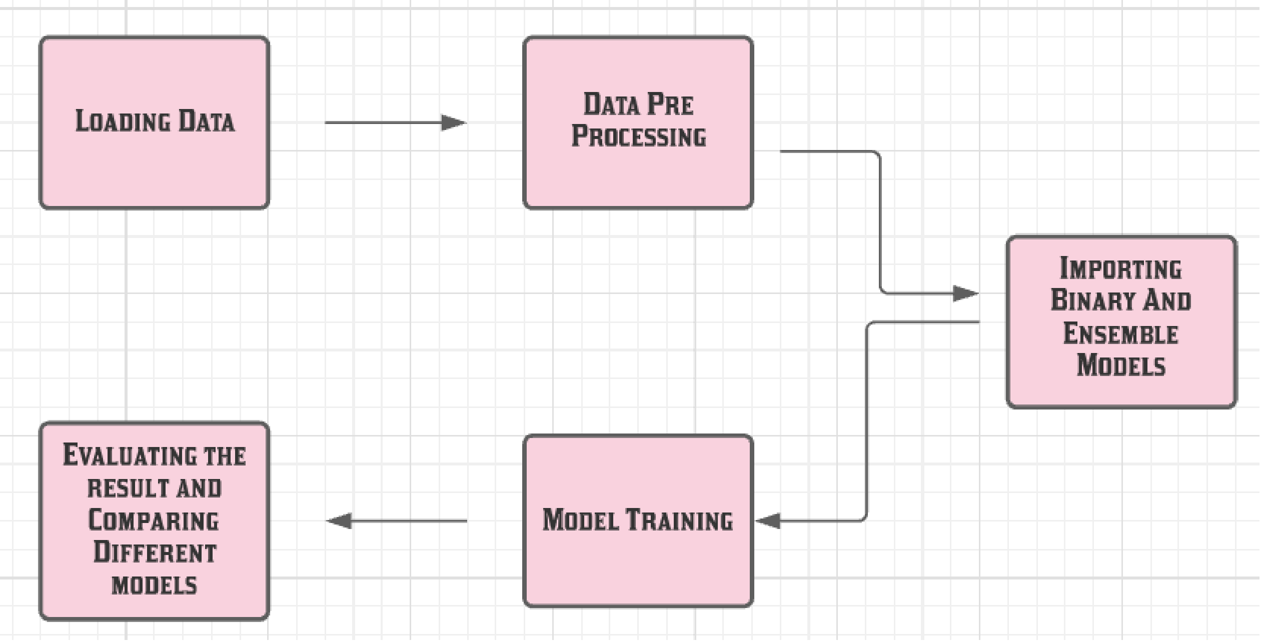
This is the just semi-Automated technique for DDoS attacks detection in large scale network

Proposed Work

Use Case Diagram


DDOS ATTTACK ARCHITECTURE
INSTALLATION
Build tools installation
Building from source
TESTING
Sending from the attacker network to targets network
Set up a python server
Use the models to detect DoS attack
ATTACKER
SEVRVER
SYSTEM THAT IS ATTACKED USING DDOS

Low Level Diagram

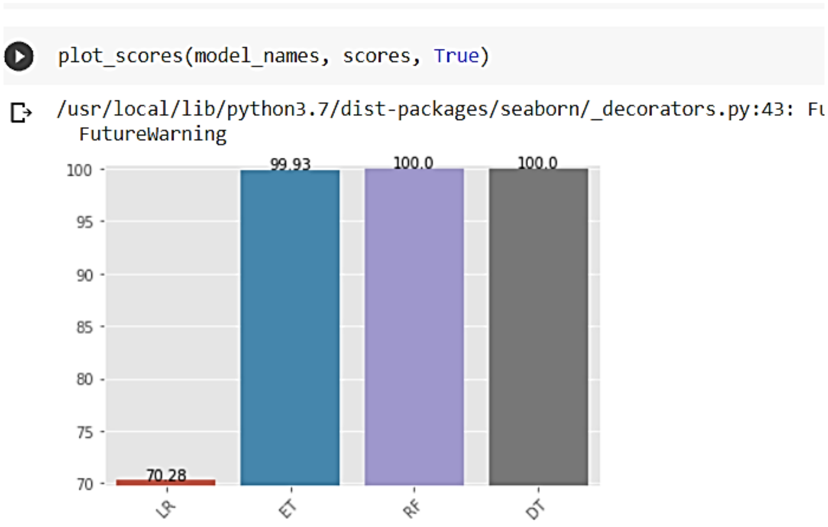

Loading Data
Data Pre Processing
Importing Binary And Ensemble Models
Model Training
Evaluating the result and Comparing Different models

In the above given method, we have added the dataset to our data frames and then we have pre processed the data in our data frames.

After this we have imported ensemble models. After this we have trained these models based on our training dataset. After this we have tested the data based on our testing data, based on which we have calculated the accuracy and precisions of our model.

Results

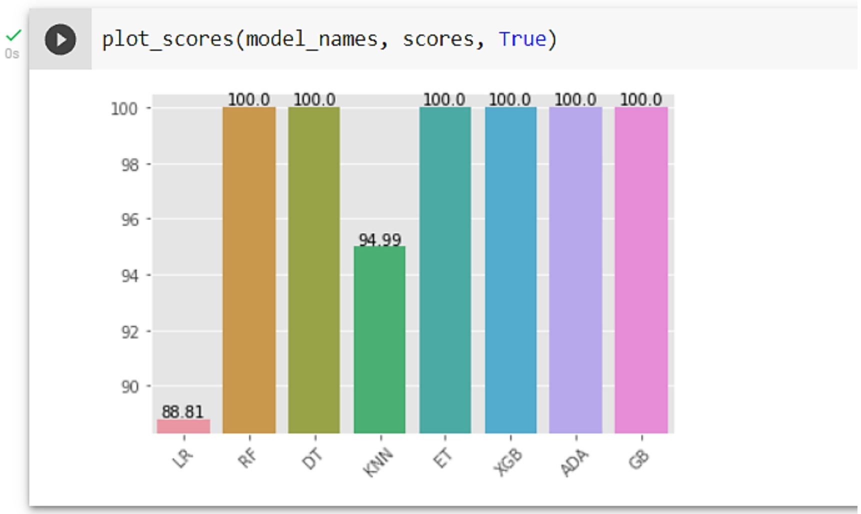
Review 2



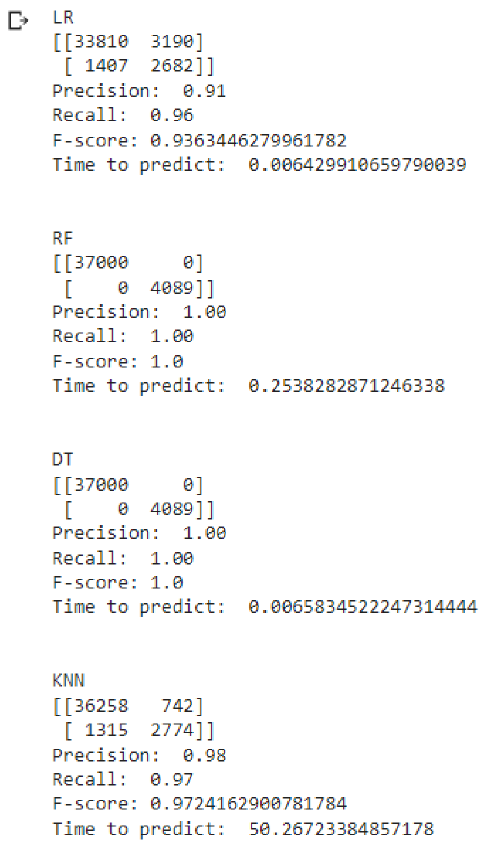
```
plot_scores(model_names, scores, True)
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: Fu
  FutureWarning
```

Review 3



```
plot_scores(model_names, scores, True)
```

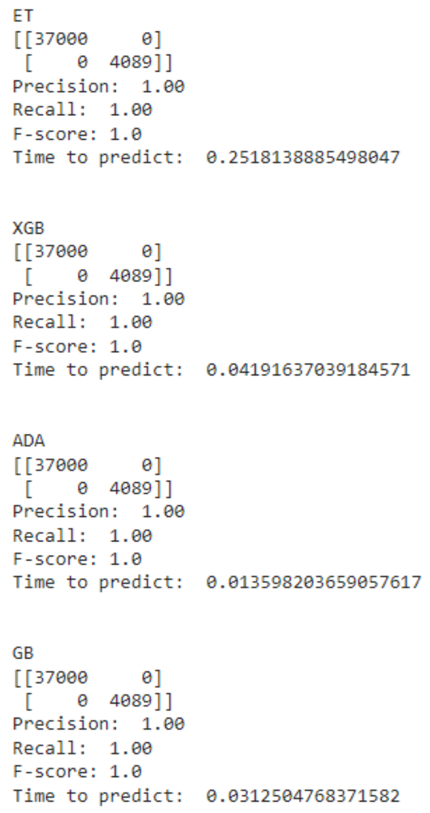


```
LR
[[33810  3190]
 [ 1407  2682]]
Precision:  0.91
Recall:  0.96
F-score: 0.9363446279961782
Time to predict:  0.006429910659790039


RF
[[37000     0]
 [    0  4089]]
Precision:  1.00
Recall:  1.00
F-score: 1.0
Time to predict:  0.2538282871246338


DT
[[37000     0]
 [    0  4089]]
Precision:  1.00
Recall:  1.00
F-score: 1.0
Time to predict:  0.0065834522247314444


KNN
[[36258   742]
 [ 1315  2774]]
Precision:  0.98
Recall:  0.97
F-score: 0.9724162900781784
Time to predict:  50.26723384857178
```



```
ET
[[37000     0]
 [    0  4089]]
Precision:  1.00
Recall:  1.00
F-score: 1.0
Time to predict:  0.2518138885498047


XGB
[[37000     0]
 [    0  4089]]
Precision:  1.00
Recall:  1.00
F-score: 1.0
Time to predict:  0.04191637039184571


ADA
[[37000     0]
 [    0  4089]]
Precision:  1.00
Recall:  1.00
F-score: 1.0
Time to predict:  0.013598203659057617


GB
[[37000     0]
 [    0  4089]]
Precision:  1.00
Recall:  1.00
F-score: 1.0
Time to predict:  0.0312504768371582
```

Tabled Result:

Models	Precision	Recall	F-score
Logistic Regression	0.91	0.96	0.936
Random Forest Classifier	1.00	1.00	1.00
Decision Tree Classifier	1.00	1.00	1.00
K Neighbours Classifier	0.98	0.97	0.972
Extra Trees Classifier	1.00	1.00	1.00
XGB Classifier	1.00	1.00	1.00
Ada Boost Classifier	1.00	1.00	1.00
Gradient Boosting Classifier	1.00	1.00	1.00

Conclusion:

From the above result table, we can see that ensemble models have performed far better than the classical models. Their accuracy and precision are far better than that of the classical models, in addition to which the prediction time is also less than those of the classical models.

Python Code

Ipynb file for ISAA Review

References

- DDoS Attack Detection Based on Simple ANN with SMOTE for IoT Environment **Yan Naung Soe, et al. [2020]**
- Machine Learning based DDOS Detection **S.Shanmuga Priya et al.[2020]**
- DDoS Intrusion Detection Through Machine Learning Ensemble **Saikat Das et al. [2019]**
- A DDoS Attack Detection Method Based on Hybrid Heterogeneous Multiclassifier Ensemble Learning **Bin Jia, et al. [2017]**
- A Novel DDoS Attack Detection Method Using Optimized Generalized Multiple Kernel Learning **Jieren Cheng1 et al. [2020]**
- A New Framework for DDoS Attack Detection and Defense in SDN Environment **Liang Tan et al. [2017]**
- Low-Rate DDoS Attack Detection Based on Factorization Machine in Software Defined Network **WU ZHIJUN et al. [2020]**
- A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks **SHAHZEB HAIDER et al. [2020]**
- IoT DoS and DDoS Attack Detection using ResNet **Faisal Abbas et al. [2020]**
- A DDoS Attack Detection System: Applying A Hybrid Genetic Algorithm to Optimal Feature Subset Selection **Abid Abbas et al. [2020]**
- DoS Attack Detection System using Apache Spark **Heena Mulla et al. [2021]**
- DDoS Attack Detection Algorithm Based on Hybrid Traffic Prediction Model **Yuze MENG et al. [2018]**
- LORD: Low Rate DDoS Attack Detection and Mitigation Using Lightweight Distributed Packet Inspection Agent in IoT Ecosystem **Pradeepkumar Bhale**
- An Evolutionary SVM Model for DDOS Attack Detection in Software Defined Networks **KSHIRA SAGAR**
- DoS/DDoS Attack Detection Using Artificial Neural Networks **Osman Ali et al. [2018]**
- Web DDOS Attacks Detection Using Multinomial Classifier **Mrs. Shital K et al. [2016]**