International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 4, Issue 4, April – 2020, Pages: 78-87

A Distributed Denial of Service Attack with IP Information Prediction Model Based on Bayesian Belief Network

Alile Solomon Osarumwense[1] and Otokiti Kareem Osayamen[2]

[1,2] Department of Computer Science, Faculty of Physical Sciences, University of Benin, Benin City, Edo State. P.M.B. 1154, Nigeria.
solomon.alile@physci.uniben.edu, kareem.otokiti@uniben.edu

Abstract: *DDoS attacks are attacks executed in stages with the sole aim of consciously restricting computing devices that receives from and provide services, resources to other devices on the network from exercising its responsibility as a client, service and resource provider by continuously engaging these computing devices with data sent simultaneously from advanced persistent threat (APT) devices. These attacks are so smartly designed that they are able to evade detection from most network instruction detection systems and they are capable of infiltrating complicated defenses. In this paper, we proposed and simulated a Bayesian Belief Network Model to predict DDoS Attacks with IP Information. The model was designed using Bayes Server and tested with data collected from cyber security repository. The model had a 99.47% prediction accuracy.*

Keywords: DDoS Attack; Prediction, IP Information; Bayesian Belief Network

1.0 INTRODUCTION

In today's world, the extensive use of internet has brought about simplicity in accessing information on servers, which in turn make services and resources available to other computing devices located on the same network via the network of networks [1]. Despite the benefits of unlimited access to data, service and resources made available on the network by these computing devices, they are subject to security threats that hinder access to the aforementioned benefits. With the amplified threat to network security, cybersecurity experts have tried to battle these increasing threats by designing complicated intrusion detection systems.

Of all these threats, the distributed denial of service attack has been regarded as the one of the underrated, ruinous form of attack which has amassed so much power that it causes unavailability of data; with the attack been utilized more by advanced persistent threat as affirmed [2,3,4].

Distributed Denial of Service (DDoS) attack is defined as a type of network attack that deliberately restricts computing devices that obtain from and render services, resources to other devices on the network from performing its role as a client, service and resource provider by bombarding these computing devices with data sent concurrently from advanced persistent threat (APT) devices [5,6].

Moreover, DDoS attacks are executed by advanced persistent threat in 4 stages or steps to exploit weaknesses of the system or network. The steps include lockdown the target computing device on the network, infect target host with malicious software for vulnerability, install malicious program on target host and execute network attack on the compromised host [7].

One dreadful trait of DDos attack is that when it attacks a host computing device on a network, there stand the chance of compromising the whole network. In cases of this magnitude, to safeguard the network, the compromised host is positioned on a restriction list called the blacklist.

In [8], a blacklist is defined as cluster of entities such as software programs, MAC addresses and IP addresses that are restricted from relating with other computing devices on a network. One way to identify computing devices with malicious intent is to utilize the IP address of the said device which distinctively recognizes the malicious device on the network. An IP address of computing device can be classified as malicious, if the device associated to a network via its means of identification is utilized in the perpetuation of a network attack. Attacks such as IP Spoofing attack, data manipulation attack and Reconnaissance attack just to name a few.

In [9], they reported a DDos attack on a popular news broadcasting organization. The servers of this organization were attacked which rendered all their websites unavailable, with the spread of this attack to the content services connected to the news agency servers which failed to load. The attack crippled the organization and they incurred a huge loss.

In time past, several techniques have been utilized in detecting DDos attacks in the works of [10, 11, 12,13,14,15, 16 and 17] but they generated a lot of false negative during testing and were unable to detect IP addresses involved in the perpetuation of DDoS attacks and its various kinds.

In this paper, Bayesian Belief Network (BBN) was utilized in detecting DDos attacks and its various kinds with IP information. BBN is a complex probabilistic network that combines expert knowledge and observed datasets. It maps out cause and effect relationships between variables and encodes them with probability that signify the amount in which one variable is probable to influence another. In this paper, BBN was our technique of choice because of its capability to make predictive inference.

2. Related Works

Several studies have been conducted on detecting DDos attacks on a network using Artificial Intelligence.

In [10], a system to detect DDos attacks which utilized Naives Bayes, K -Nearest neighborhood and Principle Component Analysis (PCA) was proposed. The system classified DDos attacks form traces of traffic flow. The system demonstrated a high level detection rate. However, the system failed to detect attacks perpetuated using malicious and normal IP addresses from where attack emanated from and destination IP address which is the target host address. The K-Nearest neighbor algorithm computation cost is very high, it does not perform well with large dataset and also the algorithm can handle noisy data and missing values in dataset; Principal components in PCA are not easily interpretable and understandable.

In [11], a system to predict zombies involved in a perpetuated DDos attack using artificial neural network (ANN) scheme was developed. The system showed the capability of detecting zombies in the execution of DDos attack with a high detection rate. However, the system failed to detect malicious and normal IP addresses of the zombies utilized in the perpetration of a DDos attack and different types of DDos attacks. The system network learning process is time-consuming and the solutions derived from the learning process are usually difficult to interpret.

In [12], a system that employed Artificial Neural Network algorithm to detect DDos attacks. The system demonstrated the ability to detect DDos attacks and some of its type based on characteristic patterns to differentiate genuine traffic from DDos attacks with a 98% detection accuracy. However, the system failed to detect malicious and normal IP addresses utilized to perpetrate this attack. The ANN algorithm learning process is time-consuming with the solution derived difficult to comprehend.

In [13], a system that combined artificial neural network (ANN) and chaos theory to develop a system that detected DDos attacks was proposed. The system showcased its ability to differentiate DDos attacks traffic from normal traffic. However, the system failed to detect various kinds of DDos attacks and malicious and normal IP addresses employed in execution of the said attack. Furthermore, the ANN algorithm learning process takes quite a long time with guarantee of success and the solution produced is difficult to comprehend; solutions produced by chaos theory loses long term predictability.

In [14], a lightweight deep learning DDoS detection system called LUCID which is based on conventional neural network was designed. The system was able to classify malicious network traffic from actual traffic with high detection accuracy. However, the system failed to detect various types of DDos attacks and IP addresses utilized to perpetrate the attack, the system neural network is difficult to interpret, requires large data to train, time and capital intensive and utilizes a lot of memory to execute the network.

A Hybrid system to detect DDos Attacks which utilized Multilayer Perceptron (MLP), Naïve Bayes and Random Forest was proposed in [15]. The system detected DDos attacks and classified 4 types of DDos attacks with a 98.63% detection accuracy. However, the system failed to detect IP addresses employed to perpetrate a DDoS attack, detect other types of DDoS attacks. Multilayer Perceptron algorithm is time-consuming, capital intensive and requires a lot of memory to actualize the network; Random Forest Algorithm solution is not easily interpretable and requires a lot of time during creation of large trees for learning.

In [16], a system that utilized random forest algorithm model based on machine learning was developed. The system detected DDoS attack with a good detection rate. However, the system failed to detect the different kinds of DDos attacks and IP addresses involved in the perpetration of the said attack. Also, the solutions produced using random forest algorithm is quite difficult to deduce and takes a long time to create large trees for learning.

In [17], a prototype system which employed Support Vector Machine (SVM) to detect DDos Attacks was developed. The prototype showcased the ability to detect DDos with a high detection accuracy of 99%. However, the system failed to detect the different types of DDos Attacks and IP addresses involved in perpetrating the said attack. Also, the SVM algorithm is not suitable for handling large datasets and solving a target problem that has overlapping features.

3. Bayesian Belief Network

Bayesian Belief Network (BBN) is directed acyclic graphical model that employs probability to demonstrate conditional dependencies that prevail amongst nodes on a graph [18]. It is a complex probabilistic network that merges expert knowledge and experimental datasets. It plans out route of cause and effect relationships between variables and encodes them with probability that signify the amount in which one variable is probable to sway another. Bayesian Belief Network is based on the Bayes theorem which relies on probability.

The Bayes theorem is represented in the mathematical equation below:

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)} \tag{1}$$

Where,
P(a) is the probability of event “a” happening without any information about event “b”. It is called the “Prior”.
P(a/b) is the conditional probability of event “a” happening given that event “b” has already occurred. It is otherwise called the “Posterior”.
P(b/a) is the conditional probability of event “b” happening given that event “a” has already occurred. It is called the “Likelihood”.
P(b) is the probability of event “b” happening without any information about event “a”. It is called the “Marginal Likelihood”.
The Naive Bayes classifiers are often represented as a type of directed acyclic graph (DAG). The Directed Acyclic Graph (DAG) comprises of vertices representing random variables and arrows connecting pairs of nodes.
Figure 1 shows a pictorial representation of a Bayesian Belief Network.

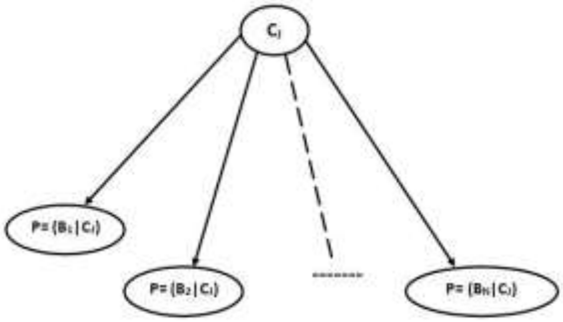


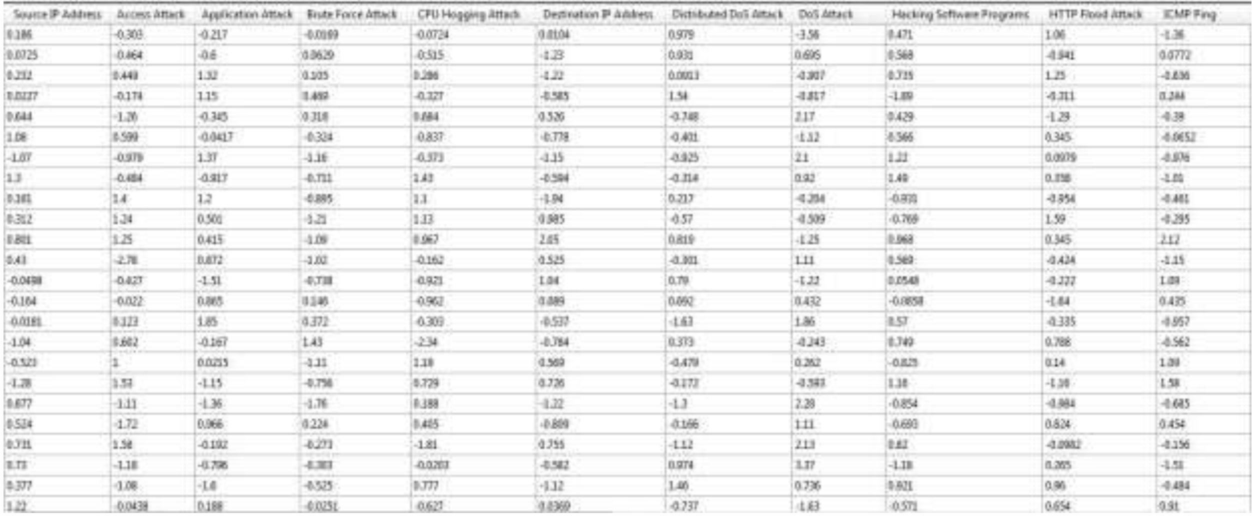
Figure 1: A Pictorial Representation of a Bayesian Belief Network

Some advantages of this model are: it is pretty fast in making inferences, the resulting probabilities are easy to interpret, the learning algorithm is quite simple to comprehend and the model adequately combines with utility functions to make optimal inferences. In this paper, we intend to detect DDos attacks and its various types with IP information using Bayesian Belief Network (BBN).
A model consisting of 25 nodes where each node represents a form of attack will be designed using Bayes Server. A cybersecurity dataset will be used to train and test the system. Using the Pareto Principle, 80% of the dataset will be used to train the model while the remainder will be used in testing the model. The aim of the model is to achieve a high level of detection accuracy with the use of IP information and various types of DDos attacks.

4. Methodology

Simulation, Result and Discussion

The dataset used in training, testing and predicting DDoS attacks with IP information was retrieved from [19]. The dataset consist of 25 attacks and each attack has a value which represents the probability of such attack in causing DDos attack. These attacks are These attacks are Access Attack (AA), Application Attack (APPA), Brute Force Attack (BFA), CPU Hogging Attack (CHA), Distributed Denial Of Service (DDoS), Denial of Service (DoS), ,Hacking Software Programs (HSP), Hyper Text Transfer Protocol Flooding Attack (HTTPA), Internet Control Message Protocol Packet Internet Groper (ICMP Ping), IP Spoofing Attack (IPSA), Network Ping (NP), Network Time Protocol Amplification Attack (NTPAA), Ping of Death Attack (PINGDA), Reconnaissance Attack (RA), Rerouting Attack (RERA), Session Attack (SESA), Slowloris Attack (SLOWA), Smurf Attack (SMA), Tribe Flood Network Attack (TFNA), Tribe Flood Network 2000 Attack (TFN2A), Trinity Attack (TRINA), User Datagram Protocol Flood Attack (UDPFA), Tribe Flood Network Attack (TFNA), and two columns which indicates the class Source IP Address (S.IPAddr) and Destination IP Address (D.IPAddr) associated to DDoS attack. The figure 2 below displays a sample the dataset.



Source IP Address	Access Attack	Application Attack	Brute Force Attack	CPU Hogging Attack	Destination IP Address	Distributed DoS Attack	DoS Attack	Hacking Software Programs	HTTP Flood Attack	ICMP Ping
0.186	-0.303	-0.217	-0.0169	-0.0724	0.0104	0.979	-3.56	0.471	1.06	-1.36
0.0725	-0.464	-0.6	0.0629	-0.515	-1.23	0.931	0.695	0.568	-0.941	0.0772
0.232	0.448	1.32	0.105	0.286	-1.22	0.0013	-0.307	0.735	1.25	-0.836
0.0227	-0.174	1.15	0.469	-0.327	-0.585	1.54	-0.817	-1.89	-0.311	0.244
0.644	-1.26	-0.345	0.318	0.684	0.526	-0.748	2.17	0.429	-1.29	-0.38
1.06	0.599	-0.0417	-0.324	-0.837	-0.778	-0.401	-1.12	0.566	0.345	-0.0652
-1.07	-0.979	1.37	-1.16	-0.373	-1.15	-0.925	2.1	1.22	0.0979	-0.876
1.3	-0.484	-0.917	-0.711	1.43	-0.584	-0.314	0.92	1.49	0.358	-1.01
0.381	1.4	1.2	-0.885	1.1	-1.94	0.217	-0.284	-0.931	-0.954	-0.461
0.312	1.24	0.501	-1.21	1.13	0.985	-0.57	-0.509	-0.769	1.59	-0.295
0.801	1.25	0.415	-1.09	0.967	2.65	0.819	-1.25	0.968	0.345	2.12
0.43	-2.78	0.872	-1.02	-0.162	0.525	-0.301	1.11	0.569	-0.424	-1.15
-0.0498	-0.427	-1.51	-0.738	-0.921	1.04	0.79	-1.22	0.0548	-0.227	1.09
-0.164	-0.022	0.865	0.146	-0.962	0.089	0.092	0.432	-0.0858	-1.84	0.435
-0.0181	0.123	1.85	0.372	-0.303	-0.537	-1.63	1.86	0.57	-0.335	-0.957
-1.04	0.602	-0.167	1.43	-2.34	-0.784	0.373	-0.243	0.749	0.788	-0.562
-0.521	1	0.0255	-1.11	1.18	0.569	-0.479	0.262	-0.825	0.14	1.09
-1.28	1.53	-1.15	-0.756	0.729	0.726	-0.172	-0.583	1.18	-1.18	1.58
0.677	-1.11	-1.36	-1.76	0.188	-1.22	-1.3	2.28	-0.854	-0.884	-0.685
0.524	-1.72	0.966	0.224	0.465	-0.889	-0.166	1.11	-0.693	0.824	0.454
0.731	1.58	-0.192	-0.273	-1.81	0.755	-1.12	2.13	0.82	-0.0982	-0.156
0.73	-1.18	-0.796	-0.383	-0.0203	-0.582	0.974	3.37	-1.18	0.265	-1.51
0.377	-1.08	-1.6	-0.525	0.777	-1.12	1.46	0.736	0.921	0.96	-0.484
1.22	-0.0438	0.188	-0.0251	-0.627	0.0369	-0.737	-1.63	-0.571	0.654	0.91

Figure 2: Snapshot of Dataset

The Bayesian model was designed using Bayes-Server platform. The Bayesian Belief Network for predicting
DDos attack with IP information was designed such that the nodes on the network are linked based on the probability of an attack resulting to another. To blacklist an IP address on the network, we analyze the attack perpetuated by the device. In our model for a source and destination IP address to be denoted as a malicious IP, such IP must have performed any of the following attacks; Access Attack, Application Attack, Brute Force Attack , CPU Hogging Attack, Distributed Denial Of Service, Denial of Service , Hacking Software Programs, Hyper Text Transfer Protocol Flooding Attack , Internet Control Message Protocol Packet Internet Groper, IP Spoofing Attack, Network Ping, Network Time Protocol Amplification Attack, Ping of Death Attack, Reconnaissance Attack, Rerouting Attack , Session Attack , Slowloris Attack, Smurf Attack, TCP Syn Flood Attack,Tribe Flood Network Attack , Tribe Flood Network 2000 Attack, Trinity Attack, User Datagram Protocol Flood Attack and Tribe Flood Network Attack respectively. The source and destination (malicious) IP address are categorized into various classes with each class indicating the hazardous level of such IP address.

Figure 3 shows the BBN model for detecting DDoS attack with IP Information.

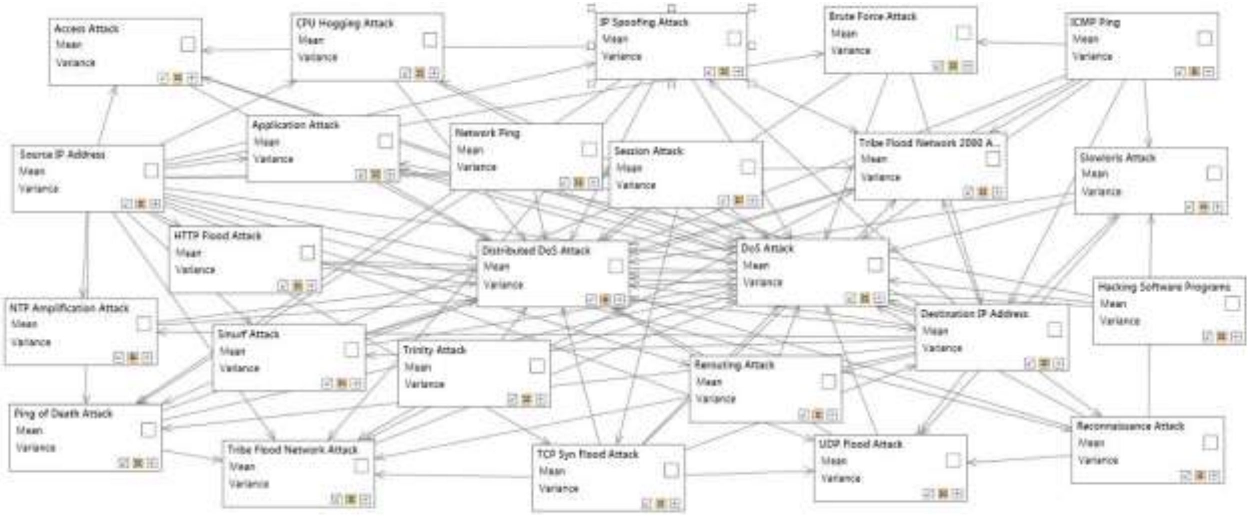


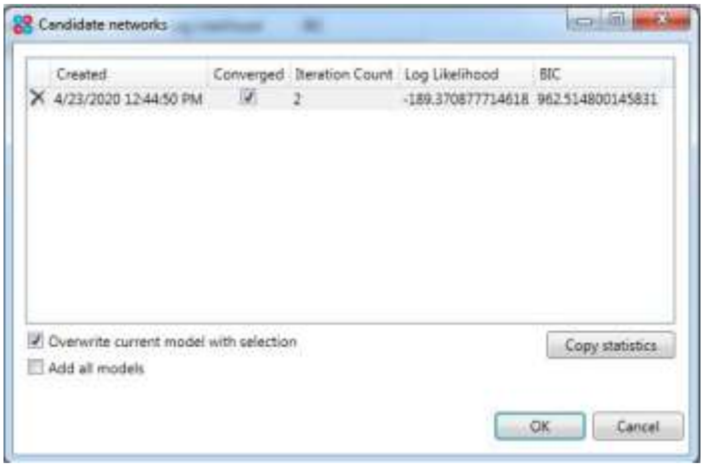
Figure 3: Bayesian Belief Network Model for Detecting DDoS attacks with IP Information.

So, to mathematically represent our model we have:

$$\text{DDos Attack} = \prod_{i=1}^{25} P(\text{Attack}_i | \text{Parents}(\text{Attack}_i) \qquad (2)$$

Where,
Attack: Node with an attack
Parents (Attacki) = Nodes that converge on Attacki..

The dataset was used to train and test the model. Upon completion of training and testing the BBN model, the test data converged at time series 2. The log likelihood value for each case was recorded.



Candidate networks

	Created	Converged	Iteration Count	Log Likelihood	BIC
✗	4/23/2020 12:44:50 PM	☑	2	-189.370877714618	962.514800145831

☑ Overwrite current model with selection
☐ Add all models

Figure 4: Convergence of Bayesian Belief Network Model for DDoS Attack with IP Information at Time Series 2.

Figures 5, 6,7,8,9 and10 shows log likelihood batch query chart for predicting DDoS attack with IP Information, feature importance chart for nodes in the model, the in-sample anomaly detection chart, the log likelihood attack graph for detecting DDoS attack with IP Information, the mesh query plot for the loglikelihood chart and Multi-stage attack detection results chart respectively.

The result generated from the simulation indicated that the network was able to predict 99% DDoS attack with IP information on the dataset accurately and it had a log likelihood of -27.81 on the test dataset.

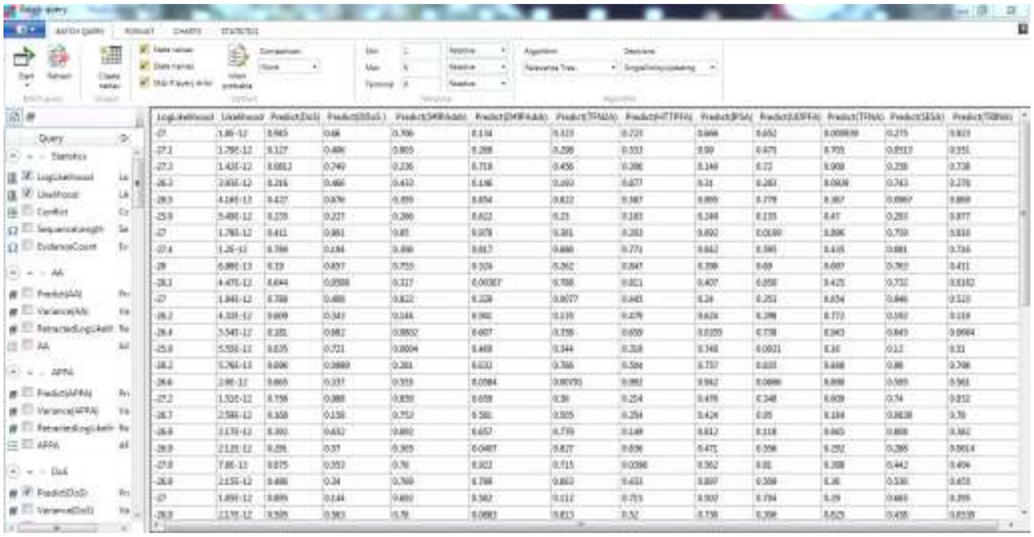


Figure 5: The Loglikelihood Chart Batch Query for Detecting DDoS Attack with IP Information

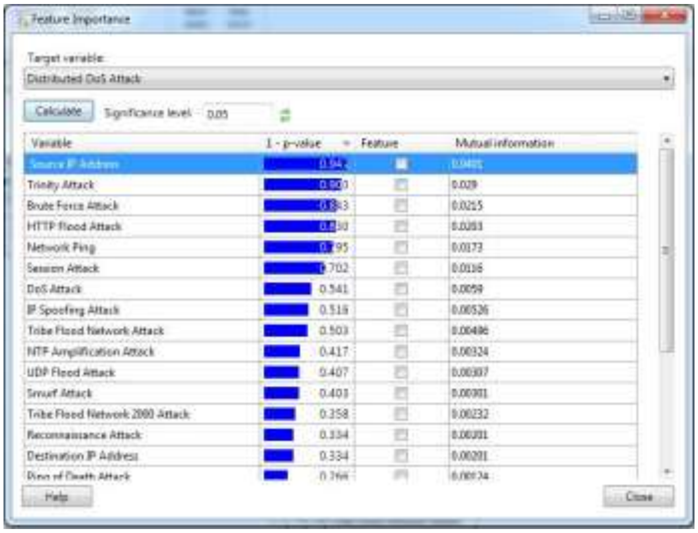
This loglikelihood chart batch shows the result of the test data.

In Experiment 1: The value of Predict(DDoS) was 0.66 compared to 0.660270987378089, Predict(S.IPAddr)=0.706 compared to 0.706222843330884 and Predict(D.IPAddr)=0.134 compared to 0.133532047308253.

Experiment 2: The value of Predict(DDoS) was 0.406 compared to 0.405827262132895, Predict(S.IPAddr)=0.805 compared to 0.804881363796469 and Predict(D.IPAddr)=0.268 compared to 0.267713508435473.

Experiment 3 The value of Predict(DDoS) was 0.749 compared to 0.749225924487481, Predict(S.IPAddr)=0.236 compared to 0.236028234904242 and Predict(D.IPAddr)=0.719 compared to 0.719412038663326 up to Experiment 50. Hence, the system results showed a 0.01% value difference between the prediction results and original test data of 100% resulting to 99% prediction accuracy.

The figure 6 below shows the Feature Importance Chart for Nodes in the Model



Feature Importance

Target variable: Distributed DoS Attack

Calculate Significance level: 0.05

Variable	1 - p-value	Feature	Mutual information
Source IP Address	0.942	☑	0.0401
Trinity Attack	0.900	□	0.029
Brute Force Attack	0.883	□	0.0215
HTTP flood Attack	0.830	□	0.0203
Network Ping	0.795	□	0.0173
Session Attack	0.702	□	0.0116
DoS Attack	0.541	□	0.0059
IP Spoofing Attack	0.519	□	0.00526
Tribe Flood Network Attack	0.503	□	0.00486
NTP Amplification Attack	0.417	□	0.00324
UDP Flood Attack	0.407	□	0.00307
Smurf Attack	0.403	□	0.00301
Tribe Flood Network 2000 Attack	0.358	□	0.00232
Reconnaissance Attack	0.334	□	0.00201
Destination IP Address	0.334	□	0.00201
Ping of Death Attack	0.266	□	0.00124

Figure 6: The Feature Importance Chart for Nodes in the Model

The Feature Importance Chart shows p-value of the variable (nodes), Feature and Mutual information in reference to the multi-stage attack Node.

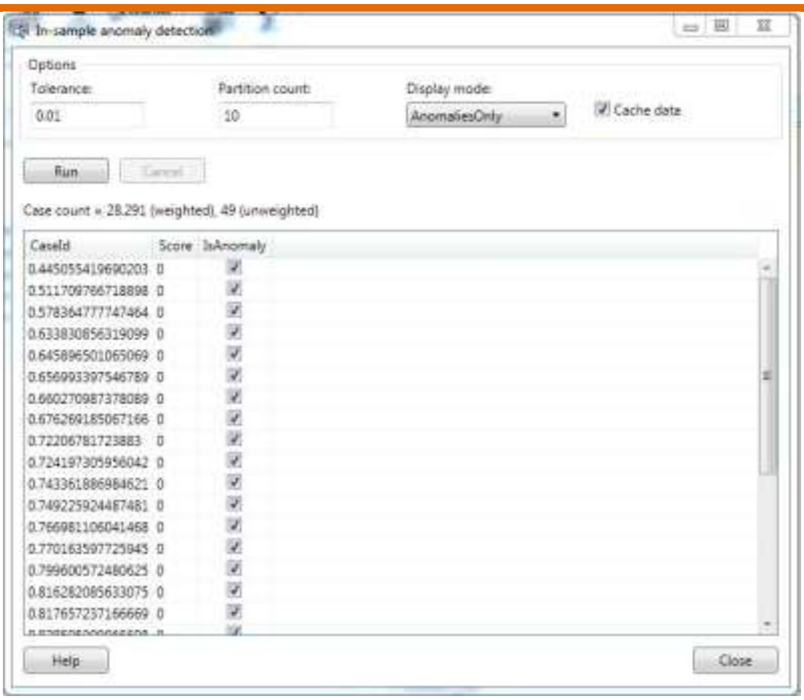
The p-value signifies the likelihood (probability)of the nodes being involved in the execution of a DDoS attack.

The Feature box shows if that particular node is fully involved in the said attack.

The Mutual information shows the relationship with nodes directly connected to one another (i.e. in this case the direct relationship of the nodes with the DDoS attack node)and assigned a value.

The Significance Level signifies the margin of error in the detection of DDoS attack.

The figure 7 below showcases the In-sample Anomaly Detection Chart.



Tolerance: 0.01
Partition count: 10
Display mode: AnomaliesOnly
☑ Cache date

Case count = 28.291 (weighted), 49 (unweighted)

CaseId	Score	IsAnomaly
0.445055419690203	0	☑
0.511709766718898	0	☑
0.578364777747464	0	☑
0.633830856319099	0	☑
0.645896501065069	0	☑
0.656993397546789	0	☑
0.660270987378089	0	☑
0.676269185067166	0	☑
0.72206781723883	0	☑
0.724197305956042	0	☑
0.743361886984621	0	☑
0.749225924487481	0	☑
0.766981106041468	0	☑
0.770163597725945	0	☑
0.799600572480625	0	☑
0.816282085633075	0	☑
0.817657237166669	0	☑

Figure 7: The In-sample Anomaly Detection Chart

The In-sample Anomaly Detection Chart shows 50 experimental results of detecting DDoS attack. Each Case is assigned an ID(Identification value) which is the value of the Predict(DDoS) in Figure 5. The IsAnomaly checkbox is checked to identify that each case is an executed DDos attack. The 50 cases of DDoS attack has a case count value of 28.291 (weighted) which signifies the importance of the cases leading to a execution of a multi-stage attack and 49 case count value signifies the number of cases in the pool of data available to the system for detection of DDos attack excluding the DDoS attack column in the dataset pool. The tolerance is the margin of error that could be encoutered as regards to the detection of the DDoS attacks.

The figure 8 shows the Loglikelihood Attack Graph for Detecting DDoS Attack with IP Information

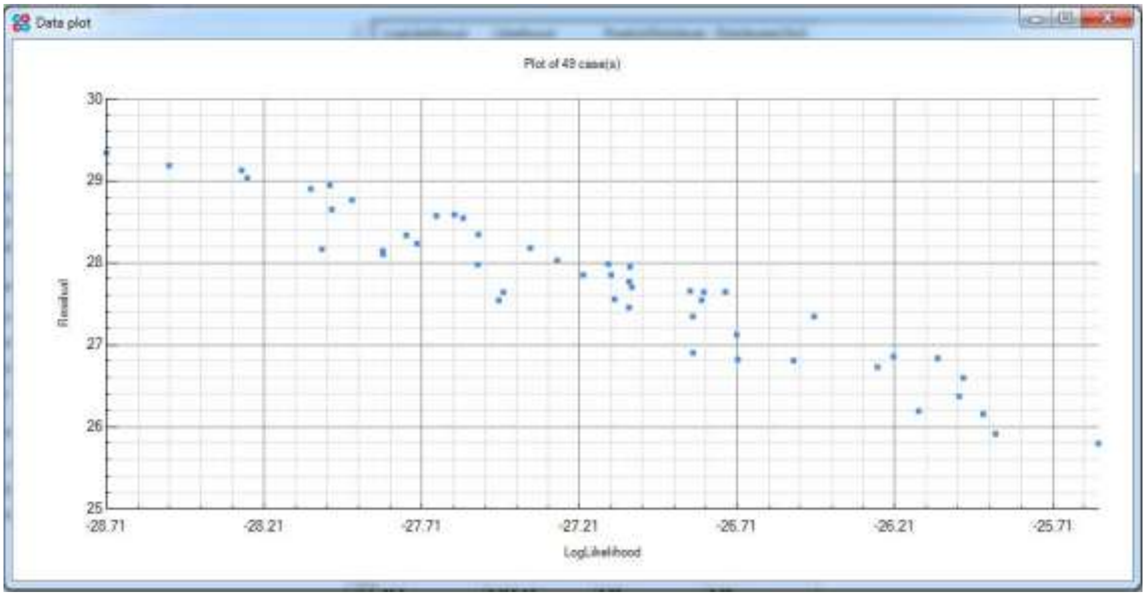


Figure 8: The Loglikelihood Attack Graph for Detecting DDoS Attack with IP Information

This loglikelihood graph for detecting DDoS attack with IP Information displays the residual values on the vertical axis plotted against the loglikelihood values on the horizontal axis which are independent variables. A residual value is a measure of how much a regression line vertically misses a data point. Regression lines are the best fit of a set of data. The lines are categorized as averages; a few data points will fit the line and others will miss.

In this graph, it shows that 50 experimental cases resulted in value of 29.28, 29.24, 29.19, 29, 28.98, 28.96…….. and 25.95 respectively.

Ideally, residual values should be equally and randomly spaced around the horizontal lines. Evaluating the system' experimental results values obtained from the horizontal lines on the graph, it can be seen that the point where the highest residual value and the loglikelihood independent variable attained meets at -28.71 on the horizontal line with 30 being the highest value that can be reached on the vertical line.

The residual value attained is 29.28 and loglikelihood independent value is -28.71, the difference between both values is 0.57 which is the predicted value.

Hence, in this system the highest residual value, a loglikelihood independent value can attain is 30. With 30, being the 100 % residual value mark, to get our prediction accuracy percentage, we have highest residual value subtracted from predicted value i.e. 100% -0.57= 99.47% residual loglikelihood percentage value.

The figure 9 displays the Loglikelihood of a Single Attack (Ping of DeathAttack, [PingofDA]) Being involved in a DDoS Attack.

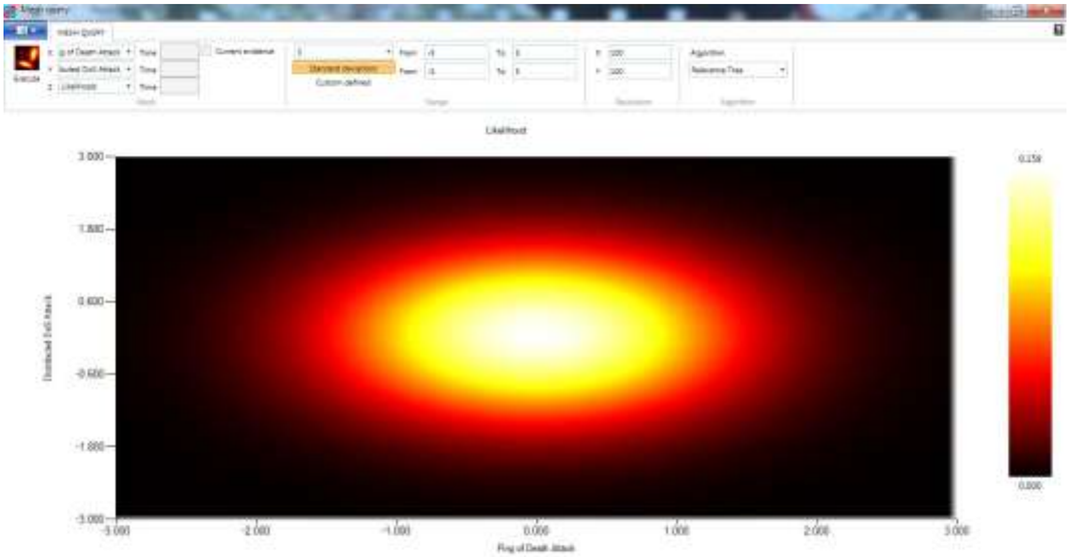


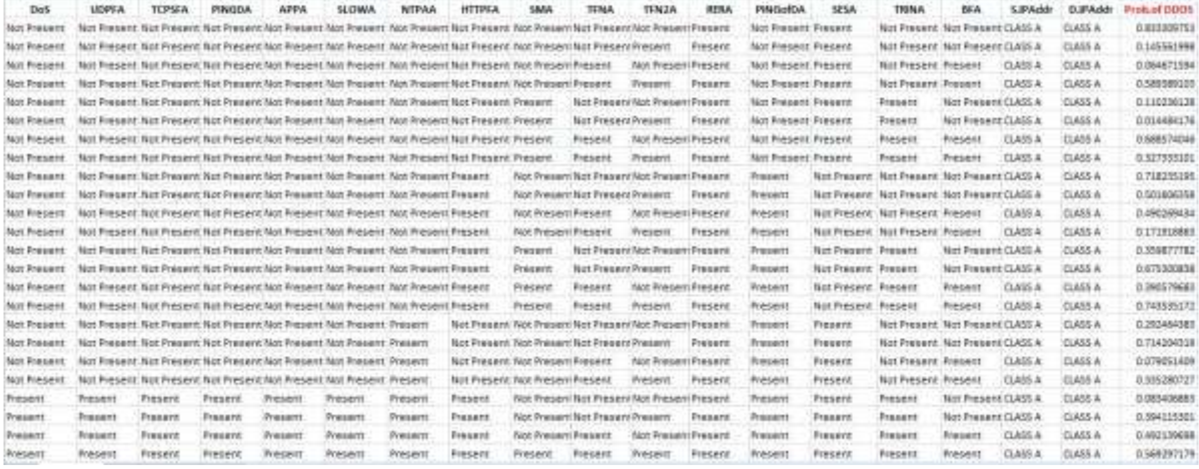
Figure 9: The Mesh Query Plot for the Loglikelihood of a Single Attack (Ping of Death Attack, [PingofDA]) Being involved in a DDoS Attack.

The mesh query plot shows the loglikelihood/likelihood of a node in this case Ping of Death Attack (PingofDA) being involved in the execution of a DDoS attack. The Node (DDoS attack) is plotted on the Y-axis and the other node Access attack (AA) plotted along the X-axis.

In this context, the Red contour signifies the likelihood of an AA being involved in a DDoS attack with the contour ranging from interval (-1.600 to -1.600) on the Y-axis and interval (-1.500 to1.500) on the X-axis.

The Yellow contour shows the loglikelihood of an AA being involved in a DDoS attack with the contour ranging from interval (-0.600 to 0.600) on the Y-axis and interval (-1.000 to 1.000) on the X-axis.

The figure 10 below showcases the DDos attack with IP Information Prediction Results.



DoS	UDPFA	TCPSFA	PINGDA	APPA	SLOWA	NTPAA	HTTPFA	SMA	TFNA	TFN2A	RERA	PINGofDA	SESA	TRINA	BFA	S.IPAddr	D.IPAddr	Prob.of DDOS
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Present	Not Present	Not Present	CLASS A	CLASS A	0.833309753
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Not Present	Present	Not Present	Not Present	CLASS A	CLASS A	0.145561998
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Present	Not Present	Present	Not Present	Present	CLASS A	CLASS A	0.084671594
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Present	Not Present	Present	Not Present	Present	CLASS A	CLASS A	0.588589120
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Present	Not Present	Present	Present	Not Present	CLASS A	CLASS A	0.110236128
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Present	Present	Not Present	Present	Present	Not Present	CLASS A	CLASS A	0.014484176
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Not Present	Present	Not Present	Present	Present	Present	CLASS A	CLASS A	0.688574046
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Present	Present	Not Present	Present	Present	Present	CLASS A	CLASS A	0.327355101
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Not Present	Present	Present	Not Present	Not Present	Not Present	CLASS A	CLASS A	0.718235195
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Present	Present	Present	Not Present	Not Present	Not Present	CLASS A	CLASS A	0.501806358
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Present	Not Present	Present	Present	Not Present	Not Present	Present	CLASS A	CLASS A	0.490269484
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Present	Present	Present	Present	Not Present	Not Present	Present	CLASS A	CLASS A	0.171816883
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Not Present	Not Present	Present	Present	Not Present	Present	Not Present	CLASS A	CLASS A	0.358877782
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Not Present	Present	Present	Present	Not Present	Present	Not Present	CLASS A	CLASS A	0.675300838
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Present	Not Present	Present	Present	Not Present	Present	Present	CLASS A	CLASS A	0.390579683
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Present	Present	Present	Present	Not Present	Present	Present	CLASS A	CLASS A	0.743535172
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Not Present	Not Present	Present	Present	Present	Not Present	Not Present	CLASS A	CLASS A	0.292484383
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Not Present	Present	Present	Present	Present	Not Present	Not Present	CLASS A	CLASS A	0.714204318
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Present	Not Present	Present	Present	Present	Not Present	Present	CLASS A	CLASS A	0.079051409
Not Present	Not Present	Not Present	Not Present	Not Present	Not Present	Present	Not Present	Not Present	Present	Present	Present	Present	Present	Not Present	Present	CLASS A	CLASS A	0.335280727
Present	Present	Present	Present	Present	Present	Present	Present	Not Present	Not Present	Not Present	Present	Present	Present	Present	Not Present	CLASS A	CLASS A	0.083406885
Present	Present	Present	Present	Present	Present	Present	Present	Not Present	Not Present	Present	Present	Present	Present	Present	Not Present	CLASS A	CLASS A	0.594115301
Present	Present	Present	Present	Present	Present	Present	Present	Not Present	Present	Not Present	Present	Present	Present	Present	Present	CLASS A	CLASS A	0.492139698
Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	Present	CLASS A	CLASS A	0.569297179

Figure 10: DDoS Attack with IP Information Prediction Results Chart

This chart shows the probabilities of network attacks resulting to a DDos attack namely Application Attack, Brute Force Attack, Denial Of Service, Hyper Text Transfer Protocol Flooding Attack, Network Time Protocol Amplification Attack, Ping of Death Attack, Rerouting Attack, Session Attack , Slowloris Attack, Smurf Attack, TCP Syn Flood Attack, Tribe Flood Network Attack, Tribe Flood Network 2000 Attack, Trinity Attack and User Datagram Protocol Flood Attack with IP information (class of source and destination IP addresses) respectively
This detection results showed the probability of having all the aforementioned attacks involved in a DDoS attack denoted as:
P(DDoS Attack| BFA, DoS, HTTPA, NTPAA, PingofDA, RERA, SESA, SLOWA, SMA, TCPSynFA, TFNA, TFN2A, TRINA,UDPFA) = 0.569297179
From the experiment it can be seen that our model has a higher residual log likelihood value which is 29.28 and a prediction accuracy of 99.47%. Comparing the log likelihood from the experiments conducted by [12, 15, 17] which are 98%, 98.63% and 99% respectively, it is obvious our model has a better prediction accuracy. The higher prediction accuracy achieved by our model could be due to the size of the dataset used in training and testing the model.

5. Conclusion

DDoS attack detection is very complicated because of its smart design. To safeguard a network, network security experts need to improve on existing technologies for detecting DDoS attacks. In this paper, we utilized a Bayesian Belief Network model to predict DDoS attack. The network had 25 nodes with each node representing a unique attack. The model was trained and tested and it had an accuracy of 99.47% in predicting DDoS attack with IP Information. The system can be deployed on computer network infrastructures to provide information which will be used to safeguard computer networks. It will also bring about improvement in the following areas: DDoS attack prediction, DDoS attacks detection, blacklisting of malicious IP addresses and computer network security in general. Future research should be geared towards improving DDoS attacks prediction using MAC address and devices that utilizes VPN.

REFERENCES

[1]Theamegroup (2020): "Network Security Threats:5 ways to protect Yourself". Retrieved from URL: https:// wwww.theamegroup.com/network-security-threats/
[2] Wilczek, M.(2019): "DDos:An Underestimated Threat" Retrieved from URL: https://www.darkreading.com/vulnerabilities-threats/ddos-an-underestimated-threat/a/d-id/1336423
[3] Hao, M (2019): "DDos Attacks and Mitigation". Retrieved from URL: https://www.nsfocusglobal.com/ddos-attacks-and-mitigation/
[4] Mehta,R. (2017): Distributed Denial of service Attacks on Cloud Environment, Int. J. Adv. Res. Comput. Sci. 8(5), 2017. pp 2204–2206.
[5] Santhi,B. and Bharathi,G.J.(2012) Study on Distributed Denial-of-Service Attack, Research Journal of Applied Sciences 4(10), 2012. pp 1366- 1370.
[6]. Y. Xie and S.Z. Yu (2009) Monitoring the Application-layer DDoS Attacks for Popular Websites, IEEE/ACM Transactions on Networking 17(1), (2009). pp 15-25.

[7] CloudDDoS (2017): “Four Stages of DDoS Attacks”. Retrieved from URL: https://www.cloudddos.com/about/news/47.html
[8] Rouse, M. (2016): “Blacklist Definition”. Retrieved from URL: www.techtarget.com/definition/blacklist/.
[9] Korolov, M (2016): "DDos Attack on BBC may been biggest in history". Retrieved from URL:http://www.csoonline.com/article/3020292/cyber-attacks-espionage/ddos-attack-on-bbc-may-have-been-biggest-in-history.html
[10] Umarani,S.and Sharmila, D.(2014):"Predicting Application Layer DDoS Attacks Using Machine Learning Algorithms".World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:8, No:10, 2014.pp 1912-1917.
[11] Gupta1,B.B.,Joshi,R.C. and Misra,M.,.(2011):"Predicting Number of Zombies in a DDoS Attack Using ANN Based Scheme". International Journal of Network Security, Vol.13, No.3, Nov.2011, pp 216-225.
[12] Alan Saied, Richard E. Overill, Tomasz Radzik (2016): "Detection of known and unknown DDoS attacks using Artificial Neural Networks". Neurocomputing 172 (2016), pp 385–393.
[13] Anjali. M and Padmavathi,B.(2014): "DDoS Attack Detection based on Chaos Theory and Artificial Neural Network".(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014,
pp 7276-7279.
[14] Doriguzzi-Corin, R., Millar, S., Scott-Hayward, S., Martinez-del-Rincon, J. and Siracusa, D. (2020): "LUCID: A Practical, Lightweight Deep Learning Solution for DDoS Attack Detection". IEEE 2020,https://doi.org/10.1109/TNSM.2020.2971776. pp 1-14
[15] Alkasassbeh, M., Hassanat, A.B.A., Al-Naymat, G., and Almseidin, M. (2016): "Detecting Distributed Denial of Service Attacks Using Data Mining Techniques". (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016. pp 436-445.
[16] Jiangtao, P., Yun, C. and Wei J. (2019):"A DDoS Attack Detection Method Based on Machine Learning".IOP Conferenece Series: Journal of Physics: Conf. Series 1237 (2019) 032040. IOP Publishing doi:10.1088/1742-6596/1237/3/032040. pp 1-5
[17] HoyosL, M.S., Isaza, G.A.E., and. Vélez, J.I. and Castillo, L.O. (2016):"Distributed Denial of Service (DDoS) Attacks detection using Machine Learning Prototype". Article in Advances in Intelligent Systems and Computing · January 2016 DOI: 10.1007/978-3-319-40162-1_4. pp 1-8.
[18]Ben-Gal, I. (2007): “Bayesian Networks”. Encyclopedia of Statistics in Quality and Reliability. John Wiley and Sons, Ltd. Retrieved April 11th 2020 from www.eng.tau.ac.il/bengal/BN.pdf/ pp 1-2.
[19] Cybersecurity IDS Dataset (2020): “Cybersecurity Intrusion Detection System Dataset”. Retrieved 21st April 2020, from URL: http://www.cybersecurity.unsw.adfa.edu.au/ADFA%20IDS%20Datasets/.