Securing Modern Web Services from Distributed Denial of Service using SVM

Abdullah Aljumah

College of Computer Engineering & Sciences Prince Sattam Bin Abdulaziz University, KSA

Abstract
The model network has entered our home through internet and has made our lives very comfortable and this huge world so small by allowing users to have access to any part of the world. The openness and increasing accessibility of the network has already increased the threats for data and the services provided by the network and one such mean is DDoS attack which prevents the legitimate users from accessing the services provided by the server. In this research article we have provided a detection mechanism called vector support mechanism (sometimes known as support vector machines). The main idea behind this SVM is to integrate already discovered attack pattern and train SVM with the help of artificial neural network. After applying various scenarios and hits, we present a highly efficient detection method for DDoS.

Keywords:
DDoS, ANN, SVM, Security, AI.

1. Introduction

Distributed Denial of Services (DDoS) is a deliberate attempt to bring down, make off line or at least degrade performance of any web site or deprive legitimate users from using Data or Services provided by particular web portal is known as Denial of Services Attack when this is launched from more than one place simultaneously, it is called Distributed Denial of Services [1]. In Distributed Denial of Services, a number of compromised devices sometimes millions in number attack a particular website through some controller
in a coordinated manner [2]. Main methodologies incorporated in attacks give them following classes

1. TCP Connection Attacks: This type of attacks is an attempt to consume all possible connection to network infrastructure devices such as firewalls, load-balancers, application servers, web servers etc.
2. Volumetric Attacks: This type of attack introduce congestion or choking in the network to exhaust bandwidth with target network or devices and between target server and Internet.
3. Fragmentation Attacks: This type of attack send a large volume of fragmented TCP and UDP packets to target server and keep it busy in reassembling the incoming stream and eventually sucking all its capability.
4. Application Attacks: This is an attempt to sabotage a particular service of a server by overwhelming it with requests while leaving other undisturbed. This particular behavior makes detection and mitigation difficult for such attacks [3].

An attacker first tries to compromise other machines on the network. After gaining access to many machines, he uses it to launch attack. As any single computer in incapable of launching a DDoS attack powerful enough to sabotage any web services in current cyber infrastructure scenario [4]. A way to achieve powerful attack is to magnify the volume of packets being bombarded to certain routes or servers. There are many ways to achieve this amplification and following are some examples.
DNS amplification: Attacker uses spoofed victim's IP to query DNS server, which in response returns enormous amount of data, when performed in a coordinated manner DNS amplification delivers around 70-fold magnification [5].
Chargen amplification: It is exploitation of an old technique used in printers connected through networks to send an array of random characters to ascertain connection validity, now also used for streaming IP of victim [6].
Distributed Denial of Service (DDoS) has been choice of attack for attackers from the beginning of cyber-attacks and had a tumultuous shadow on enterprises for last many years. DDoS attacks results are readily and widely perceived downtimes faced by important services, and have been ranked among the top cyber security threats by experts for the last several years [7]. According to Kaspersky Lab report about 20% of global enterprises have suffered DDoS attacks [8] . DDoS has been evolving and growing with advent of technology in severity, frequency of attacks by novices with readily available attack tools and in vastness with network bandwidth. According to DoS monitoring agency Arbor, maximum speed of DoS in first half of 2016 reached overwhelmingly 579 Gbps, whereas average speed of such attacks reached to 986Mbps [9]. Motivational factors behind DDoS attacks may vary from business rivalry, political ideology, cyber war among countries to even personal feuds or quest of some novices. Immediate and direct result of DDoS is Unavailability of Target services which causes several other indirect effects viz. Loss of business and Goodwill.

Manuscript received October 5, 2017
Manuscript revised October 20, 2017

Data and services provided by organizations to their customers has been evolved with time in the form of cloud computing. Cloud computing has been adopted by all major IT and IT enabled service providers across the globe [10]. as cloud is used for mission critical data and services, attacks targeting clouds and pattern of attacks has been changed also. Active Threat Level Analysis System (ATLAS) of Arbor has reported a shift in attack mechanism, more and more attacks in 2016 have used DNS as their protocol a shift from 2015 when NTP and SSDP were used most [9].

Economic Denial of Sustainability (EDoS) is another term being used for DdoS attacks on service providers because of their tendency to cause substantial business losses. These losses are direct sequences of web server being taken off line by the attack. Downtime is also indicator of severity of attack as well as service providers' preparedness to cope with such circumstances. However, sign of relieve is that except few, most of the attacks (about 90%) are for less than 1 hour [9].

A large number of research have been under taken for DDoS attacks detection and mitigation. A very comprehensive survey has been done on DDoS detection and Mitigation with traditional methods [13].

Although cost factor remained major motivation for DDoS but there are other reasons also. As mitigation of DDoS attack is a costly resource consuming phenomenon both in terms of man power and computational resources. This special scenario leaves system vulnerable to launch other type of attack e.g. Data breach etc. Modus-operandi of such attackers are first launch DDoS and when all organization expertise is consumed in mitigation of DDoS, simultaneously execute e.g. data breaches such type of attacks are called 'Smoke-screening attacks' [11]. With all attention paid to mitigation of DDoS other crucial resources and services became susceptible to other attacks.

Auto scaling, multi-tenancy and services migration due to mitigation of DDoS in cloud environment produces some other additional effects [12].

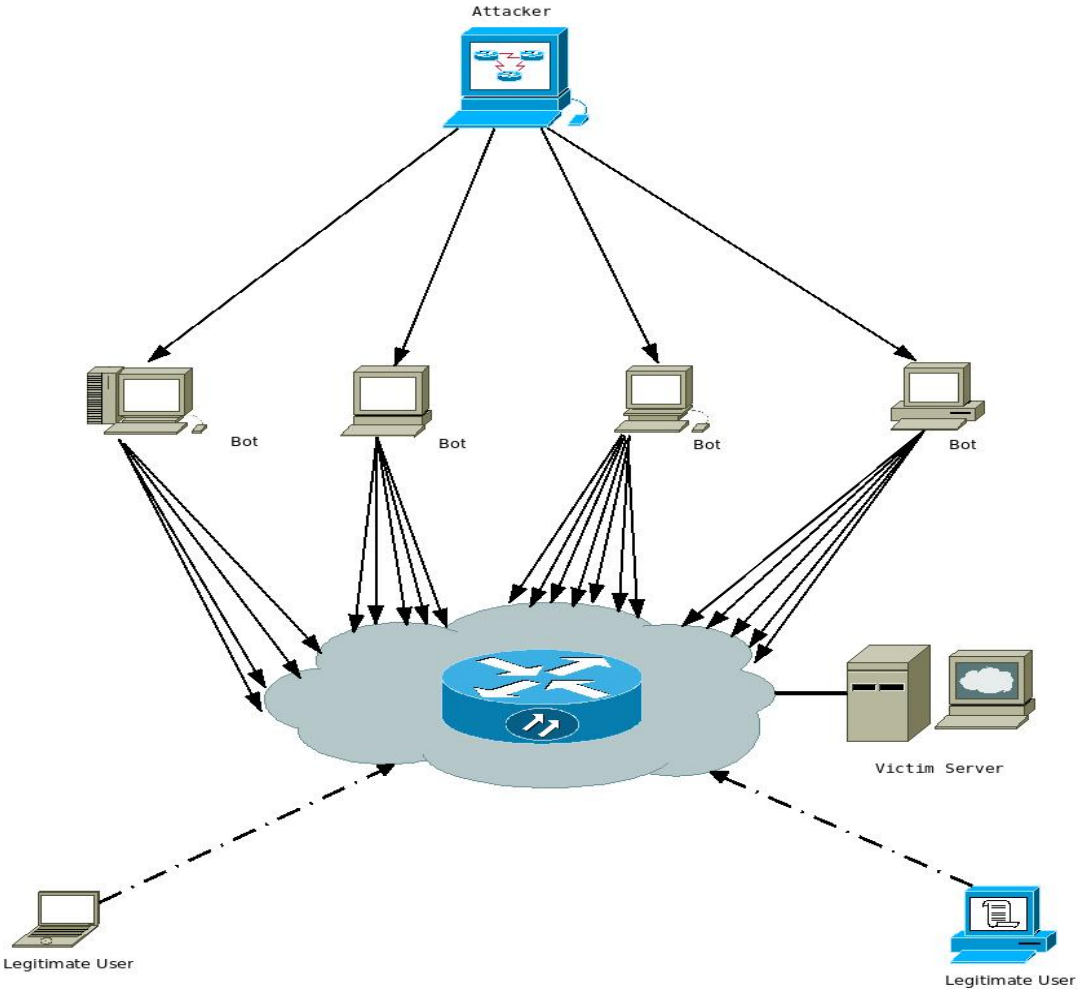

Fig 1: DDoS Attack

2. Related Work

ST Zargharet al. In their survey presented a comprehensive classification on types of DDoS attacks. They have discussed in detail about available techniques and measure available for detection, prevention of DDOS attacks and counter measures used for mitigation of DDoS attacks. [13]

Wang H et al. have found that maximum number of mitigation techniques depend upon distinguishing malign request pattern originating from zombies or bots to that of benign requests from legitimate users. These methods rely upon anomalous patterns or behaviors of attackers requests traffics. On the other hand, some other techniques work differently on proactive prevention mechanisms utilizing cryptographic authentication methods [14].

Wu et al. (2010) have tried to mold game-theory into DDoS defense technique specially where availability of bandwidth has been under attack, here player1 (attacker) tries to overcome firewall safety by optimizing botnet size whereas player2 (defender) strives for optimal rules to separate legitimate traffics from that of rogue one as shown in the figure 1. Strategy presented in work is able to describe the interaction of both players but is largely a too abstract for real life utilization [15].

Wei Zhou et al. have come up with a powerful data mining algorithm to find out Application Layer DDoS (AL-DDoS) attacks, which is able to distinguish between AL-DDoS and Flash Crowds. Experiments and simulation support the viability of proposed algorithms and models incorporated into, to heavy traffics backbone networks. Model proposed is intelligent enough to realize minute maneuvers by attackers in network traffic and trigger the defense mechanism [16].

Bing W. Et al. Tried combination of SDN (Software Defined Network) with cloud computing infrastructures. Experimentally they have showed that in fact it is feasible to utilized this combination for better detection and mitigation of DDoS attack. They have proposed an scheme for detection and mitigation of DDoS attacks viz. DaMask which is an amalgamation of two modules working in cohesion. First one for detection of anomalous behavior of likely attacker (DaMask-D) and second one for mitigation of detected attack based upon graphical interference model (DaMask-M) [19].

Cloud has entered in almost every aspect of computing now a days, Shea R , Liu J have studied effect of DDoS attack in cloud environments with virtualized servers. Their study shows how DDoS attack affects Virtualized servers in cloud environment. They have presented a comparative study of performance degradations under DDoS attack in both virtualized and non-virtualized environment [17].

Zhao S et al. have put forward a method for DDoS detection and defense based on monitoring the utilization thresholds of Virtual Machines. Their mitigation approach is to migrate the affected virtual machine to some isolated servers kept as reserved resources and bringing back them once attack is over [18].

2.1 SVM

Support vector Machines are supervised learning machines with inbuilt learning algorithm that filter and analyze data and recognize data pattern that in-turn is used for regression, analysis and classification. SVM categorize the data by support vector with the help of kernel function, SVM can furnish a general method that fits hyper plane to perform linear categorization of different patterns [20]. SVM can be provided a linear or polynomial function during the training process that selects support vectors. The difference that isolates the key data points decides the number of free parameters used by SVMs but does not include the number of input features [21]. Therefore, proving that SVM need not to reduce the number of key features so as to shun extra fitting – a recognizable benifit in application like intrusion detection and the least probability of generating errors adds credibility to its characteristics.

Support vector machine is supervised machine learning algorithm for classification, and widely used in data mining, network security etc. SVMs performs binary classification by predicting the optimal hyperplane with maximum margin, data points in the input space are classified as positive side and negative side of the hyperplane [22]. Predicting the hyperplane is done only if the data points in feature input space are associated with a property which indicates them as belongs to class A or class B in other words to say they are linearly separable.

For the given the data set $D = \{(xi\ yi)x_i \in R^P, yi \in \{-1,1\}\}_{i=1}^{N}$

Samples for training are $(x_1\ y_2), (x_i y_i), \ldots\ldots, (x_N y_N)$.

Above $x_i \in R^P$ where P is the dimension of the sample data and $y_i \in \{-1, 1\}$and $i \in \{1, N\}$

First find the optimum hyperplane with no errors using the equation $W^T x + b = 0$. Once the hyperplane is created in the feature space, unknown data items can be classified as members of class A or Class B as follows for any given element.

Class A $W^T x + b > 0$ and class B $W^T x + b < 0$

SVM first trains the system using training data set to establish several support vectors in the feature space, these support vectors will form the SVM model, these support vectors are the data points on the boundary of optimum hyperplane with maximum margin as illustrated in the figure-1B.

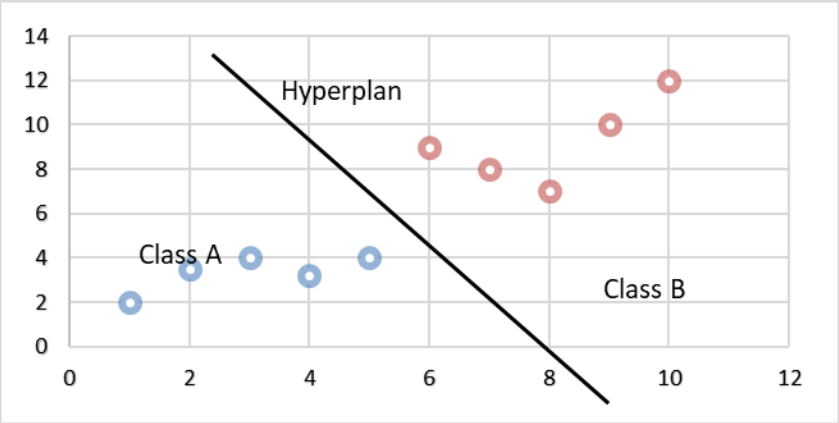

Fig 1B Linear dataset with hyperplane

When the input space has nonlinear classes as shown in figure-1C then classification is achieved by mapping the data to a higher dimensional space where linear patterns are possible, then linear classification model is applied in the new input space. Kernels method enables to map the data to higher dimension space where it is possible to use linear models of classification to separate the two classes in nonlinear feature space.

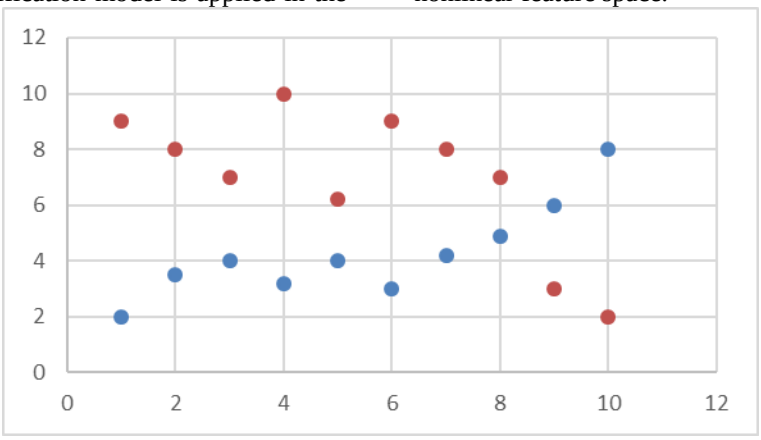

Fig 1C Nonlinear data set

Different kernel functions were proposed for various application domains, most commonly used kernel functions are linear function, polynomial function, sigmoid function, and radial basis function. Above mentioned kernels do not treat the features of data as different. SVM kernel function $K(x_i\ ,\ x)$ is added with weights to measure the importance of features in the given space . based on this idea a new kernel function is formulated as $K(wx_i\ ,\ wx)$. where w is a vector consisting of weights of features of given data set. After adding the weights, a nonlinear discriminant function with feature weights is given below to classify an element x.

$$f(x) sign(w.x + b) = sign(\ \sum_{i=1}^{N} \propto_i y_i K(wx_i, wx) + b)$$

Where $f(x)$ is the decision function such that $y_i = f(x_i)$ and $y_i \in \{-1, 1\}$and $i \in \{1, N\}$.

Kernel function K find the similarities between trained data item wx_i and given unlabeled data item x . Kernel function $K(wx_i, wx)$main types are as given below

Linear kernel:$K(x_i, x_j) = x_i^T x_j$;

Polynomial Kernel:$K(x_i x_j) = (\gamma\ x_i^T x_j + r)^P\ \ , \gamma > 0$;

Radial Basis Kernel function: $K(x_i x_j) = exp(-\gamma||x_i - x_j||^2)\ \ , \gamma > 0$;

2.2 DDoS detection using support vector network

The extraction of input vectors from raw TCP/IP shack in the data and preprocessing and the outcome value is single and indicates the current pattern is a DDoS attack or not. DDoS detection SVM (N) is comprised of three steps:

Preprocessing: by applying an automatic parser, raw TCP/IP shack data is turned in to an appropriate form.

Training: various kinds of attacks along with genuine traffic is flooded to train SVM. We have two classes with 37 features, one is DDoS attack data and the other one is normal.

Testing: the performance tested SVM ensures that it has obtained satisfactory classification capability.

Table 1: Comparison Table

Experiment	Accuracy %	Training time in seconds	Testing time in seconds	Train/test datasets
DDoS/Normal	99.73	24.08	15.31	11592/81865
DDoS/Rest	99.27	22.78	1.93	5091/6891
Smurf/Rest	100	4.78	2.54	5091/6891
Neptune/ Rest	99.97	21.21	0.91	5091/6891
Back/ Rest	99.77	9.01	2.89	5091/6891
Land/ Rest	99.97	0.83	0.17	5091/6891
PoD/ Rest	99.98	3.28	1.67	5091/6891
Tear Drop/ Rest	99.69	16.17	0.08	5091/6891

Selecting Physiognomy in intrusion detection is an important issue and in a broader sense the IDS physiognomy can be monitored for detection; which are useless, less significant or truly useful respectively? This is a vital question as by eliminating the useless physiognomy the detection accuracy can be improved and the computation process can be faster. Thus the overall performance of IDS can be improved. Now in a scenario where there is no useless physiognomy we can still improve the performance of IDS by concentrating on the most vital ones without disturbing the detection accuracy in important ways statistically.

Selecting and prioritizing the feature for intrusion detection is a problem that is parallel in behavior to many engineering problem and following listed are the types.

1- Possessing huge amount of inputs i.e D (D_1, D_2, …D_n) of different importance values (weights) i.e. some values of D are important, not so important and purposeless respectively.

2- Non-existence or inefficient mathematical formula or analytical model that implicitly narrates and relates input and output relation A=B (D)

3- Having a confined and limited set of evaluation and test data that is used to build a model for modeling, evaluation and prediction.

Since we lack a systematic prototype or method, we can use empirical methods to intuit the comparative significance of the input variables. And examination of each possibility is needed for a complete analysis e.g. analyzing the correlation and dependence of two variables at a time and increasing the number of variables e.g. taking three values a time etc. but this is not feasible because it requires 2n tests and since the quality of existing data may not be good in evaluating the input space thus this is not infallible too. Thus, we use the method of erasing single feature at a specific unit of time to arrange the input features as per the rank as shown in table 2 and select the best and vital features for DDoS attack detection as shown in fig 3.

Table 2. Detection Accuracy for DDoS using Performance ranking method

	Features	Accuracy %	Train Time	Test Time
All	41	99.24	22.86	1.91
Imp.	18	99.23	22.8	1.83
Sec. and Unimp.	33	99.24	19.7	2.09

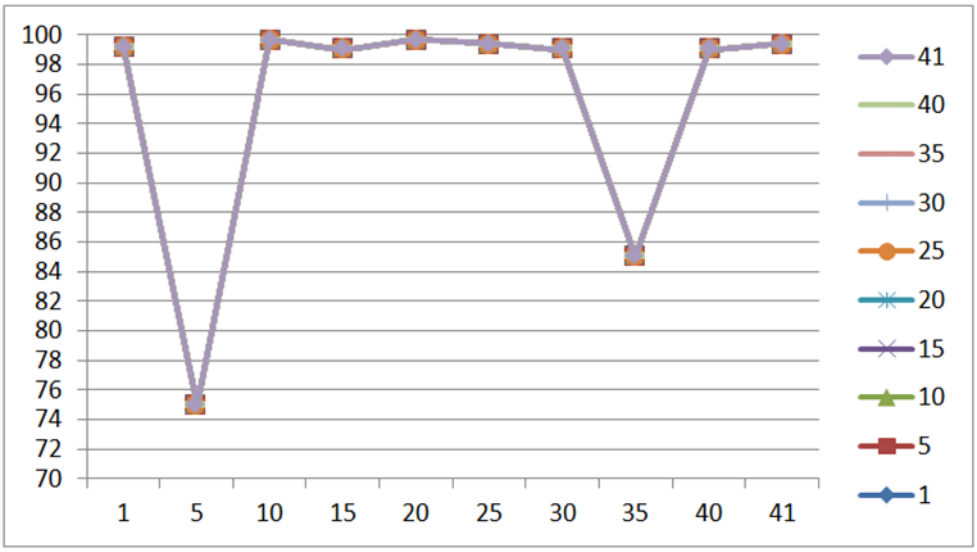

Fig 3: DDoS Detection Accuracy Obtained By Deleting One Characteristic at a time

2.3 Ranking Using Performance

We start with a Model independent approach where our ranking technique rely on performance based input that is to say we gradually minimize features from input data one at a time and the residue set is used to train and subsequently test of our classifier. Accuracy of classifier is compared to that of classifier with all features on i.e. Original classifier in terms of relevance performance identifiers. This iterative method ends with a Ranked Feature List as shown in figure 2 according to performance comparison set of rules. Following is the procedure:

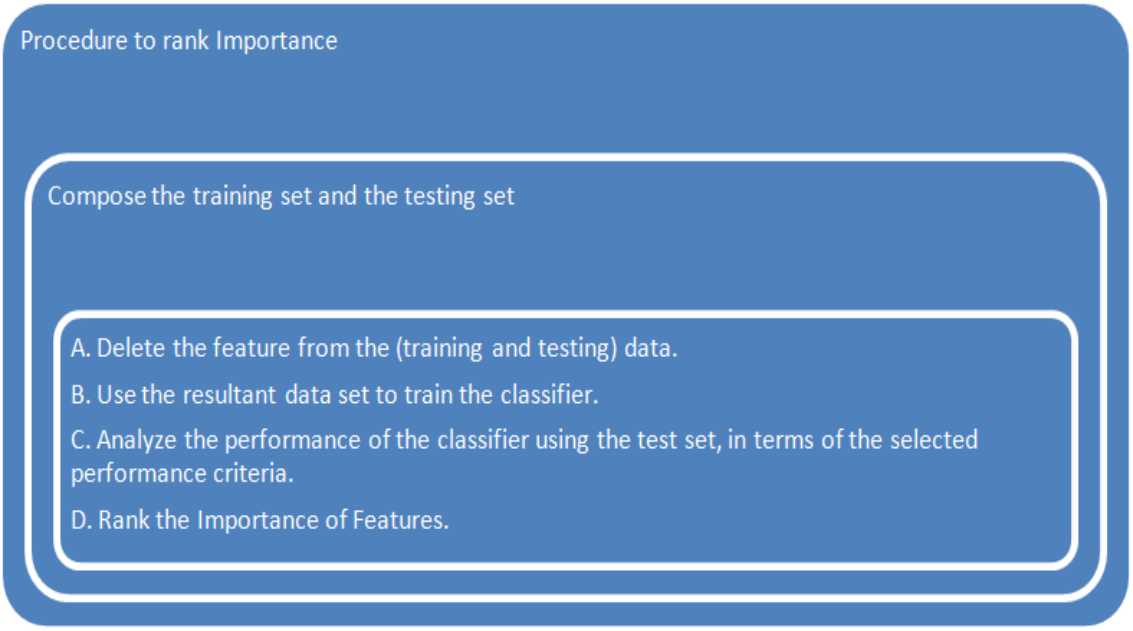

Fig 2: Procedure for Ranking Using Performance

3. Selective Ranking Method

The SV decision function holds the hidden details of the features and their role in
categorization and with these details we can prioritize their importance as shown in the following equation

$$Y(X) = \sum ZiXi + C$$

Class of X is determined by Y (X), positive for the +ve value of Y (X), -ve for the negative value of Y (X) respectively. Y (X) is grossly dependent upon the individual values of X and Zi.|Zj| discrete value of Zj determines classification's robustness. Ith feature becomes a key factor for positive class if Zi comes out to be a large positive number whereas it becomes base for -ve class for large negative values of Zi. Furthermore, it becomes insignificant to the classifier if Zi tends to or closes to zero on either side of scale. This forms the basis of ranking performed by decision function of support vector mechanism.

Ordering input is executed by training the classifier with original dataset and the importance of features is ranked using the decision function of classifier. Following procedure is used to achieve this

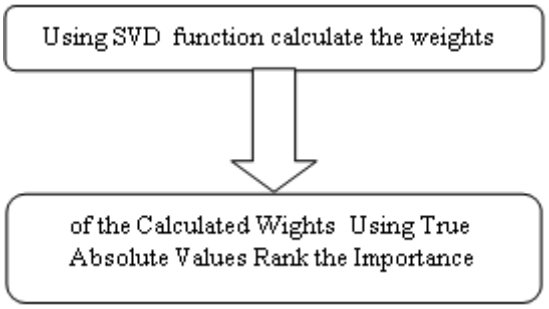

Fig 4: Procedure for selective ranking

Using this procedure, the features are categorized into three sections as shown in the table 3

Table 3: Categorized Features

Important	1,4,5,17,22,23,25,27,29,31,35,37,40
Secondary	2,3,6,9,12,24,26,33
important	7,8,10,11,13,14,15,16,18,19,20,21,28,30,32,34,36,38,39,41

Table 4. Detection Accuracy For DDoS using selective ranking method

	No. of Features	Accuracy %	Training Time	Testing Time
All	41	99.24	22.78	1.91
Imp.	12	99.17	18.91	1.01
Sec. and Unimp.	331	99.65	73.49	1.49

As shown in the table 4 these two ranking methods provided the consistent output for selecting the important features. The most vital features as certified by the two methods are given in the table 5.

Table5: Certified features obtained from ranking methods

S.No.	**Feature**	**Description**
1	Duration	*Connection length between source and destination hosts*
2	Source bytes	*No. of bytes of outbound traffic from source host to destination.*
3	Destination bytes	*No. of bytes of inbound traffic to source from destination host.*
4	Count	*Connection count to the specific host in given time interval*
5	Same service rate	*Proportion of connections with same service to the host in a specific period of time.*
6	Connection with syn errors	*Proportion of syn error connections in a given time interval.*
7	Connection same service syn errors	*Proportion of connection with same service and syn errors in it in a specific time interval.*
8	Destination-host count	*Inbound connection count by the same destination host in a time interval*
9	Destination-host name source port rate	*Proportion of the port connections between source and destination host in in-bound traffic in time frame*
10	Destination-host syn error rate	*Proportion of syn error connections between source and destination host in the time interval.*
11	Destination-host same service error rate	*Count of errors in same service connection in a time interval*

4. Conclusion

In this research article we used DARPA intrusion evaluation dataset to implement SVM for detection of DDoS attack patterns and validation of their performance. We also ordered and ranked the DDoS detection relevant features using heuristic methods. The performance of SVM is truly good in an IDS application especially in DDoS detection and SVM performs better than other existing machine learning techniques like ANN in vital aspects of detection accuracy, scalability and training time. More importantly the detection accuracy of SVM is markedly better than ANN while the training time of ANN is markedly higher than SVM. The evaluation results proved that the SVMs can achieve the detection results more than 99.00% for every instance of attack. However, the evaluation results also showed that the detection accuracy decreases slightly while using selected and vital features only and this implies that we should include sensitivity selector in an IDS. Depending on the system, series and the security essentials, the distinctive collection of features can be implemented for DDoS detection engine.

Acknowledgement

This research was funded and conducted at Prince Sattam bin Abdulaziz University, Alkharj, Saudi Arabia during the academic year 2017 under research number 2017/01/7091.

References:

[1] Tariq Ahamad, Abdullah Aljumah," Detection and Defense Mechanism against DDoS in MANET", Indian Journal of Science and Technology, Vol 8, No 33, December 2015.

[2] Tariq Ahamad, “Detection and Defense Against Packet Drop Attack in MANET” International Journal of Advanced Computer Science and Applications(IJACSA), 7(2), June 2016.

[3] A. S. Pimpalkar and A. R. B. Patil, "Detection and defense mechanisms against DDoS attacks: A review," 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, 2015, pp. 1-6.

[4] Abdulaziz Aldaej and Tariq Ahamad, “AAODV (Aggrandized Ad Hoc on Demand Vector): A Detection and Prevention Technique for Manets” International Journal of Advanced Computer Science and Applications(IJACSA), 7(10), October 2016.

[5] H. Yu, X. Dai, T. Baxliey, X. Yuan and T. Bassett, "A visualization analysis tool for DNS amplification attack," 2010 3rd International Conference on Biomedical Engineering and Informatics, Yantai, 2010, pp. 2834-2838.

[6] Z. Han, X. Wang, F. Wang and Y. Wang, "Collaborative detection of DDoS attacks based on chord protocol," 2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012), Las Vegas, NV, 2012, pp. 1-4.

[7] Abdullah Aljumah, Tariq Ahamad,"A Novel Approach for Detecting DDoS using Artificial Neural Networks". International Journal of Computer Science and Network Security, 132 VOL.16 No.12, December 2016.

[8] Kaspersky DDOS intelligence report for Q3 2016 https://securelist.com/analysis/quarterly-malware-reports/76464/kaspersky-ddos-intelligence-report-for-q3-2016/

[9] Abdullah Aljumah, “Detection of Distributed Denial of Service Attacks Using Artificial Neural Networks” International Journal of Advanced Computer Science and Applications(IJACSA), 8(8), 2017. http://dx.doi.org/10.14569/IJACSA.2017.080841.

[10] Burt C. Majority of organizations run mission-critical apps in the cloud: verizon report. URL: http://www.thewhir.com/web- hosting- news/ majority- of-organizations/- run- mission- critical- apps- in- the- cloud-verizon- report , 2015

[11] Abdullah Aljumah, Tariq Ahamad,"A Novel Approach for Detecting DDoS using Artificial Neural Networks". International Journal of Computer Science and Network Security, 132 VOL.16 No.12, December 2016.

[12] Abdullah Aljumah, Tariq Ahamad,"Futuristic Method to Detect and Prevent Blackhole Attack in Wireless Sensor Networks", International Journal of Computer Science and Network Security,VOL.17 No.2, February 2017

[13] S T Zargar, J Joshi, D Tipper, A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks, IEEE Communications Surveys & Tutorials (Volume: 15, Issue: 4, Fourth Quarter 2013)

[14] Wang H , Jia Q , Fleck D , Powell W , Li F , Stavrou A . A moving target DDos defense mechanism.ComputCommun 2014;46:10–21 .

[15] Wu Q, Shiva S, Roy S, Ellis C, Datla V. On modeling and simulation of game theory-based defence mechanisms against DoS and DDoS attacks. In: Proceedings of the 2010 spring simulation multiconference, SpringSim ’10. San Diego, CA, USA: Society for Computer Simulation International; 2010159:1e159:8.

[16] Wei Zhou , WeijiaJia, Sheng Wen, Yang Xiang, Wanlei Zhou, Detection and defense of application-layer DDoS attacks in backbone web traffic, Science Direct http://dx.doi.org/10.1016/j.future.2013.08.002

[17] Shea R , Liu J . Performance of virtual machines under networked denial of service attacks: experiments and analysis. Syst J IEEE 2013;7(2):335–45 .

[18] Zhao S , Chen K , Zheng W . Defend against denial of service attack with VMM. In: Grid and cooperative computing, 2009. GCC’09.Eighth international conference on.IEEE; 2009. p. 91–6 .

[19] Bing Wang, Yao Zheng, Wenjing Lou, Y. Thomas Hou, DDoS attack protection in the era of cloud computing and Software-Defined Networking, Science Direct, http://dx.doi.org/10.1016/j.comnet.2015.02.026

[20] Jun Liu, Xiaoming Liu, Bo Zheng and J. Tang, "Design and implementation of code security inspection system based on SVN," 2011 International Conference on Computer Science and Service System (CSSS), Nanjing, 2011, pp. 330-333.

[21] V. Filippov, "Dressing Subversion: ViewVC and SVN-Searcher," 2009 5th Central and Eastern European Software Engineering Conference in Russia (CEE-SECR), Moscow, 2009, pp. 102-107.
[22] M. Schwind and C. Wegmann, "SVNNAT: Measuring Collaboration in Software Development Networks," 2008 10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services, Washington, DC, 2008, pp. 97-104.