International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET207459

Detection of DDoS Attack in TCP protocol using Hybrid Machine Learning Techniques

Prof. Vinod Desai[1], Aravind Pradhani[2], Sheetal Majukar[3]

[1]Assistant Professor, Department of Computer Science and Engineering, Angadi Institute of Technology and Management Belagavi, Department of Computer Science and Engineering, Savagaon, Karnataka, India

[2,3]Student, Angadi Institute of Technology and Management Belagavi, Dept. of Computer Science and Engineering, Savagaon, Karnataka, India

Article Info
Volume 7 Issue 4
Page Number: 253-258
Publication Issue :
July-August-2020

Article History
Accepted : 05 Aug 2020
Published : 12 Aug 2020

ABSTRACT

Recently, damage caused by DDoS attacks increases year by year. Along with the advancement of communication technology, this kind of attack also evolves and it has become more complicated and hard to detect using flash crowd agent, slow rate attack and also amplification attack that exploits a vulnerability in DNS server. Fast detection of the DDoS attack, quick response mechanisms and proper mitigation are a must for an organization. An investigation has been performed on DDoS attack and it analyzes the details of its phase using machine learning technique to classify the network status. In this paper, we propose a hybrid KNN-SVM method on classifying, detecting and predicting the DDoS attack. The simulation result showed that each phase of the attack scenario is partitioned well and we can detect precursors of DDoS attack as well as the attack itself.

Keywords: Distributed denial of services (DDoS), Machine learning classifiers, Security, Intrusion detection, Prediction, support vector machine (SVM), k-nearest neighbor (KNN), KNN-SVM

I. INTRODUCTION

Three aspects usually involve in computer related issues such as integrity, confidentiality and availability. Security threats fall into three categories such as breach of confidentiality, failure of authenticity and unauthorized denial of services [1]. Distributed Denial of Services (DDoS) become the major problem and it gives the latest threat to the users, organizations and infrastructures of the internet. This type of intrusion (DDoS) attacker attempts to disrupt a target, by flooding it with illegitimate packets, exhausting its resource and overtaking it to prevent legitimate inquiries from getting through. According to the security report of Arbor 2005-2010. This paper analyzes current research challenges in DDoS by evaluating machine learning algorithms for detecting and predicting DDoS attack, which includes feature extraction, classification, and clustering. Besides, various hybrid approaches have been employed. It is illustrated that these evaluation results of research

challenges are mainly suitable for machine learning technique.

This paper is organized as follows. Section 2 provides a related study on an overview of machine learning techniques and briefly describes a number of related techniques for intrusion detection. Section 3 compares related work based on the types of classifier design, the chosen baselines, datasets used for experiments, etc. Conclusion and discussion for future research are given in Section 4.

II. RELATED WORKS

Lately, there are many reports that show the involvement of DDoS attack on commercial or government website [4]. Along with the advancement technique of DDoS attack, the studies on detection also evolve and as a result, various methods have been suggested to counter DDoS attack. As we know, DDoS attack can be classified into anomaly-based, congestion-based and others [3].

A network traffic controller using machine learning (ML) techniques was proposed in 1990, aiming to maximize call completion in a circuit-switched telecommunications network [1]. This was one of the works that marked the point at which ML techniques expanded their application space into the telecommunications networking field. In 1994, ML was first utilized for Internet flow classification in the context of intrusion detection. It is the starting point for much of the work using ML techniques in Internet traffic classification that follows.

Gavrilis et. al [4] utilized RBF-NN detector which is a two-layer neural network. It uses nine packet parameters and the frequencies of these parameters are estimated. Based on the frequencies, RBF-NN classifies traffic into attack or normal class. In this study, the IP spoofing characteristic which is one of the most definite DDoS attack evidences is not considered for a correct attack detection. Regarding UDP type attacks, the detection efficiency is lower than that of TCP type attacks and is apparently low in the beginning period of attacks. Defining k-means center which minimizes the quantization error is also a difficult task.

The hybrid technique proposed by Ming-Yang Su et.al [5] is a method to weigh features of DDoS attacks and it analyzed the relationship between detection performance and number of features. The study proposed a genetic algorithm combined with KNN (k-nearest-neighbor) for feature selection and weighting. All initial 35 features in the training phase were weighted, and the top ones were selected to implement Network Intrusion Detection System (NIDS) for testing. A fast mechanism to detect DDoS attack is by extracting features from the network traffic, so that all these features come from the headers, including IP, TCP, UDP, ICMP, ARP and IGMP.

According to the framework of Genetic Algorithm (GA), the proposed NIDS is described by three parts in the section. The first subsection will present all features that are considered in the study; the second subsection will state the encoding of a chromosome and the fitness function; the third subsection will provide details on the selection, crossover, and mutation in the GA. There is also an evaluation on machine learning technique on DDoS attack, proposed by Suresh [10] which indicates that Fuzzy c-means clustering gives better classification and it is fast compared to the other algorithms.

III. PROPOSED WORK

In order to handle this DDoS attack in TCP protocol, we have proposed a combination of two machine learning based model with Support Vector Machine (SVM) and *k*-Nearest Neighbors (KNN). In this section, we discuss the details of methods that have been utilized in this work for detection and prediction of DDoS attack. There are k-nearest neighbor (KNN) and support vector machine (SVM) or known as KNN-SVM.

1. **Support Vector Machine (SVM):**

In classification and regression, Support Vector Machines are the most common and popular method for machine learning tasks [16]. In this method, a set of training examples is given with which each example is marked belonging to one of the two categories. Then, by using the SVM algorithm, a model that can predict whether a new example falls into one category or the other is built.

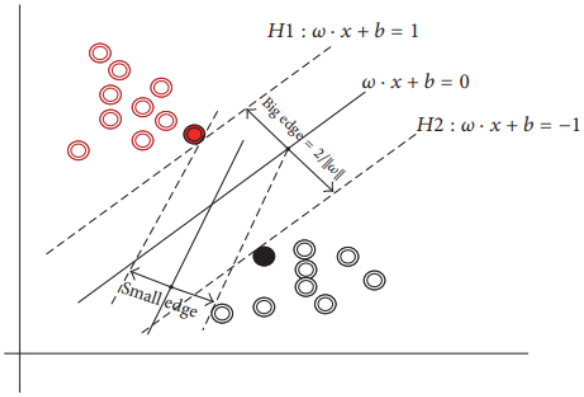

Fig 1.1 : SVM Training Examples

2. **K-Nearest Neighbor (KNN)**

A k-NN algorithm has been shown to be very effective for a variety of problem domains including text categorizing [17]. It determines the class label of a test example based on its k neighbor that is close to it. The similarity score of each neighbor document to test document is used as the weight of categories of the neighbor document. Referring to fig. 1, it has been effectively used to calculate the distance among neighbors.

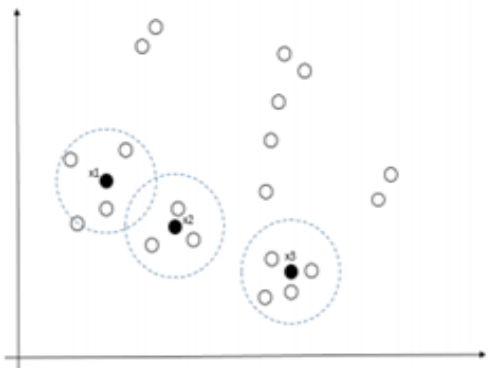

Fig 1.2 : KNN Training Examples

IV. EXPERIMENTAL SETUP

1. SVM Algorithm

Input: Captured network traffic with both attack and normal connections.
Output: An Optimal SVM Classifier.

Train the SVM classifier with training samples, when the controller starts.
Categorize the different types of packet as TCP, ICMP, UDP using traffic classifier module.

Construct an optimal hyper plane to classify the attack.
If the classifier detects the connection as attack.
Block the connection and update the rules in flow table.
Else
Allow the connection Wait for next connection
Repeat above two steps for all the incoming connection.

2. KNN Algorithm

Input: Captured Network traffic with both attack and normal connections.
Output: An Optimal KNN Classifier.

Train the KNN classifier with training samples, when the controller starts.

Categorize the different types of packet as TCP, UDMP, ICMP using traffic classifier module

.

1. Begin
2. if t ∈ T then
3. Calculate D
4. if D >0.5
5. Network is attacked by DDoS;

6. else
7. Network is normal;
8. end if
9. end if
10. End

To analyze the performance of proposed system, we have implemented SVM, KNN along with our proposed hybrid algorithm using SVM and KNN stated in respectively to detect DDoS attack. For our simulation, we have created a custom network topology with six switches and 21 host systems as shown in using mini net, an emulation tool, which helps to create a various types of DDoS attacks are studied to select the traffic parameters that change unusually during such attacks. There are eight features extracted from both datasets using information gain rank. Then we rank all the features to identify which one is more relevant. Many machine learning problems can actually enhance their accuracy by applying features selection and extraction. This situation intensively indicates that feature selection is also important for ranking [10]. Information gain is applied to measure the importance of each feature. The information gain of a given attribute X with respect to the class Y is the reduction in uncertainty about the value of Y, after observing values of X. The uncertainty about the value of Y is measured by its entropy defined as

$$H(Y) = -\sum_i P(Y_i) \log_2((P(Y_i))$$

where P(Yi) is the prior probabilities for all values of Y

where P(Yi) is the prior probabilities for all values of Y. The uncertainty about the value of Y after observing values of X is given by the conditional entropy of Y given X defined as The performance of the attack detection is displayed by the detection rate (DR) and false alarm rate (FAR); the formulas are calculated as the values:

DR = DD/(DD + DN)

In this formula, DD indicates that the attack flow is detected as an attack flow, and DN means that the attack flow is detected as a normal flow.

FAR = FD/(FD + TN)

In the formula, FD means that the normal flow is detected as an attack flow, and TN indicates that the normal flow is detected as a normal flow.

V. PERFORMANCE EVALUATION

The performance of proposed hybrid algorithm is analyzed using the parameters such as accuracy, detection rate and false alarm rate calculated using Equations 1, 2& 3respectively . The test shows the comparison of detection rate, accuracy, false alarm rate for SVM, KNN, hybrid SVM-KNN. Since SVM is a supervised machine learning model, it should be trained with labeled data only. And it is very complex to detect the new attack. In case of KNN, it is a unsupervised machine learning model, it can able to detect the new attacks. But in case of KNN, false alarm rate will be high. In order to avoid the drawbacks in SVM and KNN, we have used hybrid model.

Table 1 : Different Methods Used

Method Used	Correct Classification %	Detection Time (In Second)
SVM	96.4	0.23
KNN	96.6	0.26
Decision Tree	95.6	0.25
K-Mean	96.7	0.20
Naive Bayesian	92.9	0.52
Fuzzy C Mean	98.7	0.15

In our model, first the traffic is passed through the SVM module and attacks are identified. To detect the new kind of attacks, resultant traffic from SVM module is again passed through the KNN module. Once the attack is detected, particular connection will be closed and rules are updated in flow table. Our proposed hybrid model provides high accuracy of 96.77%, high detection rate of 90.45%, and low false alarm rate of 0.032

Table 2 : Comparison of Different Algorithms

Algorithm	TP (%)	TN (%)	FP (%)	FN (%)
SVM	82.03	87.17	5.67	4.76
kNN	84.33	88.45	6.97	5.55
SVM-kNN	85.49	89.24	2.51	0.83

Table 3 : Experimental Results

Method	TP	FP	TN	FN	F-Measure
SVM	281	18	253	20	0.96
KNN	280	20	243	30	0.97
Decision Tree	277	22	218	55	0.96
K-Mean	285	15	273	0	0.97
Naive Bayesian	292	10	256	17	0.97
Fuzzy C Mean	298	2	270	3	0.99

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a hybrid machine learning model to detect DDoS attack in SDN environment in TCP protocol. We also analyzed our proposed work based on the three performance metrics such as accuracy, detection rate, and false alarm rate. KNN, an unsupervised machine learning algorithm, which works well for detection of attacks compared to SVM algorithm in TCP protocol. But by using our proposed hybrid machine learning model (SVM- KNN), we have achieved more accuracy, detection rate and low false alarm rate compared to simple machine learning model. In future, we will try to implement ensemble machine learning models to detect DDoS attack in data plane by imposing the security rules in the flow table.

VII.REFERENCES

[1]. Kreutz, Diego, et al. "Software-defined networking: A comprehensive survey." Proceedings of the IEEE 103.1 (2015): 14-76.

[2]. Zargar, SamanTaghavi, James Joshi, and David Tipper. "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks." IEEE communications surveys & tutorials 15.4 (2013):2046-2069.

[3]. Yavuz CANBAY and Seref SAGIROGLU, "A Hybrid Method for Intrusion Detection" In IEEE 14th International Conference on Machine Learning andApplications",2015.

[4]. A.Saboor and B.Aslam, " Analyses of Flow Based Techniques to Detect Distributed Denial of Service Attacks" In Proceedings of 12th International Bhurban Conference on Applied Sciences & Technology (IBCAST), 13th-17th Jan,2015. pp354-362.

[5]. Saurav Nanda, Faheem Zafari, CasimerDeCusatis, Eric Wedaa and Baijian Yang, "Predicting Network Attack Patterns in SDN using Machine Learning Approach", In IEEE Conference on Network Virtualization and Software Defined Networks (NFV- SDN),2016.

[6]. Gisung Kim, Seungmin Lee, Sehun Kim "A novel hybrid attack detection method integrating anomaly detection with misuse detection", - journal on Expert Systems with Applications – [Online].

[7]. Niyaz, Quamar, Weiqing Sun, and Ahmad Y. Javaid. "A deep learning based DDoS detection system in software-defined networking (SDN)." arXiv preprint arXiv:1611.07400(2016).

[8]. Barki, Lohit, et al. "Detection of distributed denial of service attacks in software defined networks." Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on. IEEE,2016.

[9]. The most popular types of DNS attacks. [Online]. Available: https://securitytrails.com/blog/most-popular-types-dns-attacks (visited on 01/06/2020).

[10]. D. Smith, Portmapper is preying on misconfigured servers to amplify attacks, Sep. 2015. [Online]. Available: https://blog.radware.com/security/2015/09/port mapper-preying-on-serve rs/ (visited on 01/25/2020).

[11]. Akamai, Attackers using new MS SQL reflection techniques, Feb. 2015. [Online]. Available: https://blogs.akamai.com/2015/02/plxsert-warns- of- ms- sql- reflection- attacks.html (visited on 01/08/2020).

[12]. J. M. Alonso, R. Bordon, M. Beltran, and A. Guzman, "LDAP injection techniques," in IEEE Singapore International Conference on Communication Systems, Guangzhou, China, Nov.2008, pp. 980–986.

[13]. Microsoft, MS03-034: Flaw in NetBIOS could lead to information disclosure, Sep. 2019. [On-line]. Available: https://support.microsoft.com/en-us/help/824105/ms03-034-flaw-i n-netbios-could-lead-to-information-disclosure (visited on 01/16/2020).Cloudflare, NTP amplification DDoS attack. [Online]. Available: https://www.cloudflare

Cite this article as :

Prof. Vinod Desai, Aravind Pradhani, Sheetal Majukar, " Detection of DDoS Attack in TCP protocol using Hybrid Machine Learning Techniques, International Journal of Scientific Research in Science, Engineering and Technology(IJSRSET), Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 7, Issue 4, pp.253-258, July-August-2020. Available at doi : https://doi.org/10.32628/IJSRSET207459
Journal URL : http://ijsrset.com/IJSRSET207459