HNSJ Humanities & Natural Sciences Journal
ISSN: (e) 2709-0833
www.hnjournal.net

Peer-Reviewed Journal
Indexed at: NSP - SJIF
Impact Factor 2022 = 4.91

RESEARCH TITLE

Predicting of DDoS Attack on DNS Server using Logistic regression Algorithm

Abdusalam Yahya[1], AHMED MOHAMMED OMAR[2]

[1] Elmergib University, Faculty of IT , Libya. Email: abdusalamabduallahyahya74@gmail.com
[2] Elmergib University,Faculty of Engineering . Libya. Email: amaesheebah@elmergib.edu.ly

HNSJ, 2025, 6(2); https://doi.org/10.53796/hnsj62/31

Received at 07/01/2025 **Accepted at 15/01/2025** **Published at 01/02/2025**

Abstract

The internet heavily relies on the Domain Name System (DNS) to perform essential functions for users worldwide. However, Distributed Denial of Service (DDoS) attacks on DNS servers pose significant challenges to this functionality. In this paper, we introduce a method for detecting DDoS attacks on DNS servers using a logistic regression algorithm. We selected response time and packet length as key features for predicting attacks. Our dataset was generated using VMware and Wireshark tools. The Python programming language was employed to implement and evaluate the model. The results indicate that the proposed model effectively and accurately detects DDoS attacks on DNS servers.

Key Words: DDOS Attack, Domain Name Service, Logistic regression algorithm

Introduction

In recent years, Distributed Denial of Service (DDoS) attacks have been rising rapidly, creating a serious threat to the availability and integrity of online services.[1].A DDoS attack refers to a situation in which a legitimate service, system, or network becomes unavailable or inaccessible to its intended users[2]. It can cause the interruption of regular services in a number of ways. Their two main goals are to utilize the available network bandwidth and overwork the targeted server or infrastructure to the point where it become unavailable to legitimate users [3].

The Domain Name System (DNS) is an application layer protocol that plays a crucial role in the functionality of the internet by enabling the bi-directional translation of domain names and IP addresses, facilitating seamless communication online [4]. Historically, DNS was operated by the Internet Assigned Numbers Authority (IANA), and its operational functionality was transitioned to the Internet Corporation for Assigned Names and Numbers (ICANN) in 1998, which is a nonprofit organization under California law [5].

DDoS attacks on DNS servers are categorized into two types [6]. One type is a flooding DDoS attack, which is conducted by sending a large number of DNS requests. This attack aims to exhaust the resources of the DNS server, making it inaccessible to legitimate users. A normal DNS server is unable to differentiate between spoofed traffic and legitimate user traffic; therefore, it will accept all incoming traffic and respond to each request. As a result, the DNS server may become overloaded and start dropping requests overall. Figure 1 illustrates a flooding attack against a DNS server. The attacker generates multiple spoofed DNS requests to disrupt the standard DNS function and overwhelm its resources, primarily memory and CPU [7].

DDoS Attack Classification

A DDoS attack utilizes server machines to initiate a coordinated DoS attack on one or more victims. All DDoS attacks can be categorized into two main classes: bandwidth-consuming attacks and resource exhaustion attacks. The aim of a bandwidth attack is to overload the target network or host with unwanted traffic, limiting the flow of legitimate traffic. Flooding attacks and amplification attacks are two types of this category [8]. In a resource depletion attack, the attacker focuses on...

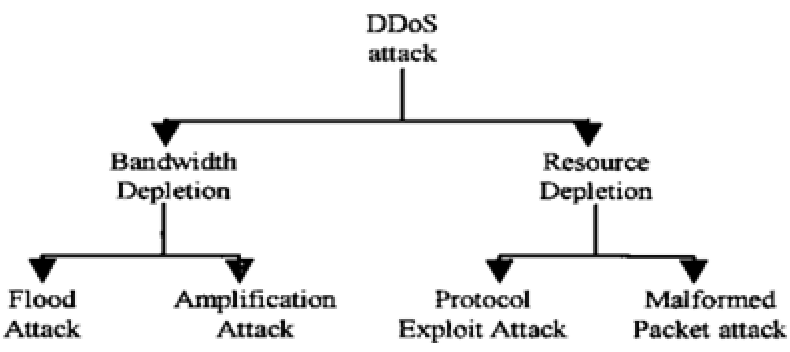


Figure1: DDoS Attack Classification

One kind of (D)DoS attack that typically attacks recursive DNS servers is a DNS flood; it was first researched several years ago [8]. A collection of hacked client devices involves in DNS flooding, which involves sending a high amount of legitimate DNS queries to a DNS server until all of its resources—memory, CPU, or bandwidth—are exhausted. Two varieties of DNS flooding exist [9].

Logistic regration

is a type of supervised mahine learning , and it cosists of two main steps :training where the mode is traind using labled data, and classfication where the model is used to predict if network traffic is normal or not [10]. Due to it is simplesty ,easy to implement and it is low computational requirements it was utlized in this study to predect if a network traffic is normal or attack traffic[11].

Review of Litreture

The authors of [12] have proposed a proactive security technigue that measures DDOS attack volume fpr the purose of addressing the limitation of the response time of reactive security system.the.

To detect eleven different DDoS attacks the researchers.in [13] have used six machine learning algorithm .The CICDDoS2019 dataset was used , and the stud involved eleven different database files in CSV format.The performance of Logistic regression, Decision tree, Random Forest, Ada boost, KNN, and Naive Bayes were evaluted using the eleven dataset and they conclude that , and determined the best classification algorithms for detection.They concluded that the Decision tree and random forest algorithms gave have low efficiency compared to . Logistic regression, Ada Boost, KNN, and NB show good results. in [14] a mathmatical model to detect DDOS was proposed , Logistic Regression and Naive Bayes were used. The CAIDA 2007 Dataset was used to train and test the algoriths. Weka tools was used for the implementation and result analysed .they concluded that the performance of logistic regration was better than Naves Bayes algoitm. Based on Random Forest the researchers in [15] have applied a noval technogue to reduce the DDoS attack traffic on the top level Domain Name System on the internet. The accuracy of the classifer was 99.2% .

Methodology

The approach for detecting DDoS traffic included multiple steps. First, two datasets (normal and DDoS attack) were generated by setting up a network that consisted of a Windows 2012 DNS server, a Windows XP client, and one Kali Linux machine, all configured and run in VMware. To simulate normal DNS server traffic, a script was created to continuously run the nslookup command from the Windows XP client to the DNS server. Wireshark was used to capture network traffic, and the file was saved and exported as a CSV file for future use in the prediction stage using the logistic regression algorithm. In addition, to generate the DDoS attack dataset, an attack was launched from the Kali Linux machine to the DNS server using the hping3 tool, and, as in the previous step.

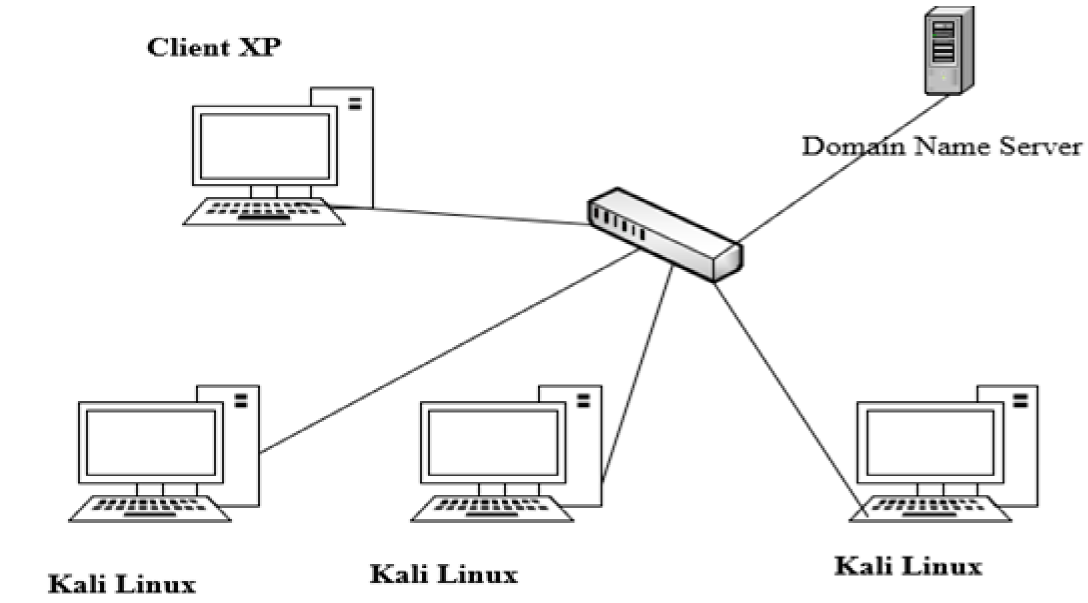
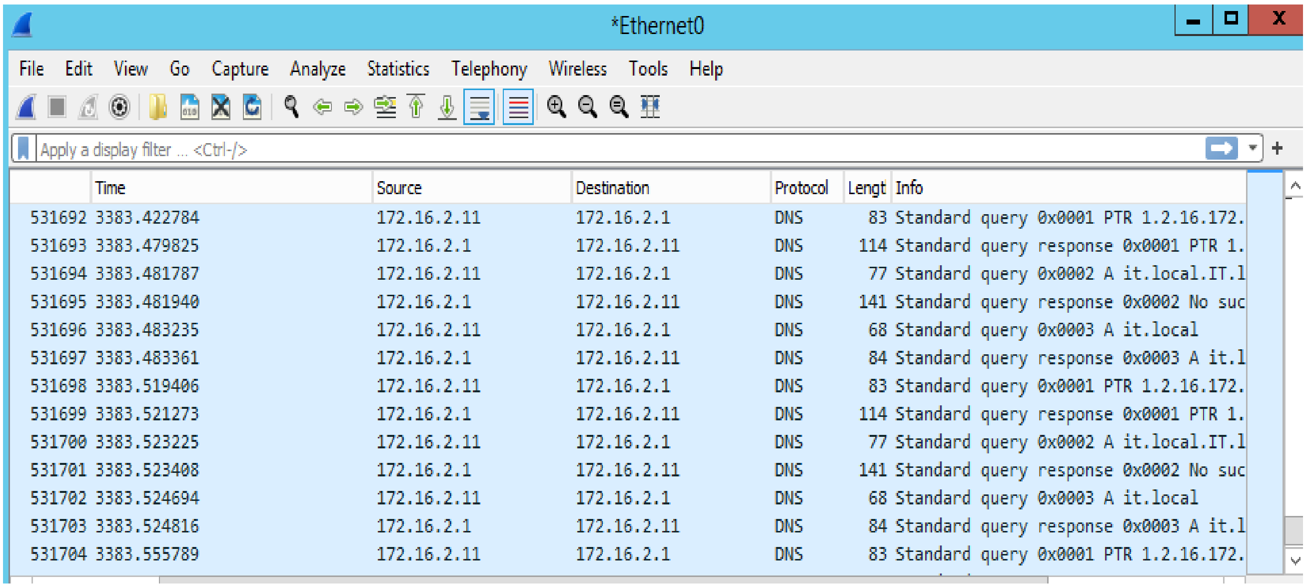


Figure2 : Experiment Setup

Wireshark is an open soursce tool . It is an example of a disruptive technology that has been created and maintained by a worldwide team of protocol specialists. Recelntly , it became the most populer network traffic captering nad anaysier tool [16]. In this paper it was used to capture DNS traffic in normal and attack cases . The caputred files were exported form wireshark as csv files .Figure two illustrated a sample of generated traffic in wireshark .



	Time	Source	Destination	Protocol	Lengt	Info
531692	3383.422784	172.16.2.11	172.16.2.1	DNS	83	Standard query 0x0001 PTR 1.2.16.172.
531693	3383.479825	172.16.2.1	172.16.2.11	DNS	114	Standard query response 0x0001 PTR 1.
531694	3383.481787	172.16.2.11	172.16.2.1	DNS	77	Standard query 0x0002 A it.local.IT.l
531695	3383.481940	172.16.2.1	172.16.2.11	DNS	141	Standard query response 0x0002 No suc
531696	3383.483235	172.16.2.11	172.16.2.1	DNS	68	Standard query 0x0003 A it.local
531697	3383.483361	172.16.2.1	172.16.2.11	DNS	84	Standard query response 0x0003 A it.l
531698	3383.519406	172.16.2.11	172.16.2.1	DNS	83	Standard query 0x0001 PTR 1.2.16.172.
531699	3383.521273	172.16.2.1	172.16.2.11	DNS	114	Standard query response 0x0001 PTR 1.
531700	3383.523225	172.16.2.11	172.16.2.1	DNS	77	Standard query 0x0002 A it.local.IT.l
531701	3383.523408	172.16.2.1	172.16.2.11	DNS	141	Standard query response 0x0002 No suc
531702	3383.524694	172.16.2.11	172.16.2.1	DNS	68	Standard query 0x0003 A it.local
531703	3383.524816	172.16.2.1	172.16.2.11	DNS	84	Standard query response 0x0003 A it.l
531704	3383.555789	172.16.2.11	172.16.2.1	DNS	83	Standard query 0x0001 PTR 1.2.16.172.

Figure 3: Simple of Captured Traffic Using Wireshark

Then , using the python programming languqge, the recorded files from wireshark were exported, and data processing , feature extraction and lableling were performaned on them..DNS response time and packet length in byte were selected.as key features in this study .Next labeling process was conducted , with lable zero to identify normal DNS traffic and lable one used for DDoS attack . in addition , the two dataset were combined and splitted to traing and testing datasets .Finaly , the logistic regration algoritms was implemented and run in python .

Evaluation of the Model

The logistic regration algorithm is implemented in python and its performane was evaluted using Precision , Recall and F1-socre evaluate the performance of the classifier following metriccs were chosen .Precision provides a clear explanation of the proportion of accurately predicted cases that resulted in positive outcomes. The main goal of recall is to describe the percentage of true positive cases that are accurately identified. When both precision and recall ratings are required for the model's evaluation, the F1 Score is employed [17].

Table 1:Results of Logistic Regration Algorithm

	Precision	Recall	F1-score
Normal Traffic (0)	1.00	0.86	0.92
Attack traffic (1)	0.88	1.00	0.93

Using Classification_report function which is included in scikit-learn in python a classification report was genreated as follows

Table 2: Classification Report

Metric	Precision	Recall	F1-score
Accuracy	-	-	0.93
Macro Average	0.94	0.93	216.183
Weighted Avero;age	0.94	0.93	216.183

Conclusion

In this study, the logistic regression model is used to predict DDoS attacks on DNS servers. Two original datasets were generated using VMware and Wireshark tools. Two scenarios were conducted: one to generate normal DNS traffic and another to generate DDoS attack traffic using the Hping3 tool. The Python programming language was used to implement the algorithm. The evaluation metrics used were precision, recall, and F1 score. The results for normal traffic prediction were 100%, 86%, and 92% respectively for the three algorithms. For DDoS attack traffic, the results were 88% for precision, 100% for recall, and 93% for the F1 score.

References

[1] A. K. Sharma and R. Kumar, "A Comprehensive survey of DDoS Attacks: Evolution, Mitigation and Emerging trend," in *2024 3rd International conference on Power Electronics and IoT Applications in Renewable Energy and its Control (PARC)*, 2024, pp. 185-188 :IEEE.

[2] T. E. Ali, Y.-W. Chong, and S. J. A. S. Manickam, "Machine learning techniques to detect a DDoS attack in SDN: A systematic review," vol. 13, no. 5, p. 3183, 2023.

[3] J. J. N. S. Nazario, "DDoS attack evolution," vol. 2008, no. 7, pp. 7-10, 2.008

[4] H. Gao *et al.*, "An empirical reexamination of global DNS behavior," in *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*, 2013, pp. 267-278.

[5] N. Palladino, M. Santaniello, N. Palladino, M. J. L. Santaniello, Power,, and I. i. t. M. I .G. A. I. Transition, "Introduction: The IANA Transition and Internet Multistakeholder Governance," pp. 1-20, 2021.

[6] Y. Xie, S. Tang, X. Huang, C. Tang, and X. J. C. C. Liu, "Detecting latent attack behavior from aggregated Web traffic," vol. 36, no. 8 ,pp. 895-907, 2013.

[7] T. Ni, X. Gu, and H. J. T. I. J. o. E. E. Wang, "Detecting DDoS attacks against DNS servers using time series analysis," vol. 12, no. 1, pp. 753-761, 2014.

[8] L. Fang, H. Wu, K. Qian, W. Wang, and L. Han, "A Comprehensive Analysis of DDoS attacks based on DNS," in *Journal of Physics: Conference Series*, 2021, vol. 2024, no. 1, p. 012027: IOP Publishing.

[9] R. Alonso, R. Monroy, and L. A. J. S. Trejo, "Mining IP to domain name interactions to detect DNS flood attacks on recursive DNS servers," vol. 16, no. 8, p. 1311, 2016.

[10] T. En-Najjary, G. Urvoy-Keller, M. Pietrzyk, J.-L. J. E. Costeux, Department of Networking, and F. Security: Biot, "Traffic classification: Application-based feature selection using logistic regression," 20.10

[11] R. Bapat *et al.*, "Identifying malicious botnet traffic using logistic regression," in *2018 systems and information engineering design symposium (SIEDS)*, 2018, pp. 266-271: IEEE.

[12] D. Kwon, H. Kim, D. An, and H. Ju, "DDoS attack volume forecasting using a statistical approach," in *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, 2017, pp. 1083-1086: IEEE.

[13] K. B. Dasari and N. J. I. d. S. d. I. Devarakonda, "Detection of Different DDoS Attacks Using Machine Learning Classification Algorithms," vol. 26, no. 5, pp. 461-468, 2021.

[14] K. Kumari and M. J. J. o. B. D. Mrunalini, "Detecting Denial of Service attacks using machine learning algorithms," vol. 9, no. 1, p. 56, 2022.

[15] L. Chen, Y. Zhang, Q. Zhao, G. Geng ,and Z. J. P. c. s. Yan, "Detection of dns ddos attacks with random forest algorithm on spark," vol. 134, pp. 310-315, 2018.

[16] U. Banerjee, A. Vashishtha, and M. J. I. J. o. c. a. Saxena, "Evaluation of the Capabilities of WireShark as a tool for Intrusion Detection," vol. 6, no. 7, pp. 1-5, 2010.

[17] O. Oluwasanya, A. J. U. J. o. E. Braimah Joachim, and A. Sciences, "Credit card fraud detection using logistic regression and isolation forest algorithms," vol. 2, no. 1, pp. 187-195, 2023.