DDoS Attack Detection System for WhatsApp

Jithin S
Assistant Professor
Computer Science and Engineering
St. Thomas College of Engineering and Technology, Mattannur
Kerala, India
jithinsubash@stthomaskannur.ac.in

Arpit Ramesan
Computer Science and Engineering
St. Thomas College of Engineering and Technology, Mattannur
Kerala, India
arpitramesan777@gmail.com

Akhil K
Computer Science and Engineering
St. Thomas College of Engineering and Technology, Mattannur
Kerala, India
akhilajith821@gmail.com

Anupama U
Computer Science and Engineering
St. Thomas College of Engineering and Technology, Mattannur
Kerala, India
anupamaulhas@gmail.com

Anooja V
Computer Science and Engineering
St. Thomas College of Engineering and Technology, Mattannur
Kerala, India
anoojaprakashan@gmail.com

***Abstract*—Distributed denial-of-service (DDoS) attacks are one of the most common and severe threats ever in today's evolving cyber security environment. Their ability to shut down network services while imposing millions of dollars in financial damage has proven effective, therefore prevention is essential for both enterprises and governmental organizations.**

In this project, our primary goals are to identify potential DDoS attacks and develop effective solutions to prevent them. The application gets integrated into the network and is constantly induced to check the network's incoming traffic and this further utilizes the possible features to avoid or block such possibilities prior to any obscure or bigger threats. Methods like complete IP blocking, black hole filtering, casting, and an alert notification system can be implemented for the user to be alerted and take necessary measures for securing their system from overloading or any other faults happening to their systems. In the given scenario the device and data of the user are considered to be a prime concern. Apart from that, this project mainly focuses on tackling the problem of getting attacks to "WhatsApp" as multiple messages at a time in the form of Bomb messaging. Bomb messaging is considered to be a logical DDoS attack, in which an attacker sends an unknown load of random messages to the user, overloading the user system. We are trying to tackle this issue and possibly improve upon it in the future.

***Index Terms*—Bomb message, WhatsApp, DDoS, Attacks**

I. Introduction

A denial-of-service attack (DoS attack) is a cyberattack that aims to make a computer or network resource unavailable to its intended users by temporarily interrupting the operations of a host that is linked to a network. Typically, a denial of service attack involves bombarding the targeted computer or resource with unnecessary requests to overburden the system and prevent some or all valid requests from being performed. The incoming traffic that floods the victim during a distributed denial-of-service assault (DDoS attack) comes from a variety of sources. Since there are numerous sources, it is necessary to employ more advanced mitigation techniques to prevent this kind of attack.

The three most common forms of DDoS attacks include protocol attacks, volumetric attacks, and application layer attacks. DDoS attacks are difficult to prevent, but DDoS defense systems, rate limiting, real-time packet analysis, and web application firewalls can provide some protection. The most basic DoS attack uses brute force to overwhelm the victim with an excessive amount of packets, overtax its connection bandwidth, or exhaust its system resources. The attacker's capacity to produce the massive flux of packets is what allows for bandwidth-saturating floods. Today, distributed denial-of-service attacks using a botnet are a popular method of achieving this. An application layer DDoS attack is typically carried out to target specific objectives, such as blocking transactions and database access. Attacks on the application layer have the potential to interfere with services like information retrieval or website search capabilities.

II. Problem Definition

A distributed denial-of-service (DDoS) attack is a malicious attempt to disrupt a server, service, or network's regular traffic by overloading the target or its surrounding infrastructure with an excessive amount of Internet traffic. Due to their scattered nature, these attacks are also very challenging to fight against. It might be challenging to distinguish between regular web traffic and DDoS attack requests. There are certain defenses you can use to lessen the likelihood of a successful DDoS attack. As you can see, a DDoS assault may cause serious harm to any website, costing victims money, reputation, and SEO rating, as well as raising the likelihood that they will be hacked. This might amount to millions of data packets and thousands of IP addresses that are continually changing during a DDoS attack, which state tables must keep track of. The vast amount of memory and processing power needed to do that fast for every packet is incredibly high, and the majority of firewalls can't handle the load.

A Denial of Service (DoS) attack method's primary objective is to prevent a website from being accessible:

1) The speed at which the website responds to valid requests may decrease.
2) It is possible to completely disable a website, preventing authorized people from using it.

Information about website visitors cannot be stolen by DDoS attacks. A DDoS attack's main objective is to overburden the website's resources. DDoS attacks, however, can be utilized for extortion and blackmail. Website owners could be required to pay a ransom to cease a DDoS attack, for instance. DDoS attacks may also be motivated by political, hacktivist, terrorist, or even commercial rivalry. A DDoS attack can be launched against an organization by anyone with financial or ideological motivations.

III. Related Works

With the aid of CNN and machine learning techniques, numerous researchers have conducted their studies on DDoS attack detection. This study provides a summary of the earlier suggested techniques for detecting DDoS attacks. The primary goal is to identify the drawbacks of the earlier studies and make recommendations for the upcoming work. Numerous algorithms using different methodologies have been created, but there hasn't been a literature review that compiles the practices that are currently in use. In *DDoS attack detection and classification via Convolutional Neural Network (CNN)*, the convolutional neural network (CNN) technique is demonstrated to identify and categorize DDoS traffic into legitimate and malicious data with a 99% accuracy using two separate datasets. In *Detection of Distributed Denial of Service Attacks in SDN using Machine learning techniques*, to identify malicious communications, they suggest using a machine learning technique called Decision Tree and Support Vector Machine (SVM). In *Analysis and Detection of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques*, an attack was made on the owncloud environment using Tor Hammer, and an intrusion detection system was used to create a new dataset. In *Design and Implementation of IoT DDoS Attacks Detection System based on Machine Learning*, they suggested an IoT DDoS attack detection system based on machine learning that includes IoT devices, IoT gateways, SDN switches, and cloud servers.

IV. Proposed System

In our system, we are trying to implement 3 main components. These components are capable of managing the necessary data required for the user to be safe in WhatsApp, without the risk of being attacked by an attacker and getting the user system overloaded. So the main components are as follows:

1) The main detection system: This system is responsible for constantly detecting messages that come from people to the user in WhatsApp. Based on the message or text, the system classifies the message as spam or normal text. This is done by a message count factor. Based on the number of messages that the person sends, also if there exist special characters that lag the system performance, it is also classified as spam.
2) The notification system: Based on the classification, this system sends notifications to the user if there exist spam messages. Otherwise, no notifications are sent. The notification is either through an sms based system that is available online through online methods, a mail response regarding the attack and action taken for it, or simply a WhatsApp message from the developer team saying so and so a person has attempted to attack. The last mentioned function will be implemented by a bot. These mentioned functions will only be implemented if the base model or design gets properly and fully implemented and if further improvements are going to be applied. The user has to log in to their WhatsApp accounts in WhatsApp web primarily to implement or start detecting the messages. Messages would be detected using Python as a scripting tool for interacting with the web user interface for WhatsApp.
3) The block and report system: Using the WhatsApp interface by utiliZing python scripts, we can use the block and report button for directly blocking and reporting the spam message that is sent by the attacker. This is done when the data is detected to be successful. We can interact with the interface that is set in the WhatsApp website using Python. This usage will be able to provide us with interaction with the interface, without any direct contact with the WhatsApp system.

V. System Architecture

The architecture diagram shown below represents a scenario where an attacker sends spam messages to a user, and there is a detection system in place to identify and block those messages. Here's a breakdown of each component: The attacker is the entity that is attempting to send spam messages to the user. The attacker may be a person, a bot, or some other type of program that is designed to send unsolicited messages.

Packets are the content of the message that the attacker is attempting to send to the user. It may include text, advertising, phishing links, or other types of unwanted content. These messages may be of any type which will be uncomfortable to the user.

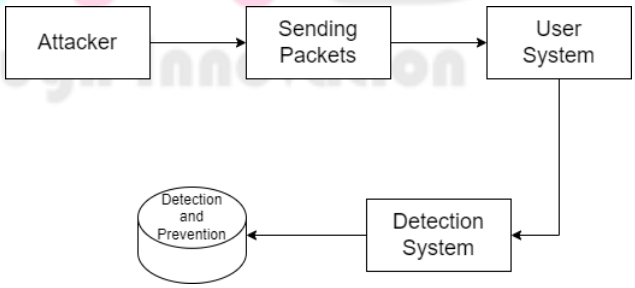

Fig. 1. Architecture Diagram

A User is the intended recipient of the spam message. The user may be an individual or an organization, and they may be using various communication channels in social media. In our case, WhatsApp is the communication channel through which the user receives the messages.

The Detection system is the system that is designed to identify and block spam messages. We used JavaScript to develop this detection system to detect and filter out unwanted messages and block the attacker and also send a notification to the user that someone has tried to spam the account.

The flow of information in this architecture is as follows: The attacker sends a spam message to the user, and the user receives it. The detection system then analyzes the message and determines whether it is spam or not. If the message is identified as spam, the detection system blocks it and sends a notification to the user about the attack. Overall, this architecture is designed to protect users from unwanted and potentially harmful messages, and to prevent attackers from successfully delivering their spam content.

VI. Conclusion and Future Work

As we have seen, distributed DoS assaults are a real danger that harms a lot of Internet users severely. Losses have progressed from being merely annoying to truly being disastrous and devastating for certain users. In this project, we are trying to implement a Mozilla Firefox running extension for monitoring DDoS attacks occurring on WhatsApp web. This extension will be able to detect the number of times a message has been sent by an attacker. Accordingly, the extension will notify the user using alerts and block the user that is attacking. The overall system is currently able to detect, classify, notify, and block any kinds of spam-based attacks towards a user that is using the system. This has been implemented for WhatsApp web, in which the user has to scan a QR code and simply have to enter the number to which the notification will be sent. The complete system works automatically and runs based on a factor mechanism. We can define the factor or count of numbers, up to which is considered or classified as spam message. This is mostly done manually by the developer.

In the near future, to gain better performance and easy access we plan to make an application for our detection system which will be able to communicate directly with the WhatsApp application.

VII. Result

The results regarding the system are as follows:

1) An accurate notification system that operates according to the speed of the network connection.
2) A Blocking system based on the classified spam message.
3) The accuracy of the classification system is about 90%, with a small margin for error caused by slow network access.
4) The detection system also has 90% accuracy due to the lag in network connectivity and hence the detection of the same may vary accordingly.

References

[1] T. Mahjabin, Y. Xiao, G. Sun, and W. Jiang, *A survey of distributed denial-of-service attack, prevention, and mitigation techniques*

[2] S. S. Mohammed, R. Hussain, O. Senko, B. Bimaganbetov, J. Lee, F. Hussain, C. A. Kerrache, E. Barka, and M. Z. A. Bhuiyan, *A new machine learning-based collaborative DDoS mitigation mechanism in software-defined network*

[3] A. Callado, C. Kamienski, G. Szabo, B. P. Gero, J. Kelner, S. Fernandes, and D. Sadok, *A survey on internet traffic identification*

[4] X. Ying, *An overview of overfitting and its solutions*

[5] A. H. Lashkari, G. D. Gil, M. S. I. Mamun, and A. A. Ghorbani, *Characterization of tor traffic using time based features*,

[6] I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, *Deep-Learning Framework to Detect Lung Abnormality–A study with Chest X-Ray and Lung CT Scan Images*, Pattern Recognition Letters (Nov 2019)

[7] R.T.Sausa, O.Marques, F.A.A.M.N.Soares.Et al., *Comparative Performance Analysis of Machine Learning Classifiers in Detection of Childhood Pneumonia Using Chest Radiographs*, Procedia Computer Science 18 (2013)

[8] P.Rajpurkar, J.Irvin, K.Zhu R.L.Ball, H.Mehta B.Yang, et al., *Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy*

[9] A. Lashkari, *CICFlowmeter-V4.0 (formerly known as ISCXFlowMeter) is a network traffic Bi-flow generator and analyser for anomaly detection*

[10] M. S. Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, *DDoSNet: A deep-learning model for detecting network attacks*

[11] A. E. Cil, K. Yildiz, and A. Buldu, *Detection of DDoS attacks with feed forward based deep neural network model*

[12] M. A. Salahuddin, M. F. Bari, H. A. Alameddine, V. Pourahmadi, and R. Boutaba, *Time-based anomaly detection using autoencoder*

[13] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, *Kitsune: An ensemble of autoencoders for online network intrusion detection*

[14] J. Chen, Y. Yang, K. Hu, H. Zheng, and Z. Wang, *DAD-MCNN: DDoS attack detection via multi-channel CNN*,

[15] S. S. Priya, M. Sivaram, D. Yuvaraj, and A. Jayanthiladevi, *Machine learning based DDOS detection*,

[16] O. Elejla, B. Belaton, M. Anbar, B. Alabsi, and A. Al-Ani, *Comparison of classification algorithms on icmpv6-based DDoS attacks detection*,

[17] O. E. Elejla, M. Anbar, B. Belaton, and S. Hamouda, *Labeled flow-based dataset of ICMPv6-based DDoS attacks*

[18] R. F. Fouladi, O. Ermis¸, and E. Anarim, *A novel approach for distributed denial of service defense using continuous wavelet transform and convolutional neural network for software-defined network*

[19] D. Peraković, M. Periša, I. Cvitić, and S. J. T. J. Husnjak, *"Model for detection and classification of DDoS traffic based on artificial neural network," vol. 9, no. 1, p. 26, 2017.*

[20] A. Sahi, D. Lai, Y. Li, and M. J. I. A. Diykh, *"An efficient DDoS TCP flood attack detection and prevention system in a cloud environment," vol. 5, pp. 6036-6048, 2017.*

[21] Dong,S., Sarem, M, *"DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Attack in Software-Defined Networks", IEEE Access, 8, 5039-5048.*

[22] Dong, S., Abbas, K., Jain, R, *"A survey on distributed denial of service (DDoS) attacks in SDN and cloud computing environments," IEEE Access, 7, 80813- 80828.*

[23] Wani, Abdul Raoof, Q. P. Rana, and Nitin Pandey, *"Analysis and Countermeasures for Security and Privacy Issues in Cloud Computing." System Performance and Management Analytics. Springer, Singapore, 2019. 47-54.*

[24] Xiao, Le, et al, *"A protocol-free detection against cloud oriented reflection DoS attacks"* Soft Computing 21.13 (2017): 37133721.

[25] Lan Li, Gyungho Lee, *"DDoS Attack Detection and Wavelets"*

[26] Laura Feinstein, Dan Schnackenberg, *"Statistical Approaches to DDoS Attack Detection and Response"*