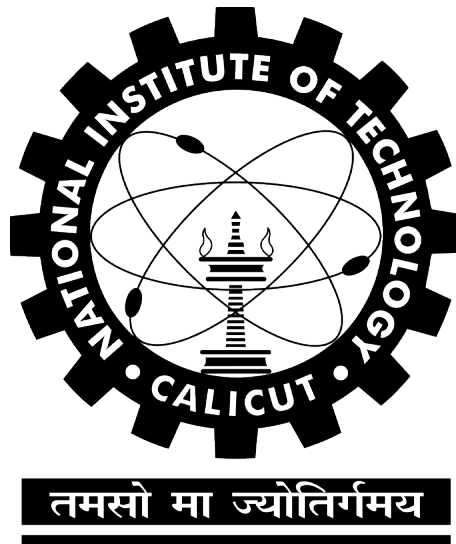


ASSIGNMENT ON

Algorithm for Detecting DDoS Attack

*Submitted by*

Shahin John JS  
B160943CS



Department of Computer Science and Engineering  
National Institute of Technology Calicut  
Calicut, Kerala, India - 673 601

May 25, 2020

# 1 INTRODUCTION

In computing, a denial-of-service attack (DoS attack) is a cyber-attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services of a host connected to the Internet. Denial of service is typically accomplished by flooding the targeted machine or resource with superfluous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled.

In a distributed denial-of-service attack (DDoS attack), the incoming traffic flooding the victim originates from many different sources. This effectively makes it impossible to stop the attack simply by blocking a single source. A DoS or DDoS attack is analogous to a group of people crowding the entry door of a shop, making it hard for legitimate customers to enter, thus disrupting trade.

The main reason behind these situations is that the network security community does not have useful traceback methods to trace attackers and detect the attacks efficiently and effectively immediately after the attack. Some of the work in DDoS attack detection takes more computation time which makes the detection system very complex.

# 2 SUMMARY

In this paper, we propose an algorithm for finding the DDoS attack. We have mixed two strategies from two existing journals to build our algorithm. The Introduction outlined the basic idea of what is DDoS attack and its significance. The below section will state our primary problem statement and what we aim to do in the paper. The motivation for carrying out our problem is given in section 4. We shall depict the scenario across various sector and how detecting DDoS attack is essential for them. In section 5, the Objective of our paper will be shown. The objectives will guide through the procedure of our methodology given in section 6. The section will neatly illustrate the design of our algorithm. The algorithm will detect DDoS Attack using Fast Entropy Approach on Flow-Based Network Traffic. The algorithm will essentially use entropy of the packets send to find the attack. To make the algorithm more effective, the algorithm shall be adaptive to the dataset. The algorithm will desperately attempt to get a DDoS attack, if not found. While attempt to reduce down one one DDos attack, if many DDoS attacks are found. For a full detailed description please refer to section 6. The facilities required for the program to run are given in section 7. In section 8, we give the outcomes of our paper. Results in section 8 and finally conclude with section 10.

# 3 PROBLEM STATEMENT

The paper is aimed at designing and implementing a detection of Distributed Denial of Service attacks algorithm. We expect our algorithm to be an improvement over the existing ones. Since detecting a DDoS attack with 100% accuracy is a difficult task, we restrict our algorithm to finding the attack under some constraint in terms of entropy. The design shall have better time complexity while compromising on exact solution.

# 4 MOTIVATION

Modern distributed denial-of-service (DDoS) attackers are leveraging the convenience of the cloud to access the computing power required to execute "mega" DDoS attacks. The consequence of a DDoS attack on a business can be far reaching and directly contributes to significant revenue loss, customer data breaches and damage to brand reputation.

However, many businesses today still believe that "it will never happen to me" and often do the bare minimum to protect their web systems and applications from such attacks. Understanding

the motivation behind a DDoS attack can be helpful in determining the risk to your business as DDoS attacks become a growing problem in frequency and size.

- **Financial Motivation:** The most prevalent motivation behind an attack is financial gain. Cybercriminals are able to ransom or extort businesses with unprotected web systems and applications who want to avoid being attacked or wish to stop an attack. Often requesting hard to acquire bitcoins as payment with the ransom being increased by the day if a business can't pay. This can lead to hours, if not days of downtime of e-commerce and other business critical applications.
- **Non-financial motivations:** Of the many non-financial reasons a cybercriminal might launch an attack the most common include a protest or "hacktivism" against a business practice or organizations whose ideologies differ from theirs. For example, groups such as Anonymous have been known to attack businesses who have been affiliated with political candidates, have been deemed to be controlling the internet and even medical organizations for what they feel is questionable medical care of minors.
- **Cover for targeted attacks:** Another motivation that we are seeing more frequently is that a DDoS attack is used as cover for other more sophisticated targeted attacks. According to a Computer Weekly article published in October 2016, "The majority of DDoS attacks (53%) resulted in additional compromise, including viruses (46%), ransomware (15%) and other malware (37%)." This same report found that 21% of these attacks resulted in customer data theft.

With the growing technologies, it is necessary to detect or mitigate the DDoS attacks.

## 5 OBJECTIVES

Our algorithm is based on the following objectives.

- Flow aggregation for doing flow based attack detection.
- Fast Entropy computation to detect DDoS attack with less computation time.
- Adaptive Threshold Algorithm to improve detection accuracy.

## 6 METHODOLOGY

- **Flow Aggregation**

Flow is a unidirectional series of IP packets of a given protocol travelling between a source and a destination IP/port pair within a certain period of time. Flow aggregation techniques are used to aggregate flows into a single flow with a larger granularity of classification giving a flow count for each connection with a unique combination of attributes given below for a packet.

- Source IP
- Destination IP
- Source Port
- Destination Port
- Protocol

Aggregated flows have a larger number of packet information that dramatically reduces the amount of monitoring data. Hence, Internet traffic flow profiling has become a useful technique in the passive measurement and analysis field. Instead of considering the packet count of each connection, the proposed method calculates the flow count of each connection at particular time interval for detecting the flooding attacks, increasing the speed of analysis and reducing the time complexity of our algorithm.

- **Entropy Approach**

In this detection approach, entropy of flow count is calculated for each connection using the formula given below.

$$H_{i,t} = -\log(x_{i,t}/(x_{1,t} + x_{2,t} + \dots + x_{n,t})) + R_{i,t}$$

If  $(x_{i,t} \geq x_{i,t+1})$  then  $r \leftarrow |\log(x_{i,t+1}/x_{i,t})|$  else  $r \leftarrow |\log(x_{i,t}/x_{i,t+1})|$ .

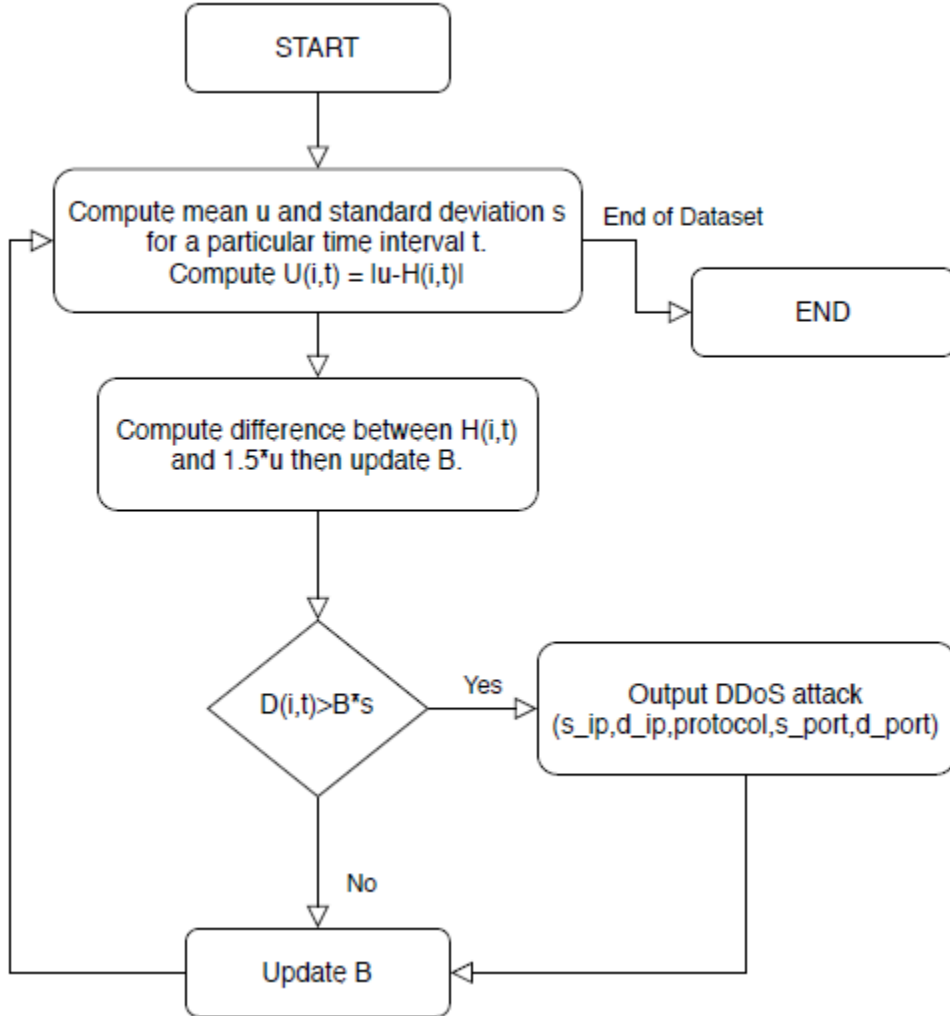


Figure 1: Algorithm Flowchart

It is essentially a standard way of calculating any entropy in a system. When there is an attack, entropy drops drastically, because there is one flow count that is dominating. In the non-attack case, the entropy will be in a constant range. Let a random variable  $x_{i,t}$  represent the flow count of a particular connection  $i$  over a given time interval  $t$ .

- *Adaptive Threshold Algorithm*

In flooding attack detection, threshold value is very important. Threshold value needs to be updated according to the packet traffic condition. On one hand, if an attacker sends malicious traffic with small change in traffic when the channel is stable, the detector cannot detect the attack with high value of  $B$ , where  $B$  is the threshold multiplication factor. Because of the steady channel condition and stealthy attack pattern, the detection facility does not work properly with highly set  $B$ . On the other hand, if the channel is burst but the detector has small  $B$ , the detector works very sensitively in this situation. As a result, the detector yields many false positives, which are not severe but a bad characteristic of the detector. In the proposed method  $B$  value is updated based on entropy value. The detection algorithm is shown in figure below. The  $B$  will be changed under the following rules:

If  $(H_{i,t} > 1.5u_t)$  then  $B \leftarrow B + 1$

If  $(H_{i,t} < 0.5u_t)$  then  $B \leftarrow B - 1$

$H_{i,j}$  refers to the fast entropy,  $u_t$  and  $s_t$  is the mean and standard deviation of flow count during a particular time interval.  $D_{i,t}$  is defined as the difference between the mean value  $u_t$  and the fast entropy  $H_{i,t}$ . While applying adaptive threshold algorithm, if  $D_{i,t}$  is greater than the product of  $B$  and  $s$  indicates flooding attack.

## 7 FACILITIES/ EQUIPMENT REQUIRED

Network traffic dataset must contain at least the following attributes in it.

- Source IP
- Destination IP
- Source Port
- Destination Port
- Protocol

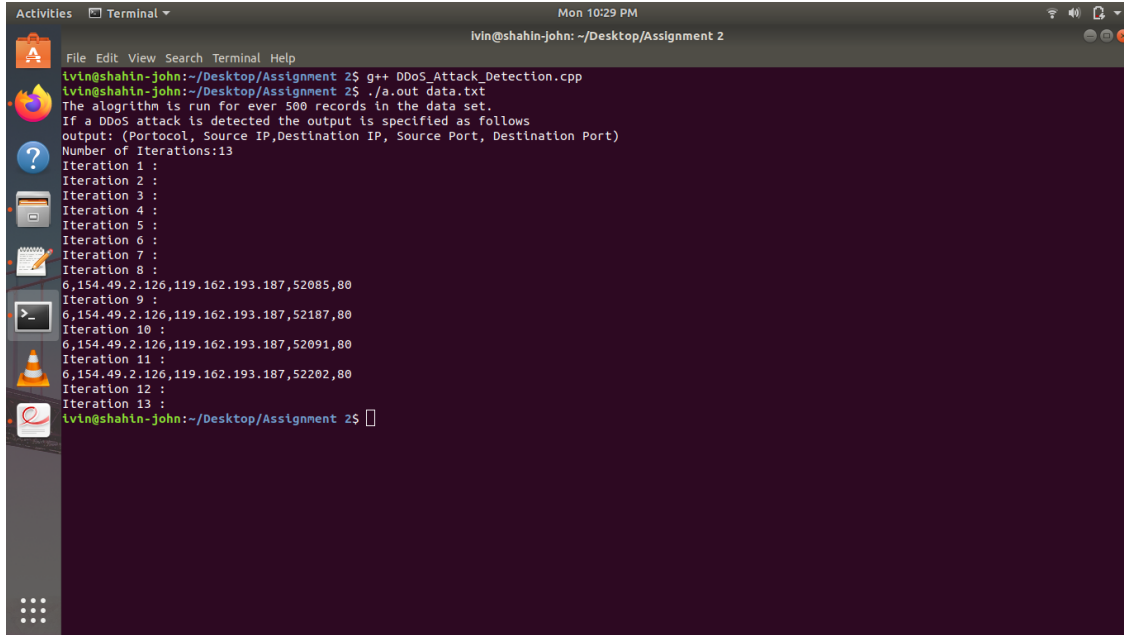
All other computations are carried out by the program compiler. Since we use the strategy of entropy, any additional data mining tools are not needed.

## 8 OUTCOME

A code given the methodology was compiled and evaluated. The algorithm was able to detect some of the DDoS attack if not all for a network traffic dataset. The adaptive algorithm was able to adapt to the entropy in the dataset for each iteration. As iterations go by the algorithm becomes more sceptical and desperately attempts to detect an attack. While, if the algorithm finds many DDoS attacks then the algorithm assumes that it had overlooked the dataset. Thus, attempts to reduce down to a single DDoS attack. This way, we can be sure that the algorithm will eventually find at least one DDoS attack in the entire dataset.

## 9 RESULT

Experiments are carried out on a network traffic dataset used in DDoS detection evaluation program. The proposed method is based on flow based analysis which requires only packet header information. In flow aggregation, the header information was aggregated in a particular time interval, which belongs to identical 5 tuple (Source IP address, Destination IP address, Source port, Destination port, Protocol Number). By flow aggregation, processing overhead is reduced and speed of analysis is increased. The algorithm was tested and has a complexity of  $O(|Dataset| * |t|)$



```
ivins@shahin-john:~/Desktop/Assignment 2$ g++ DDoS_Attack_Detection.cpp
ivins@shahin-john:~/Desktop/Assignment 2$ ./a.out data.txt
The algorithm is run for ever 500 records in the data set.
If a DDoS attack is detected the output is specified as follows
output: (Portocol, Source IP, Destination IP, Source Port, Destination Port)
Number of Iterations:13
Iteration 1 :
Iteration 2 :
Iteration 3 :
Iteration 4 :
Iteration 5 :
Iteration 6 :
Iteration 7 :
Iteration 8 :
6,154.49.2.126,119.162.193.187,52085,80
Iteration 9 :
6,154.49.2.126,119.162.193.187,52187,80
Iteration 10 :
6,154.49.2.126,119.162.193.187,52091,80
Iteration 11 :
6,154.49.2.126,119.162.193.187,52202,80
Iteration 12 :
Iteration 13 :
ivins@shahin-john:~/Desktop/Assignment 2$
```

Figure 2: Screenshot

where  $|Dataset|$  is the number of records in the dataset and  $|t|$  is the number of records selected over a period of time  $t$ . The algorithm does not produce exact solution but still is an improvement over existing methods.

## 10 CONCLUSION

An efficient DDoS attack detection method using entropy approach was designed and implemented. The flow count is calculated for each connection at particular time interval. From the observation it is clear that the fast entropy value is considerably reduced for particular connection and particular time interval of which flow count is large value compared to rest. DDoS attack is detected, when the difference between entropy of flow count at each instant and mean value of entropy in that time interval is greater than the threshold value. Since the threshold value is updated adaptively based on traffic pattern condition, the accuracy of detection is improved.