On Selection of Attributes for Entropy Based Detection of DDoS

Sidharth Sharma  
Department of Computer Science  
National Institute Of Technology  
Rourkela, India  
sidharth.sharma091@gmail.com

Santosh Kumar Sahu  
Department of Computer Science  
National Institute Of Technology  
Rourkela, India  
santoshsahu@hotmail.co.in

Sanjay Kumar Jena  
Department of Computer Science  
National Institute Of Technology  
Rourkela, India  
skjena@nitrkl.ac.in

Abstract—
Distributed Denial of service (DDoS) attack is an attempt to prevent the legitimate users from using services provided by service providers. This is done through flooding their server with the unnecessary traffic. These attacks are performed on some prestigious websites like Yahoo, Amazon and on various cloud service providers. The severity of the attack is very high, as a result the server goes down for the indefinite period of time. To detect such attempts, various methods were proposed. In this paper, an entropy-based approach is used to detect the DDoS attack. We have analyzed the effect on the entropy of all the useful packet attributes during DDoS attack and tested their usefulness against famous types of distributed denial of service attacks. During analysis, we have explained the proper choice of attributes one should make to get a better threshold during DDoS detection.

Keywords—DDoS, Entropy, SYN Flood, Attributes Selection

I. INTRODUCTION

Distributed Denial of service attack possesses the biggest threat to Internet services. The attacker creates a botnet of internet connected computers (each of them has been maliciously taken over through malware like Trojan horses) and attack through them so it is a hard to locate attackers and block them. This makes research to detect and mitigate DDoS more interesting. Attackers are finding new techniques to perform attacks day by day so as researchers are using new ideas to counter them.

Since the evolution of DDoS attacks different type of detection mechanism being proposed. Cabrera et al. [1] used network management systems that use MIB (Management Information Base) variables to detect precursors of attack, change in these MIB variables during attack used as a technique to detect attacks. This method was statistical in nature. Jeong et al. [2] Stated that during denial of service attack very few IP addresses appear which are very less in comparison to the appearance during flash events they also compared these two events on various parameters. Lee et al. [3] Proposed a unique path fingerprint scheme that represents the route of the IP packet has traversed. For each client, they create an entry for corresponding fingerprint if a spoofed packet is coming from some other route it has been discarded. Liao et al. [4] used a K-nearest neighbor classifier to classify patterns into normal or intrusive classes. They have tested with the 1998 DARPA BSM dataset. Gavrillis et al. [5] presented an approach using Radial basis function neural network detector for DDoS attacks. A small number of statistical descriptors were used to distribute behavior of DDoS and classification is achieved using RBF-NN.

In this paper, we are analyzing an entropy-based approach to detect DDoS. Many authors have worked previously on entropy based approach to detect DDoS. [6]-[9]

We have analyzed all the useful packet attributes whose entropy values could deviate during DDoS attack and explain the usefulness of these attribute selection in different type of DDoS attacks (TCP-SYN flood, UDP flood, Smurf attacks) by using some DDoS attack datasets that is explained in further sections.

II. ENTROPY BASED DETECTION

Entropy is a well-known and valuable concept in information theory. It is the measure of uncertainty associated with a random variable and describes the degree of dispersal or concentration of a distribution. It was introduced by Claude E. Shannon in "A Mathematical Theory of communication", 1948.

Entropy can be well utilized in the detection of DDoS because after analyzing DDoS attacks, researchers examined that Higher Volume of traffic, incomplete Connections, flooding of packets are characteristics of the attack. Therefore, entropy could be used to calculate the distribution randomness of the packet attributes. These attributes could be Source IP, Destination IP, Source Port, Destination Port, Length, Protocol, Flags. We can check the entropy of all useful attributes of the packet. If entropy is high, it means distribution is random.

Entropy is calculated by examining a series of packets, refer as the window. If a window consists N packets, then entropy could vary from 0 to log N. Entropy is zero when all the values of the distribution are same, it is highest when all the values are different.

If there are N elements in a window, then Entropy of random variable X is defined as H(X), where a random variable X is taking values from x1,x2,x3........xn and probability of occurrence of these values are p1,p2,p3,......pn respectively.

\[ H(X) = \sum_{i=1}^{N} p_i \log p_i \]

Where,

\[ p_i = \frac{\text{number of times } X \text{ having value } x_i}{\text{size of the window}} \]

\[ 0 \leq H(X) \leq \log N \]

\[ H(X) = \begin{cases} 0 & \text{if } p_i = 1, \forall i \in \{1,2,...,N\} \\ \log N & \text{if } p_i = 0, \forall i \in \{1,2,...,N\} \end{cases} \]
As explained earlier, size of UDP attack packets also remains similar during the attack, which considerably decrease entropy values of Length shown in Fig. 15. As the name suggests, UDP flood packets use UDP protocol, so this repetition of protocol decreases the entropy of the protocol field clearly shown in Fig. 16.

3. **Source Port and Destination Port Entropy**

As discussed earlier in this paper Source port of attack packets are usually random these days, so not a big deviation during the attack as shown in Fig. 17, so as Destination Ports. In case of Destination port attacker first check which ports are open on the server side. Then the attacker targets only those ports that are open to disturb the service given by the server. Since attacker attacks all the open ports randomly, so this hardly affects entropy of Destination ports as shown in Fig. 18.

**IV. CONCLUSION**

In this paper we have analyzed how the entropy values of different attributes changes when subjected to attack. We have checked the entropy values of attributes like Source IP, Destination IP, Flags, Length, Protocol, Source Port and Destination Port against 3 major types of Distributed Denial of Service attacks namely TCP-SYN attack, Smurf attack, UDP flood attack. In most of the research works related to entropy based detection researchers focused on Source IP attribute. We have seen that in most of cases the Source IP of the packets is spoofed, but if an attacker choose them from a smaller pool then only entropy will deviate during attack. But, experiments shows that attributes like Destination IP, Length, Flags and Protocol are better selections for actual deviation of entropy values during the attack. Through simulation we have shown that in majority of the cases these attributes are better selections for entropy based detection than Source IP.

**V. REFERENCES**


