Volume 4, Issue 9, September 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Secure Network Access by Flow Analysis Based Detection against Denial of Service Attack (DDoS)

Kanchan H. Patil*, Prof. A. B. Bagwan
Department of Computer Enginnering & Pune University
Maharashtra, India

Abstract –In a fast moving growth of Distributed Service over internet, Dist ributed Denial of service (DDoS) is more challenging and critical threat over the internet.Internet is a worldwide network that conveys an expandable range of information resources and services which lead to bulk exchange huge traffic every day. Excessive popularity of internet creates some troubles in the networks. Botnets are gradually arise. More latest and improved environment on Bot Masters aided to disable detectors by frequently watching and updated to recover flash crowds. Flash crowd and Distributed Denial of Service (DDoS) attacks are the two major events among them. Botnets are the main drivers of cyber-attacks, such as distributed denial of service (DDoS). Experienced botmasters attempt to disable detectors by mimicking the traffic patterns of flash crowds. During a flash event, a Web server should aim to differentiate DDOS attacks from genuine flash crowds. This is a challenging task to those who defend against DDoS attacks. It was found that the current attack flows are usually more similar to each other compared to the flows of flash crowds. Based on this, the flow correlation coefficient is used as a similarity metric to differentiate DDOS attack flows from genuine flash crowd flows.

Keywords– Network flow, DoS attack, DDoS attack, Botnet, Botmaster, Flash crowd, Flow similarities, Discrimination.

I. INTRODUCTION

Due to excessive and fast moving growth of internet it is becoming more vulnerable day by day. Attackers are continuously trying to gain unauthorized access over it. To gain unauthorized privileged access, one attempt by attacker is Denial of Service (DoS) attack. A denial of service, or DoS, is a category of attack in which genuine traffic will be ignored and the access to system's users will be denied. A Distributed Denial of Service (DDoS) is an improved form of a traditional DoS attack which uses distributed systems to launch on the target system, and botnets are usually the engines behind it. There are generally two types of events that often overload Web sites and degrade their services:
1. Flash crowds are due to a sudden, large surge in traffic to a particular Web site, created by legitimate requests.
2. Denial of Service attacks (DoS) contains malicious requests to subvert the normal operation of the website.

In survey [2] of the 70 largest Internet operators in the world have found that DDoS attacks have increased dramatically in recent years. In order to sustain botnets, botmasters take advantage of various ant forensic techniques such as code obfuscation, memory encryption [4], fresh code pushing for resurrection [5], peer-to-peer implementation technology [6], [7], [8], or flash crowd mimicking [9], [10] to conceal their traces.

Flash crowds are unexpected, sudden surges of access to a server, but genuine, such as breaking news. Flash-crowd attacks are the one of the dangerous form of distributed denial of service (DDoS). They flood the victim with service requests generated from many bots. Attack requests are similar in content to those generated by legitimate users. So during the flash event, the main aim of server is to identify flash crowd attacks or DDoS attacks from genuine flash crowds. This paper presents an approach using flow similarity to differentiate DDoS attacks flows from genuine flash crowds flows because it was found that the current attack flows are usually more similar to each other compared to the flows of genuine flash crowd flows.

II. LITERATURE SURVEY

Research survey over a novel flow similarity-based approach to discriminate DDoS attacks from flashcrowds, which remains an open problem to date. Lot of research have been done in the past to differentiate DDoS attacks from flash crowds. These works generally focused on extracting DDoS attack features, which was then followed by detecting and filtering DDoS attack [11],[12] packets by the known features which cannot actively detect DDoS attacks.

Use of graphical puzzles to differentiate between bots and humans is the current most popular defence mechanism against flash crowd attacks. This method involves human responses. Xie and Yu tried to differentiate DDoS attacks from flash crowds using user browsing dynamics [13] at the application layer. Oikonomou and Mirkovic tried to differentiate DDoS attacks from flash crowds by modeling human behavior. This method also works well at the application layer.

Numbers of reports are there on size and organization of botnets [6], [8]. Botnet infiltrations are implemented to collect first-hand information about their activities. Peer-to-peer-based botnet have been implemented by Wang et al for research purposes [7].

Following observations are noted regarding current botnets as:

1. For one botnet prebuilt programmed attack tools are same. Botmaster issues command to all notes in his botnet to start one attack session which can be evidenced from the literature of Botnet [3], [5], [6].
2. The attack flows those are observed at the victim end are an aggregation of many original attack flows. The aggregated attack flows share a similar standard deviation as an original attack flow and the flow standard deviation is usually smaller than that of genuine flash crowd flows.

The reason behind this phenomenon is that the number of live bots of a current botnet is far less than the number of concurrent legitimate users of a flash crowd. A botmaster has to trigger his live bots to generate many more attack packets, than that of legitimate users of a flash crowd in order to launch a flash crowd attack. This in turn results that the aggregated attack flow possesses a small standard deviation compared with that of a flash crowd.

III. IMPLEMENTATION DETAILS

A. DETECTION / DIFFERENTIATION METHOD

1. Detection method is based on flow analysis which uses feature of flow similarity to differentiate DDoS attacks from genuine flash crowds under current botnet size and organization, addressing the problem of differentiation at the network layer.
2. Differentiation algorithm works independently of specific DDoS flooding attack types.
3. Defferentiation method computes correlation coefficient [1] which makes it delay proof and effective against explicit random delay insertion among attack flows.

B. DATA FLOW ARCHITECTURE

Figure1 shows a Data flow architecture of proposed method which is described by following steps

Step1 - Capture input network flows coming towards community network to count the number of packets for every flow.
Step2 - Identify high requests forwarded to the same destination.
Step3 - Identify distinct flows among multiple flows.
Step4 - Differentiation method computes flow fingerprint and flow correlation coefficient values.
Step5 - Confirm DDoS attack and evaluate the differentiation accuracy of differentiation method.

C. DIFFERENTIATION ALGORITHM

Step1 : Get the input network flows.
Step2 : For all network flows
i) Do Clustering.
ii) Calculate Flow Strength.
iii) Calculate Flow Finger Print.

Step3 : Claculate the P_X (Flow correlation coefficient) value for Combination of network flows.
Step4 : If (P_X > Threshold)
{
DDOS attack flows
}
else
Flash crowd flows(Legitimate flows).

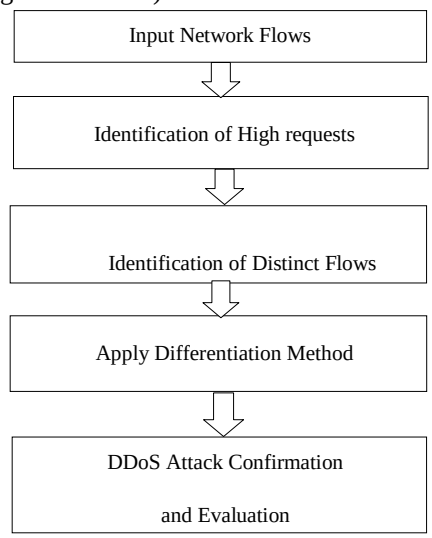

Figure1: Data Flow Architecture

D. MATHEMATICAL MODEL

1. X [1……..M] = $\{X_1, X_{2,X_3} .....X_M\}$, M≥1,

X is a network flow.

2. $X_i = \{x_{i_1}, x_{i_2}, x_{i_3} .....x_{i_N}\}$, $1 \le i \le M$ and

N ≥ 1, X_i is i^{th} network flow.

3. $X_{iK} = \{x_{i1}, x_{i2}, x_{i3} .....x_{iK}\}$, $1 \le K \le N$,

Number of packets counted in k^{th} time interval.

4. X = <Protocol, Source IP, Source Port, Destination IP, Destination Port>
Protocol = <TCP, UDP>
Source IP = <32 bit Source IP Address>
Destination IP = <32 bit Destination IP Address>
Source Port = <16 bit Source Port number>
Destination Port = <16 bit Destination Port number>

5. I_{X_i,X_j} = {0, 1}, $1 \le i, j \le M, i \ne j$, $X_i, X_j \in X$,

Similarity indicator.

I_{X_i,X_j} = {1}, DDoS attack Indicator.

E. DEFINITIONS

1. Network Flow

For a given community network, cluster the network packets that share the same destination address as one network flow. Network flow is defined as,

$$X_i = \{x_i[1], x_i[2].....x_i[N]\} \quad (1)$$

Where, X_i - given network flow, N – length of given network flow, $x_i[k](1 \le k \le N)$ - represents the number of packets that we counted in the k^{th} time interval for the network flow.

2. Flow Strength

Expectation of given flow is defined as flow strength of that flow. Flow strength represents the average packet rate of a network flow.

$$E[X_i] = \frac{1}{N}\sum_{n=1}^{N} x_i[n] \quad (2)$$

Where, $E[X_i]$ - expectation of the flow (flow strength)

3. Flow Fingerprint

Flow fingerprint is the unified representation of the given network flow.

$$X_i' = \{X_i'[1], X_i'[2]...., X_i'[N]\} = \{\frac{x_i[1]}{N * E[X_i]}, \frac{x_i[2]}{N * E[X_i]}, ..., \frac{x_i[N]}{N * E[X_i]}\} \quad (3)$$

Where, X_i' - fingerprint of flow X_i

On the basis of definition (2) and (3), a network flow and its fingerprint is related as,

$$X_i = N * E[X_i] * X_i' \quad (4)$$

Since $\sum_{k=1}^{N} X_i'[k] = 1$

Correlation between the two flows is given as,

$$r_{X_i,X_j} = \frac{1}{N}\sum_{n=1}^{N} x_i[n]x_j[n] \quad (5)$$

It may be indicated zero correlation although the two flows are completely correlated with a phase difference. The definition therefore is modified as:

$$r_{X_i,X_j}[k] = \frac{1}{N}\sum_{n=1}^{N} x_i[n]x_j[n+k] \quad (6)$$

Where, k (k = 0, 1, 2, ….., N-1) indicates the position shift of flow X_j.

4. Flow Correlation Coefficient

Flow Correlation Coefficient indicates similarity between two flows. Correlation coefficient of the two flows is defined as

$$\rho X_i, X_j[k] = \frac{rX_i, X_j[k]}{\frac{1}{N}\left[\sum_{n=1}^{N-1} x_i^2[n] \sum_{n=1}^{N-1} x_j^2[n]\right]^{1/2}}$$

F. FLOW SIMILARITY BASED DIFFERENTIATION METHOD

For a given community network there is only one server which is under attack or experiencing a flash crowd at any given time. The attack packets enter the community networks which are generated by only one botnet, therefore the finger-prints of the attack flows are the same.

The numbers of packets for every flow are counted and this information is recorded for a short term. Detection method is independent of network topology because it is using network flow.

For sampled M network flows $X_1, X_2 ....., X_M$ .Obtain the flow correlation coefficient of any two network flows, $X_i (1 \le i \le M)$ and $X_j (1 \le j \le M, i \ne j)$. An indicator for the similarity is I_{X_i, X_j} of flow X_i and X_j , and which has only two possible values: 1 indicates DDoS attacks and 0 otherwise. Let δ be the threshold for the differentiation as,

$$I_{X_i, X_j} = \begin{cases} 1, & \rho X_i, X_j[k] \ge \delta, \\ 0, & \text{otherwise}, \end{cases} \qquad (7)$$

Where, $1 \le i, j \le M$ and $i \ne j$

Generally, there may have more than two suspected flows in a community network. Therefore pairwise comparisons can be conducted. The final decision can be derived from these comparisons in order to improve the reliability of decision.

IV. RESULT

1. Figure1 shows the resultant graph for three different input files. Each file contains input network packet flows. Flow correlation coefficient is calculated for two different network flows in each file. And the flow correlation coefficient for two flows is compared with threshold value. Here the threshold value is defined in between 0 and 1(considered as 0.5 for comparison).If the correlation coefficient is greater than threshold then packets are considered as attack (DDoS) packets otherwise legitimate (flash crowd) packets.

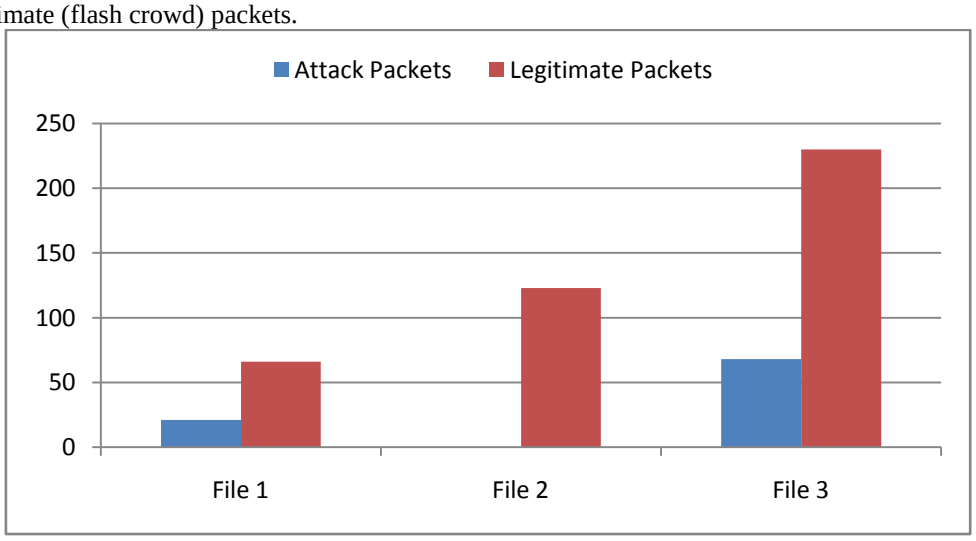

Figure 1: Resultant Graph showing Separation of Attack (DDoS) Packets from Legitimate (flash crowd) Packets.

V. CONCLUSION

Proposed method in this paper tried to differentiate distributed denial of service attacks from genuine flash crowds. Under the current conditions of botnet size and organization, it has noted that DDoS attack flows have more similarity than genuine flash crowd flows. So the flow correlation coefficient is used here to measure similarity among suspicious flows and to confirm DDoS attack. Evaluation confirms the effectiveness of the proposed method.

Future work will focus to find the possibility of organizing a super botnet, with a sufficiently large number of live bots which can the proposed method and to explore actions which there should take against attackers' actions.

REFERENCES

[1] Shui Yu, Weijia Jia, Song Guo, Feilong Tang and Yong Xiang "Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient" IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 23, NO. 6, JUNE 2012.

[2] Arbor, "IP Flow-Based Technology," 2011.

[3] B. Stone-Gross, M. Cova, B. Gilbert, G. Vigna,, M. Szydlowski, R. Kemmerer, L. Cavallaro and C. Kruegel "Your Botnet Is My Botnet: Analysis of a Botnet Takeover,"Proc. ACM Conference. Computer Communicatin. Security, 2009.

[4] A. Hackworth and N. Ianelli, "Botnets as Vehicle for Online Crime,"Proc. 18th Ann. First Confe.. 2006.

[5] C.Y. Cho, J. Caballero, V. Paxson, D. Song and C. Grier "Insights from the Inside: A View of Botnet Management from Infiltration," Proc. Third USENIX Confe. LargeScale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More (USENIX LEET), 2010.

[6] V.L.L. Thing, N. Dulay and M. Sloman "A Survey of Bots Used for Distributed Denial of Service Attacks,"Proc, pp. 226-244,2007.

[7] E. Biersack, F.C. Freiling, T. Holz, M. Steiner and F. Dahl, "Measurements and Mitigation of Peer-to-Peer-Based Botnets: A Case Study on Storm Worm,"Proc. First Usenix Workshop Large-Scale Exploits and Emergent Threats (LEET), 2008.

[8] M. Bailey, , Y. Xu, M. Karir, F. Jahanian and E. Cooke "A Survey of Botnet Technology and Defenses,"Proc. Cyber security Applications and Technology Conference. for Homeland Security, 2009.

[9] J. Jung, B. Krishnamurthy and M. Rabinovich "Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites,"Proc. 11th Int'l Conf. World Wide Web (WWW),pp. 252-263, 2002.

[10] A. Scherrer, N. Larrieu, P. Borgnat, P. Abry and P. Owezarski "Non-Gaussian and Long Memory Statistical Characterizations for Internet Traffic with Anomalies, "IEEE Transaction Dependable Secure Computing, vol. 4, no. 1, pp. 56-71, Jan.-Mar. 2007.

[11] G. Carl, G. Kesidis, S. Rai and R. Brooks "Denial-of-Service Attack-Detection Techniques, "IEEE Internet Computing, vol. 10,no. 1, pp. 82-90, Jan./Feb. 2006.

[12] Y. Chen and K. Hwang, "Collaborative Detection and Filtering of Shrew DDoS Attacks Using Spectral Analysis," J. Parallel Distributed Computing, vol. 66, no. 9, pp. 1135-1155, 2006.

[13] Y. Xie and S.-Z. Yu, "Monitoring the Application-Layer DDoS Attacks for Popular Websites,"IEEE/ACM Transaction Networking,vol. 17, no. 1, pp. 15-27, Feb. 2009.

