DENIAL OF SERVICE ATTACKS (DDOS) DETECTION AND PREVENTION BASED ON PACKETS FLOW OF THE WEBSITE

R.Rajalakshmi , A.Mani , B.Raja

M.Tech(cse) II year, Prist University, Chennai, TamilNadu-600 014, ***rajee787@gmail.com ,***

Assistant Professor, Prist University, Chennai,Tamil Nadu-600 014, India ***Invic.mani @gmail.com,***

M.E(cse) Dhanalakshmi Engg College Chennai, TamilNadu-600 014,India ***kbm.raja@gmail.com***

Abstract

Nowadays, In a fast moving growth of Distributed Service over internet by a Distributed Denial of service (DDoS) is more challenge due to a critical threat over the internet and 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. Botnets are usually the engines behind Denial of Service Attacks (DDoS). Sophisticated botmasters attempt to disable detectors by mimicking the traffic patterns of flash crowds. This poses a critical challenge to those who defend against DDoS attacks.Ultimate and foremost planning to defend against DDoS attacks by flash crowds. To propose a most relevant and correcting the problems through Discrimination method and corresponding algorithm results to found optimized results.

Keywords: flash crowds, discrimination, Distributed Denial of service (DDoS),Bot Masters.

I. INTRODUCTION

In this paper, ResearchSurvey over a novel flow similarity-basedapproach to discriminate DDoS attacks from flashcrowds, which remains an open problem to date. DistributedDenial of Service (DDoS) attacks pose a critical threatto the Internet. Motivated by huge financial rewards, such asrenting out their botnets for attacks or collecting sensitiveinformation for malicious purposes, hackers are encouragedto organize botnets to commit these crimes [2]. Furthermore,in order to sustain their botnets, botmasters takeadvantage of various anti-forensic techniques to disguisetheir traces, such as code obfuscation, memory encryption [3], peer-to-peer implementation technology [4], [5], [6], or flash crowdmimicking . Flash crowds are unexpected, butlegitimate, dramatic surges of access to a server. This is referred to as a flash crowd attack.Denial of service attacks are performed by the attackers which randomly increases thebandwidth floods to attacks the target websites to deny the service to the users. In recent years, the arrival of Distributed Denial-of- Service (DDoS) open-source bot-based attack tools facilitating easy code enhancement, and so resulting in attack tools becoming more powerful. Developing new techniques for detecting and responding to the latest DDoS attacks often entails using attack traces to determine attack signatures and to test the techniques. However, obtaining actual attack traces is difficult, because the high-profile organizations that are typically attacked will not release monitored data as it may contain sensitive information. Present a detailed study of the source code of the popular DDoS attack bots, Agobot, SDBot, RBot and Spybot to provide an in-depth understanding of the attacks in order to facilitate the design of more effective and efficient detection and mitigation techniques.DDOS attacksand flash events can both overload the server or the server's Internet connection and result in partial or complete failure. Unlike DDOS attacks, which are simply malicious requests that do not have to be handled by a Web site, flash events consist of legitimate requests. A Web serverhas the responsibility to try and handle as many of the requests as possible during a flash event.We use discrimination algorithm using the flow correlation coefficient as a similarity metricamong suspicious flows.The currentmost popular defence against flash crowd attacks is the useof graphical puzzles to differentiate between humans and bots.Flash-crowd attacks are extremely challenging because they request legitimate and business-critical content. Thus their traffic appears legitimate-like, which makes defences that detect and filter malicious traffic ineffective against flash crowd attacks. We define the security model to capture the request from each client and identify the level of network traffic generated by them is recognized internally by the website and blocks the misbehaving client by recognizing the ipaddress of the client and blocks them from access the website which minimizes the workload of the website. We differentiate flash crowd attack from DDOS attack by assigning a threshold value if the maximum packets generated by the each client for each time is monitored by the security model and blocks the misbehaving users.

To note the following facts concerning the currentbotnets after our thorough study:

1. The attack tools are prebuilt programs, which arebeusually the same for one botnet. A botmaster issues acommand to all bots in his botnet to start one attacksession. This can be evidenced from the literature of botnet [2], [4], [5].

2. The attack flows that we observe at the victim endare an aggregation of many original attack flows,and the aggregated attack flows share a similarstandard deviation as an original attack flow, andthe flow

standard deviation is usually smaller thanthat of genuine flash crowd flows. The reason forthis phenomenon is that the number of live bots of acurrent botnet is far less than the number ofconcurrent legitimate users of a flash crowd.The comparison among the proposed method and theprevious ones can be found in the online supporting material.

We found a new feature of flow similarity to defeat flash crowd attacks under current botnet size and organization. It is the first work in this field to the best of our knowledge. Within the relevant literature,flash crowd attacks continue to be a challenge.Our work sheds light on a new perspective in addressing this problem at the network layer.The proposed algorithm works independently of specific DDoS flooding attack genres. Therefore,it is effective against unknown forthcoming flooding attacks. The proposed correlation coefficient-based method is delay proof. This property is very effective against explicit random delay insertion among attack flows.

2.Definitions and Problem Setting

A Denial of Service attack is an attempt by a person or a group of persons to cripple an online service. Distributed denial-of-service attacks (DDoS) pose an immense threat to the Internet users to keep endorsed users of a website or web service from accessing it, or limiting their ability to do so.In the existing method, extracting DDoS attack features, and was followed by detecting and filtering DDoS attack packets by the known features. However, these methods cannot actively detect DDoS attacks. For a given router in a local network (e.g., a community network), we cluster the network packets that share the same destination address as one network flow. And the delays among the attack flows from different bots depend on normal Internet delays, and therefore are limited compared to fast Internet transportation facilities.

If a DDOS attack or flash crowd attack occurs during a flash event, a Web server should aim to Ignore DDOS requests and handle the legitimate requests. This requires the Web site to be able to distinguish between the two sets of requests and block both types of attacks. We characterize Flash crowd and DDoS attack along the following dimensions:
Traffic patterns: Traffic patterns as seen by the Web site are important for several reasons. Overall traffic volume determines how much a server should provision resources to keep the site operational up to a certain level. If server load exceeds its maximum tolerance level which is pre-defined by its capacity, the server begins to slow down and can be driven to a shutdown. Thus, watching traffic patterns allows us to articulate the period when an unusually large number of clients can overwhelm a site and how much time the server has from the start of an FE or DDoS to take defensive measures. In this section, we begin by presenting a number of preliminary definitions, and then discuss the setting of the discrimination problem. For simplicity, we use the terms flow and network flowinterchangeably in this paper.

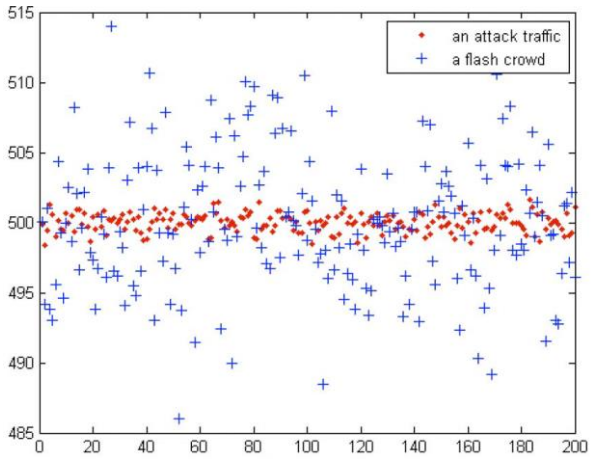

Fig 1. The difference between aggregated attack traffic and a flashcrowd traffic under the current botnet size and organization.

Definition 1 (Network Flow). For a given router in a localnetwork (e.g., a community network),we cluster the networkpackets that share the same destination address as onenetwork flow.

$$X_i = \{x_i[1], x_i[2], \dots, x_i[N]\}$$

Definition 2 (Flow Strength). For a network flow Xi, let thelength of the network flow beN(N>1). We define theexpectation of the flow as the flow strength of Xi.

$$E[X_i] = \frac{1}{N}\sum_{n=1}^{N} x_i\,[n]$$

Definition 3 (Flow Fingerprint). For a given network flow Xi with length N, its fingerprint X0i is the unified representation of Xi, namely,

$$\Gamma_{X_i,X_j}[K] = \frac{1}{N}\sum_{n=1}^{N} x_i\,[n]x_j[n]$$

Definition 4 (Flow Correlation Coefficient). Let Xi and Xj($i \neq j$) be two network flows with the same length N. We define the correlation coefficient of the two flows as

$$\rho_{X_i.X_j}[k] = \frac{\Gamma_{X_i,X_j}[K]}{\frac{1}{N}\left[\sum_{n=1}^{N-1} x_i^2[n]\sum_{n=1}^{N-1} x_j^2[n]\right]}$$

3. Similarity-Based Detection Method

In this area, to present the similarity-based detection method against flash crowd attacks. For a given community network, we set up an overlay network on the routers that we have control over. We execute software on every router to count the number of packets for every flow and record this information for a short term at every router. Under this framework, the requirement of storage space is very limited and an online decision can be achieved. Once an access surge on the server occurs, our task is to identify whether it is a genuine flash crowd or a DDoS attack. DDoS attack flow can be discriminated from flash crowds by the flow correlation coefficient at edge routers under two conditions: the length of the sampled flow is sufficiently large, and the DDoS attack strength is sufficiently strong. DDoS attack flow can be discriminated from flash crowds by the flow correlation coefficient at edge routers under two conditions: the length of the sampled flow is sufficiently large, and the DDoS attack strength is sufficiently strong. According to our proposal, when a possible DDoS attack alarm goes off, the routers in the community network start to sample the suspected flows by counting the number of packets for a given time interval, for example, 100 milliseconds. When the length of a flow, N, is suitable,

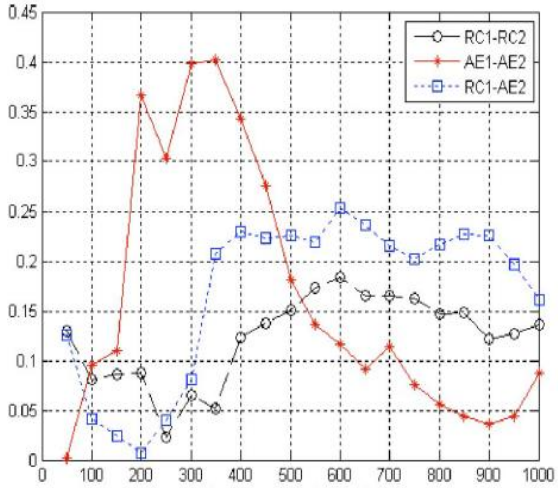

Fig 2. The flow correlation coefficient against length of flows in theWorld CUP 98 data set.

We start to calculate the flow correlation coefficient between suspected flows. Suppose we have sampled M network flows, X1, X2; . . . ,XM, therefore,

$$I_{x_i,xj} = \begin{cases} 1, & \rho_{x_i,x_j}[k] \geq \delta, \\ 0, & otherwise, \end{cases}$$

Where $1 \leq i, j \leq$ Mand $I \neq j$.

The flow correlation coefficient is used to indicate similarity between two flows. It is sometimes the case that two similar flows may have a phase difference which will decrease the correlation coefficient. Fortunately, this is easy to deal with because we can shift one flow to match the other and take the maximum value of the correlation coefficients to represent the similarity of two flows.

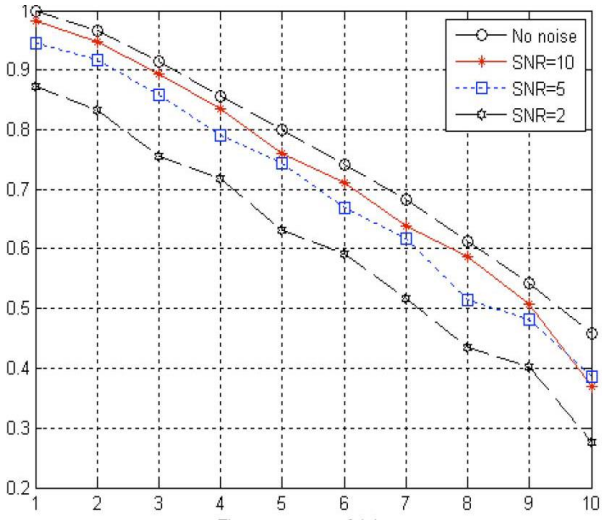

Fig 3.The flow correlation coefficient of attack flows against backgroundnoise and delays

4.Analysis onthe Proposed Method

In this section, we first prove that flash crowds and DDoSattacks can be differentiated using the Flow correlationcoefficient in theory. Following this foundation, we analyse the effectiveness of the proposed discrimination method,and prove that the threshold 8 in (8) does exist.In order to make our analysis clear, we make thefollowing assumptions:

1. There is only one server in a community networkwhich is under attack or experiencing a flash crowdat any given time.
2. The attack packets enter the community network via minimum of two different edge routers.
3. In one attack session, all the attack packets aregenerated by only one botnet, therefore the fingerprintsof the attack flows are the same.
4. The network delays are discrete and countable.

In the proposed method, the current most popular defense against flash crowd attacks is the use of graphical puzzles to differentiate between humans and bots. This method involves human responses and can be annoying to users. These behavior-based discriminating methods work well at the application layer. However, we have not seen any detection method at the network layer, which can extend our defence diameter far from the potential. We set up an overlay network on the routers that we have control over. We execute software on every router to count the number of packets for every flow and record this information for a short term at every router. Under this framework, the requirement of storage space is very limited and an online decision can be achieved.

5. Performance Evaluation

In this section, we demonstrate the effectiveness of the proposed detection method. We investigate the issue with areal data set first, followed by more general studies in order to achieve general results. The reason that we chose these two distributions for the simulation is that the Pareto distribution has been identified by researchers as the best one to represent network traffic.The Gaussian distribution is a general distribution in nature, and combinations of Gaussian distributions with different parameters can approximate other distributions two flows from the same distribution law (e.g., two Pareto flows with different parameters) is usually higher than that of two flows from different distribution laws (e.g., one from the Pareto distribution and another one from the Gaussian distribution).Here security model to capture the request from each client and identify the level of network traffic generated by them is recognized internally by the website and blocks the misbehaving client by recognizing the ip address of the client and blocks them from access the website which minimizes the workload of the website.

6. Summary and Future Work

In this paper, Attempt to discriminate flash crowd attacks from genuine flash crowds, which is a really tough and open problem for researchers. Widely found that DDoS attack flows possess higher similarity compared with that of flash crowd flows under the current conditions of botnet size and organization. To improve the flow correlation coefficient as a metric to measure the similarity among suspicious flows to differentiate DDoS attacks from genuine flash crowds. By theoretically proved the feasibility of the proposed detection method, and our experiments confirmed the effectiveness of the discrimination method within the current botnet size and organization. To discussed the possible anti detection methods from the attackersperspective. In the fu ture work the project has been enhanced to analyse and read the attackers way of attack hence it is necessary to explore which actions should take against attacker's actions. Secondly the cost of reconfiguration should be reduced by detecting the attack in proactive and reactive manner. DDoS attack sources have a form of pattern behaviour of packet transmission, with the predictability of known patterns being a very effective approach in detecting them. Here propose two methods using the correlation coefficient to detect the known patterns and tested these methods with generated data and a real dataset from the website of online movie ticket reservation.It has been found that the hidden predictable behaviour from both datasets.

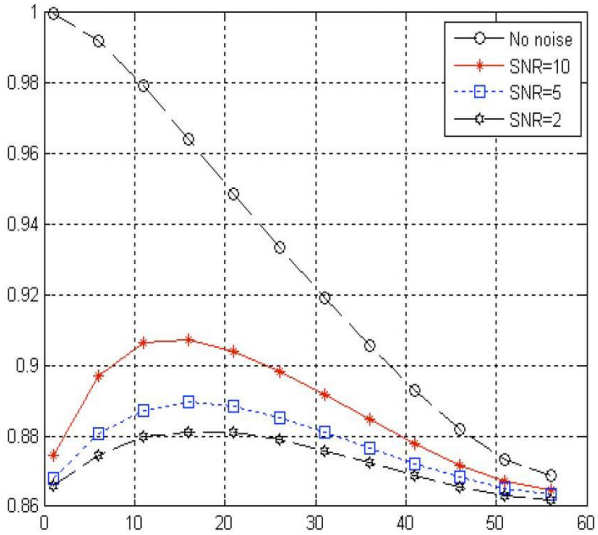

Fig 4. The flow correlation coefficient attack flows against a numberof merged attacks and background noise.

Further the work will be implemented to identify and protect the server from this two as well as from the upcoming network attacks. Implement security model to prevent sql injection to prevent data loss from database. SQL injection is an attack in which malicious code is inserted into strings that are later passed to an instance of SQL Server for parsing and execution. Any procedure that constructs SQL statements should be reviewed for injection vulnerabilities because SQL Server will execute all syntactically valid queries that it receives. Even parameterized data can be manipulated by a skilled and determined attacker.

REFERENCES

[1] ShuiYu,WanleiZhou,WeijiaJia, Song Guo, Yong Xiang, and Feilong TangDiscriminating DDoS Attacks from FlashCrowds Using Flow Correlation Coefficient, Ieee Transactions On Parallel And Distributed Systems, Vol. 23, No. 6, June 2012

[2] Arbor, "IP Flow-Based Technology," http://www.arbornetworks.com, 2011.

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

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

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

[6] V.L.L. Thing, M. Sloman, &N. Dulay, "A Survey of Bots Usedfor Distributed Denial of Service Attacks," Proc. SEC, pp. 229-240,2007.

[7] T. Holz, M. Steiner, F. Dahl, E. Biersack, and F.C. Freiling,"Measurements and Mitigation of Peer-to-Peer-Based Botnets: ACase Study on Storm Worm,"

Proc.FirstUsenix Workshop Large-Scale Exploits and Emergent Threats (LEET), 2008.
[8] M. Bailey, E. Cooke, F. Jahanian, Y. Xu, and M. Karir, "A Survey ofBotnet Technology and Defenses," Proc. CybersecurityApplicationsand Technology Conf. for Homeland Security, 2009.
[9] J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash Crowds andDenial of Service Attacks: Characterization and Implications for
CDNs and Web Sites," Proc. 11th Int'l Conf. World Wide Web(WWW), pp. 252-262, 2002.
[10] A. Scherrer, N. Larrieu, P. Owezarski, P. Borgnat, and P. Abry,"Non-Gaussian and Long Memory Statistical Characterizationsfor Internet Traffic with Anomalies," IEEE Trans. Dependable SecureComputing, vol. 4, no.1, pp. 56-70, Jan.-Mar. 2007.

