APPROACH OF DETECTION AND TRACING TECHNIQUE OF DDOS ATTACKS FROM FLASH EVENT USING FCC

[1]D.B. Pawar

[1]Assistant Professor
[1]Information Technology Department
[1] Sinhgad Institute of Technology and Science, Narhe,India

Abstract : Internet is a wide network used to combine different sectors, education, business, banks, government, entertainment, optical network technologies. It carries number of information services and resources which exchange large amount of traffic over the Internet every day. The growing needs of such applications to make it more prone towards malicious users who are trying to invade. Protection against different software attacks is one of the key challenges to maintain data integrity and privacy. Among them, Distributed Denial of Service (DDoS) and Flash Crowd attacks are the two major events. Web services require security and stability and from these two concerns there are some methods that can differentiate DDoS attack from flash crowd and trace the sources of the attack in large amount of traffic in network. But it is difficult to detect the exact sources of DDoS attacks in traffic of network when flash crowd event is also present. Due to the alikeness of these two irregularities, attacker can easily mimic the harmful flow into legitimate network traffic patterns and The existing defense mechanism fail to detect real sources of attack on time. After analyzing the characteristics of DDoS attacks and the existing Algorithms to detect DDoS attacks, this paper proposes a novel detecting and tracing algorithm for DDoS attacks based on flow correlation coefficient. In this paper, flow correlation coefficient, a theoretic parameter, is used to differentiate DDoS attack from flash Crowd and trace the sources of the DDoS attack. The proposed approach focuses majorly on the efficiency and scalablity features with minimum overhead in terms of resources and time, removal of traffic pattern dependency, increase in detection rate between DDoS and flash crowd and also trace the sources of DDoS attack.

IndexTerms **- DDoS attacks, IP Tracing and detection, Flash Crowd, Differentiation.**

I.Introduction

Distributed denial of service (DDoS) [1] is a critical threat to the user and has caused a huge economic loss to the victims. Therefore, the detection of traffic irregularity is important to secure the today's networks. Flash crowd and DDOS attack are identified and blocked by detection and prevention methods. Attack detection, tracing and prevention methods aims to secure the network by crashing servers of DDOS and flash crowd attack [2]. This proposed approach aims to increase the global security level and is the best solution to DDOS attacks in theory. Both denial of service and flash crowd attacks have the similar impact on web servers. So we demonstrate a way to differentiate between them using our FCC security model to identify the network traffic, so that web servers can attempt to serve normal clients and drop requests from clients involved in attacks and also to block the users who misbehaves in network. Attack detection aims to detect DDoS attacks and also helps to distinguish attack traffic from legitimate traffic [4]. A flash event (FE) is a large amount of traffic to a particular web site causes a dramatic increase in server load and putting severe strain on the network links leading to the server, which results in significant increase in network traffic [3]. A distributed denial of service attack is an explicit attempt by attackers to prevent legitimate users of a service from using that service[5]. DDoS attacks and flash events can both overload the server or the server's internet connection and result in partial or complete failure. This causes a critical challenge to those who defend against DDoS attacks. So to overcome this problem, we proposes a novel approach to differentiate

DDoS from flash event using the flow correlation coefficient as a similarity metric among suspicious flows [7].In a deep study of the size and organization of current botnets, it is found that the current attack flows are usually more similar to each other compared to the flows of flash crowds. Based on this, we observed that it is better to increase the rate of differentiation between DDoS and flash crowd with the help of proposed novel algorithm of differentiation using the flow correlation coefficient as a similarity metric among suspicious flows [3]. In this paper flow correlation coefficient is used to differentiate DDoS attack from Flash Crowd and trace the sources of the DDoS attack. So it will be more effective to increase the rate of distinguishing between DDoS from flash crowd and trace the sources of the DDoS attack experimentally.

The rest of the paper is organized as follows. Section II reviews literature survey. Section III describes problem statement of differentiating DDoS attack from Flash Crowd by using flow correlation coefficient and its detailed architecture .The novel tracing and detection algorithm for proposed system are described in Section V. Section VI describes implementation, expected result set, dataset and performance of FCC system .We conclude the work in Section VII.

II. Literature Survey

As we know, Internet users are increasing day by day. All important services which are based on Internet needs to be maintained properly so that the users can avail them whenever they need. If the network resources and services are not available in time then it will create a crisis. As the number of hosts in Internet is increasing, the threats to it are also increasing. DDoS attacks are one of the most deadly threats rising in internet [5].

In DDoS attacks the attacker use various means to exhaust the resources of a desired server/system so that the other requests cannot be processed and hence bring the services down. The amount of DDoS attack has been increasing drastically in recent years [7].

In this section we are presenting the different methods which are previously used for differentiation and also discussing some advantages and limitations of these systems.

- In paper [1] author used information distance technique to distinguish DDoS from flash crowd .Both these attacks are motivated different methods to measure the similarity among flows such as Abstract distance metrics, Jeffrey distance,Sibson distance,Hellinger distance.After comparison among these four metrics, it is found that the Sibson distance is the most suitable method. By applying an algorithm to the real datasets ,an accuracy around 65%and it is very efficient to improve an accuracy of the flow based discrimination strategy.
- In [2] DDoS is distinguished from flash crowd by using probability metrics. They proposed main contributions to distinguish DDoS attacks from Flash crowds as hybrid metric and the Bhattacharyya metric. The hybrid metric can reduce the false positive rate greatly. But the limitation of this method is that it is not applied in the real network situation, and so cannot find out more recognizable characteristics of IP packets .
- Paper [3] presented a packet arrival pattern for distinguishing DDoS from flash event. In this paper, two methods are used, first Behavior based detection which can discriminate DDoS attack traffic from traffic generated by real users and second Pearson's correlation coefficient which can extract the repeatable features of the packet arrivals. The major limitation is two methods are not tested with different packet information such as packet delay and changing rate of port number so that it can test with the real scenarios in real time. So there is no confirmation of the performance from the predictability test.
- In [4] discrimination of DDoS from flash crowdis done with the help of flow correlation coefficient, used as a similarity metric among suspicious flows. Limitations of this method are, the detection rate of differencing DDoS from flash crowd is less, Tracing of the sources of the DDoS attack is not given and it is very hard toidentify DDoS attack flows at sources since the traffic is not so aggregate using world cup dataset.
- In [7] authors proposed a survey of botnet technology and defense system.They described different kinds of networks that have access to different types of visibility and this has a strong impact on the effectiveness of any botnet detection mechanism. They surveyed that botnet behaviour is undiscoverable and these are moving targets.
- In paper [10] characterization and implications of flash crowd and DDoS for content distribution networks CDNs and Web Sites is presented. This method cannot used toobtain larger flash crowd logs from diverse places and experiment against instrumented servers

We surveyed on different techniques to differentiate DDOS attack from flash crowd such as Information distance, Probability Metrics, Packet Arrival Patterns and Flow Correlation Coefficient [1]-[4]. Among these techniques "Flow Correlation Coefficient" shows the better results compared to another three techniques. But after a detail study of this technique, we found some drawbacks i.e. detection rate of differencing DDoS from flash crowd is less, Tracing of the sources of the DDoS attack is not given and it is very hard to identify DDoS attack flows at sources since the traffic is not so aggregate[5].
So, the novel proposed system increases the rate of differentiating DDoS attack from Flash Crowd by using flow correlation coefficient, increases accuracy and also traces the sources of the DDoS attack.

III. PROPOSED SYSTEM

3.1Block Diagram of Proposed System

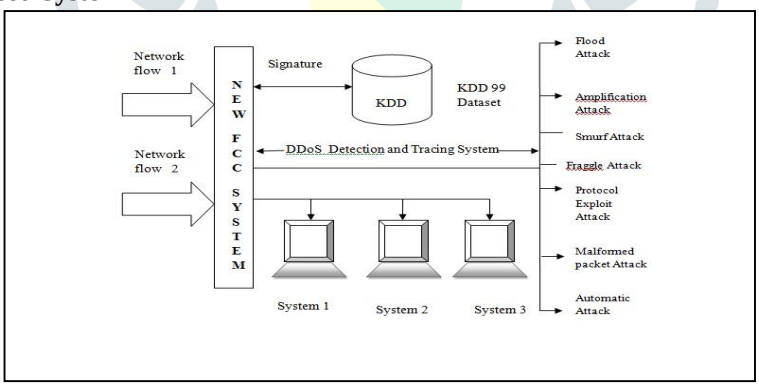


Fig.1.Block diagram of Proposed FCC System and DDOS detection and Tracing System

Two network flows with the same length are given in above fig.1.Detection algorithm using flow correlation coefficient is used to indicate similarity between two flows [4]. It is sometimes the case that two similar flows may have a phase difference which will decrease the correlation coefficient. So it 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. The new FCC system is used to increase the rate of differentiating DDoS from flash crowd.

The proposed FCC system combines parameters from KDD CUP 99 dataset such as time, Duration, Protocol, service flag, src_bytes and dest_bytes of flows at each router to distinguish DDoS from flash crowd. In this way, our novel approach aims to improve the global security level and is the best solution to DDOS attacks in theory.

Also our novel tracing system is used to trace the sources of the DDoS attack .It will detect the subtypes of DDoS attack such as flood attack, amplification attack ,smurf ,fraggle attack etc[6].

IV. Detection and Tracing Mechanisms

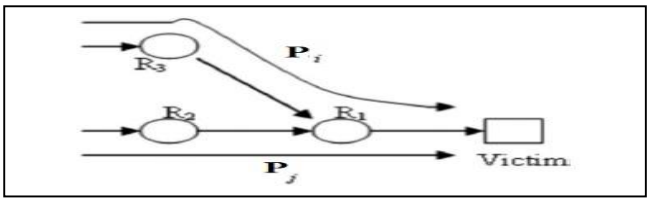
The detection mechanism contains two algorithms to differentiate the DDoS from flash crowd.

1. Network Packet Tracing Algorithm for DDoS using Fuzzy Logic Rules

2. Detection Algorithm using Fuzzy Logic Classifier.

These two phenomenon's reduce the workload over the network, detection time, storage space required for routers and increases performance and scalability.

A sample community network with flows can is given in fig.2 [4]. In the sample community network, R1, R2 and R3 are three routers where R2 and R3 are the edge routers, and we try to protect to the server that is potential victim. Consider, Pi and Pj are two incoming flows observed at R3 and R2, respectively. The two network flows merge at router R1 and both are directed to the potential victim, and enter the community network through different paths. We collect the number of packets for a given network flow with a specific time interval.



3.2 Network Packet Tracing Algorithm

This algorithm monitors the flow at each router in the network. With the help of this algorithm, each router in the network records the entire flow rate that comes either from client or attacker during non-attack, attack and flash crowd period.

In this novel packet tracing algorithm, we are combining Parameters from KDD CUP 99 dataset such as time, Duration, Protocol, service, flag, source bytes, destination bytes of flows at each router to differentiate DDoS from flash crowd [5].

For network packet tracing purpose, we analyze the four different techniques which are depends on four theorems given below.

1. Network Flow

In a local network or a community network for a given router, we collect the network packets that have the same destination address as one network flow [4].

P***i*** = (***pi*** [1], ***pi*** [2]…….. ***pi*** [N]) (1)

Here, P***i*** represents N number of packets. According to our definition of flow, a router may have many network flows at any given point in time.

2. Flow Strength

For a network flow Pi, consider the length of the network flow be N ***(pi[N] >=1)*** .

We define the expectation of the flow as the flow strength of Pi. Flow strength represents the average packet rate of a network flow. If ***pi*** is a DDoS attack flow, then we also call A[***pi***] as attack strength [9].

$$A[\boldsymbol{pi}] = 1/N \sum_{i=0}^{N} \boldsymbol{pi}\,[\boldsymbol{n}] \quad (2)$$

3. Flow Fingerprint

For a given network flow Pi with length N used to represent the fingerprint as unified representation of Pi, which describes the similarities of different flows [4][5].

$$\boldsymbol{pi} = \{\boldsymbol{pi}'\,[\boldsymbol{1}], \boldsymbol{pi}'\,[\boldsymbol{2}] \ldots\ldots \boldsymbol{pi}'\,[\boldsymbol{N}]\}$$

$$\boldsymbol{pi} = \{\boldsymbol{pi}'[\boldsymbol{1}]/\mathbf{N*A}[\boldsymbol{pi}]* \boldsymbol{pi}'[\boldsymbol{2}]/\mathbf{N*A}[\boldsymbol{pi}]\ldots\ldots. \boldsymbol{pi}'[\boldsymbol{N}]/\mathbf{N*A}[\boldsymbol{pi}]\}$$

$$\boldsymbol{pi} = \boldsymbol{N} *\mathbf{*A}[\boldsymbol{pi}]* \boldsymbol{pi}$$

$$\boldsymbol{RESpipj} = \boldsymbol{1/N} \sum_{n=0}^{N} \boldsymbol{pi}\,[\boldsymbol{n}] * \boldsymbol{pj}\,[\boldsymbol{n}]$$

$$\boldsymbol{RESpipj}[\boldsymbol{K}] = \boldsymbol{1/N} \sum_{n=0}^{N} \boldsymbol{pi}\,[\boldsymbol{n}] * \boldsymbol{pj}\,[\boldsymbol{n}+\boldsymbol{k}] \quad (3)$$

4. Flow Correlation Coefficient

Let $\boldsymbol{pi}$and $\boldsymbol{pj}$ and ($\boldsymbol{i \# j}$) be two network flows with the same length N. We define the correlation coefficient as,

$\boldsymbol{FCCpipj[K]}$=

$\boldsymbol{RESpipj[K]}$/1/N[$\sum_{n=1}^{N-1} \boldsymbol{pi}2[n] * \sum_{n=1}^{N-1} \boldsymbol{pj}2[n]]1/2$ (4)

B. FP-Growth algorithm for detection purpose:

1. Initialize I (no. of records) = 1

2. Scan each record of the rule pool set.

3. Find number of items (N), number of transactions (M)

4. Increment I by 1 and repeat step 2 until last record in the rule pool set.

5. Initialize k (number of item set) =1

6. Find frequent item set Lk from Ck of all candidate item sets

   Lk is data record in collected data set and Ck is data record in KDD data set

   Scan D and count each item set in Ck,

   If count is greater than minimum support, then it is frequent

7. Form Ck+1 from Lk; k = k + 1

   Join Lk-1 item set with itself to get the new candidate item sets, If found a non-frequent subset then remove that subset.

8. Check all rules with test dataset

9. Repeat step 6 and step 9 until Ck is empty

V. Result Analysis

A. Implementation

The experiments of this novel proposed system are performed by using weka tool system and KDD CUP 99 dataset. It contains four modules. At the beginning we have to capture packets from different networks and store it on any text file. Then we have to generate the rules based on KDD CUP 99 dataset. After that we have to trace the DDoS attack by using all theorem mentioned above based on similarities. Then we have to calculate detection ratio based on flow correlation coefficient by using FP-Growth algorithm so that we can classify DDoS attack in different types. At the end we will analyse the result expecting increase in rate of detection of DDoS from given flash event.

B. Dataset

The experiments are performed by using International Knowledge Discovery Dataset. The KDD CUP 99 dataset is publicly available and considered as a benchmark dataset for testing of various detection algorithms [5]. By using KDD CUP 99 dataset, rather than inserting the attack packets into the normal traces, the labeled attack samples which are obtained by passive monitoring [6].

The KDD CUP 99 datasets consist of two types of dataset: training dataset and testing dataset. Each record of the training data is labeled as either anomalous or normal, which denotes a specific kind of attack. The training dataset
contains a total 22 types of attacks and in the testing dataset, 395 dataset has contain additional 15 types of attacks[7].
As we are detecting, sources of DDoS attacks (Smurf, fraggle, Neptune, Teardrop and Ping of Death). After elaborating labeled dataset, it has been found that total number of 41 attributes provides the specifications of the received packets.

For this experiment, by using different attributes of packet flows such as time, duration, protocol, service, flag, source bytes, destination bytes at each router to differentiate DDoS from flash crowd.

C. Result set

We are increasing the rate of differentiating DDoS from flash event using flow correlation coefficient compared to existing system.The performance of the network is evaluated in terms of the some metrics traceback time, DDoS detection Ratio based on flow correlation coefficient and throughput based on flow correlation coefficient. The graph of comparing performance of the conventional system and proposed system is given below in Fig.3. As the rate of detection of DDoS from flash event will be increased in the proposed approach of FCC system.

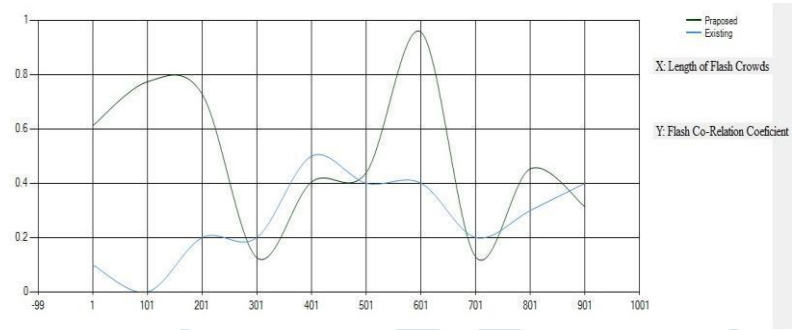


Fig.3. Flow correlation coefficient against network flow

Graph in Fig.4.shows relationship between length of simulated flash crowd and flow correlation coefficient. This indicates that the flow correlation coefficient decreases if the attack flows come from different network distributions.

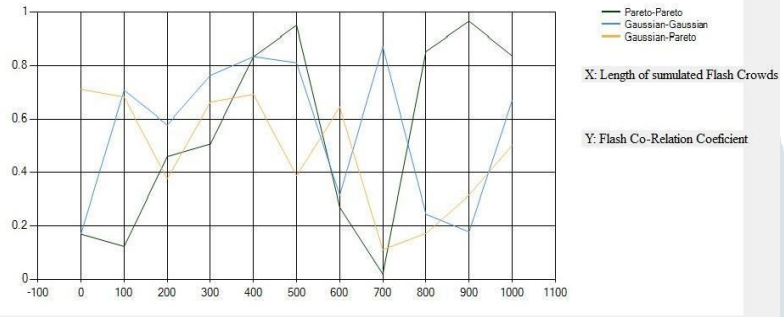


Fig.4.Flow correlation coefficient against length of simulated flash crowd

VI. Conclusion and Future Work

In this paper, we proposed an enhanced version of FCC System which is an effective and efficient detection and tracing mechanism based on flow correlation coefficient. The proposed method does not need any marking on packets and also any updating of routing software; hence it acts as an independent software module. It also reduces the problem of differentiating the flash crowd i.e. legitimate flow from DDoS attack. From this mechanism, it is proved that by combining parameters from KDD CUP 99 dataset such as time, duration, protocol, service, flag, source bytes, destination bytes of flows at each router, the DDoS attack (malicious flow) can bedistinguished from flash crowd, so that there is no probability of rising false alarm. Also, the proposed system can easily detect the actual sources of attack in time and increases effectiveness. In future, it is efficient to apply genetic algorithm to detect all network attacks globally to make network more secure.

Acknowledgement

We authors would like to thanks to Shui Yu; Thapngam, T.; Jianwen Liu; Su Wei; Wanlei Zhou for sharing their valuable knowledge.

References

[1]. Shui Yu; Thapngam, T.; Jianwen Liu; Su Wei; Wanlei Zhou, "Discriminating DDoS Flows from Flash Crowds Using Information Distance," Network and System Security, 2009. NSS '09. Third International Conference on 19-21 Oct. 2009

[2]. Ke Li; Wanlei Zhou; Ping Li; Jing Hai; Jianwen Liu, "Distinguishing DDoS Attacks from Flash Crowds Using Probability Metrics," Network and System Security, 2009. NSS '09. Third International Conference on 19-21 Oct. 2009.

[3]. Thapngam, T.; Shui Yu; Wanlei Zhou; Beliakov, G., "Discriminating DDoS attack traffic from flash crowd through packet arrival patterns," Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on 10-15 April 2011

[4]. Shui Yu; Wanlei Zhou; Weijia Jia; Song Guo; Yong Xiang; Feilong Tang, "Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient," Parallel and Distributed Systems, IEEE Transactions on June 2012

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

[6]. Kaur, G.; Varma, S.; Jain, A., "A novel statistical technique for detection of DDoS attacks in KDD dataset," *ContemporaryComputing (IC3), 2013 Sixth International Conference on* , vol., no., pp.393,398, 8-10 Aug. 2013

[7]. M. Bailey, E. Cooke, F. Jahanian, Y. Xu, and M. Karir, "A Survey of Botnet Technology and Defenses," Proc. Cybersecurity Applications and Technology Conf. for Homeland Security, 2009.

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

[9]. G. Oikonomou and J. Mirkovic, "Modeling Human Behavior for Defense against Flash-Crowd Attacks," Proc. IEEE Int'l Conf. Comm., 2009.

[10]. J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash Crowds and Denial of Service Attacks: Characterization andImplications for CDNs and Web Sites," Proc. 11th Int'l Conf. World Wide Web (WWW), pp. 252-262, 2002

[11]. V.L.L. Thing, M. Sloman, and N. Dulay, "A Survey of Bots Used for Distributed Denial of Service Attacks," Proc. SEC, pp.229-240, 2007.C. Riess and E. Angelopoulou, "Scene illumination as an indicator of image manipulation," Inf. Hiding, vol.6387, pp. 66–80, 2010.

[12]. Srikanth Kandula, Dina Katabi, Matthias Jacob, Arthur Berger., "Botz4sale:Surviving Organized Ddos Attacks That Mimic Flash Crowds.," IEEE Trans. Dependable Secure Computing, vol. 4, no. 1, pp. 56-70, Jan.-Mar. 2007.

[13]. G. Carl, G. Kesidis, R. Brooks, and S. Rai, "Denial-of-Service Attack-Detection Techniques," IEEE Internet Computing, vol. 10, no. 1, pp. 82-89, Jan./Feb. 2006.

[14]. 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. 1137-1151, 2006.

[15]. C. Patrikakis, M. Masikos, and O. Zouraraki, "Distributed Denial of Service Attacks," The Internet Protocol J., vol. 7 4, pp. 13-35, 2004