International Journal of Latest Engineering Research and Applications (IJLERA) ISSN: 2455-7137

Volume – 02, Issue – 10, October – 2017, PP – 01-06

Discriminating DDoS Attacks from Flash Crowds in IPv6 networks using Entropy Variations and Sibson distance metric

HeyShanthiniPandiyaKumari.S[1], Rajitha Nair.P[2]

[1](*Department of Computer Science &Engineering, New Horizon College of Engineering, Banglore, India)*

[2](*Department of Computer Science &Engineering, New Horizon College of Engineering, Banglore, India)*

Abstract: One of the major threats to network security is Distributed Denial of Service (DDoS) attacks. DDoS attack causes service disruptions on the victim system by sending large number of packets destined to victim. Lots of mechanisms such as packet marking and packet logging are proposed for identifying the attack sources. Entropy variation based trace back approach is better than other techniques in terms of scalability, storage space and operation overload. On the other hand, the entropy variation metric cannot differentiate legitimate flash crowd flows from DDoS attack flows since both involve sending large number of packets to the victim. To overcome this problem, in this paper, a hybrid metric of entropy variation and Sibson distance to perform the traceback process with flash crowd discrimination has been proposed. Using experiments, it has verified that the accuracy of the traceback process is improved significantly by the proposed method.

Keywords: DDoS attack, Entropy, Packet marking, Flash crowd, IPv6, Sibson distance.

1. Introduction

Cyber security is a major concern in the current Internet world. New threats are emerging day by day. Among them, Distributed Denial-of- Service (DDoS) attacks pose critical threats to the Internet1. DDoS attacks involve bombarding the victim networks with a high volume of attack packets originating from a large number of machines with the intention of causing service disruptions on the victim by exhausting the bandwidth or system resources. The attack takes place simultaneously from large number of systems.

Identifying the attack source of is a difficult problem because of the stateless and anonymous nature of Internet. The source may be a zombie, reflector or a final link in a chain. The problem of identifying the source of offending packets is called the IP trace back problem. A number of IP trace back approaches have been suggested to identify attackers. Some of the major methods are the probabilistic packet marking (PPM), the deterministic packet marking (DPM), packet logging and entropy variation [1]. Entropy variation based trace back approach is a reactive mechanism and is better than the available PPM and DPM methods in terms of scalability, operation overload and storage space. But this method may treat legitimate flash crowd as a DDoS attack. Flash crowd is the sudden surging of packets from legitimate users.

The major contributions of this work are as follows:

- The proposed method can effectively distinguish flash crowds from DDoS attacks during the trace back process.
- The discrimination process can be performed independently by any router. They do not need cooperation from other routers.
- The proposed method introduces almost no additional traffic into the network for solving the trace back problem.
- No need to create infrastructure or change the router configuration. The proposed algorithm can be integrated with the router as an additional module.
- The proposed method is scalable and practical.

2. Related Work

Yu et al.[2], proposed a trace back method using entropy variation. During non-attack periods, routers are required to observe and record entropy variations of local flows. Once a DDoS attack has been identified, the victim first identifies which of its upstream routers are in the attack tree based on the flow entropy variations it has accumulated, and then submits requests to the related immediate upstream routers and so on until the source is identified.

Baba and Matsuda.[3] presented a hop-by-hop packet logging approach for detecting DDoS attacks in IPv6 networks. The tracer or router, logs incoming packet information named, the packet feature, in a buffer memory called packet information area. It then replaces the packet's data link-level identifier with its interface

identifier and forwards the packet. Once a DDoS attack is detected, the tracer identifies adjacent node of an attack packet by searching in the packet information area. Savage et al.[4] introduced the probabilistic packet marking scheme for IP trace back. Goodrich.[5] proposed the randomize-and-link approach to implement IP trace back based on the probabilistic packet marking mechanism. According to his method, every router X calculates a checksum value, C = C (M_x), named as cord from its unique message M_x (e.g., IP address) and fragments M_x into several pieces, M_0, M_1 , . . .M_i. The router then marks the packets probabilistically with b bits where b_i = [i, C, M_i] fragments M_x into several pieces, M_0, M_1 , . . .M_i. The router then marks the packets probabilistically with b bits where b_i= [i, C, M_i].

3. Proposed System

In this paper, an entropy-based trace back system that can efficiently discriminate DDoS attack flows from legitimate flash crowd flows has proposed. The entropy metric to monitor the randomness of flows at the router was used. If there is an abnormal change of entropy, then some suspicious flows are present in the system. These flows may be DDoS attack flows or flash crowd flows. Since both DDoS and flash crowd flows involve large number of packets, the entropy metric is sensitive to both and treats both the flows as attack flows. These flows considered as suspicious flows and categorize them into DDoS attack flows and flash crowd flows. On the other hand, flash crowd flows are generated by randomly distributed users who generate packets at different rates. Hence the flow similarity among flash crowd flows will be very less. To find the flow similarity we use the information theoretic parameter, Sibson distance. The Sibson distance metric is already used for flash crowd discrimination in DDoS detection context [6]. In contrast, for our IP trace back problem, at each node on the attack path we have to perform the flash crowd discrimination and hence involving two cooperating routers for performing flash crowd discrimination at each node on the attack path is difficult to achieve and also will slow down the trace back process. But by our proposed method individual nodes can perform the discrimination process on their own without requiring any cooperating node.

3.1 System Model

In the proposed system, each trace back enabled router will have three modules for performing the trace back process. The modules are flow observer, suspicious flow identifier and flow classifier as depicted in Fig.1.

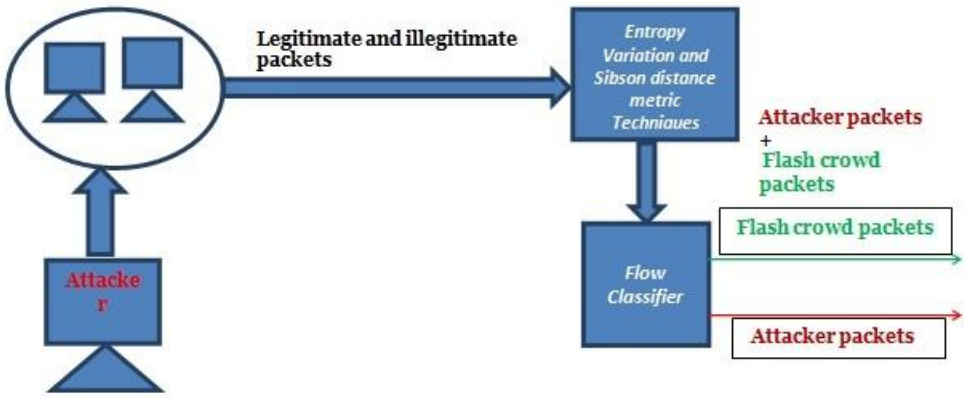

Fig.1. System model

3.1.1. Flow Observer

The flow observer module monitors the flows that are passing through the router in time intervals of ΔT [7]. The packets that are passing through a router R during time interval ΔT are categorized into different flows. The packets that are from the same upstream router and destined to the same system are grouped into a single flow. Thus, at a router R, a flow f (u, d) is identified by the upstream router u of its packets and by the destination address d of its packets. Let n be the number of flows at router R during the time period ΔT. Let $|f_i|$ denotes the number of packets in a particular flow fi during time interval ΔT. The total number of packets that passed through router R during ΔT is given by $\sum|f_i|$. For the time interval ΔT, the probability of each flow f_i can be calculated by using (1). Then, the entropy of i = 1 fi flows at the router R during ΔT, can be calculated using the probability values as (2).

In the routers the mean and standard variation of entropy values are maintained. The difference between the current entropy value, H and the mean is calculated. We call this difference, |H-M| as entropy

variation. If the entropy variation is greater than the standard deviation σ, i.e. $|H-M|>\sigma$ then the change in the randomness of the system is abnormal. Here the standard deviation value σ is used as the threshold. If there are no such events then, $|H-M|\leq \sigma$ holds with higher probability.

Algorithm for flow observer:

1. Identify flows $f_1, f_2, \ldots f_n$;
2. During time period ΔT, for each flow f_i calculate count of packets $|f_i|$;
3. After Δ T is over, calculate probability of each flow $f_i$as follows,

(1): $Pi=\frac{|fi|}{\sum_{i=1}^{n}|fi|}$

4. Compute entropy of flows,

$$(2): H = -\sum_{i=1}^{n} p_i.\log p_i$$

5. If$|H- M|\leq \sigma$,
   a) progress mean and standard variation of entropy as follows,

$$(3): M_t = \sum_{i=1}^{k} \alpha_i.M_{t-i}$$

$$(4): \sigma_t = \sum_{i=1}^{k} \alpha_i.\sigma_{t-i}$$

   b) Iterate the process

3.1.2. Suspicious Flow Identifier

In the case of drastic change of entropy, the suspicious flow identifier is invoked to identify the flows responsible for the extensive change of entropy [8]. Let F be the set of flows at the router R during ΔT and S be the set of suspicious flows detected by the suspicious flow identifier. First the flows in F are sorted in descending order f_1', f_2' … f_n', then the flow with the maximum packet count, f_1' is added to the suspicious flow set S and is removed from the set of flows, F. The entropy is calculated for the flows in F and the entropy variation value is computed. If the newly computed entropy variation is still greater than the threshold σ, then the next flow with the maximum packet count is added to the suspicious flow set S and the process is repeated until the entropy variation becomes normal. On the other hand, if the entropy variation is less than the threshold, then the flows in the suspicious flow set are finalized as the flows responsible for the abnormal entropy change and are given to the flow classifier for differentiating flash crowd flows from DDoS attack flows.

Algorithm for Suspicious Flow Identifier:

1. Sort the flows f1, f2, …fn in descending order f1', f2'…fn' ;
2. For i = 1 to n,
   {
   a) Add fi' to suspicious flow set S;
   b) Calculate H(F – fi);
   c) If |H(F - fi) - M|> σ, then continue;
   d) Else break;
   }

3.1.3. Flow Classifier

The flow classifier module categorizes the suspicious flows given to it as input into flash crowd flows and attack flows.

Algorithm for Suspicious Classifier:

1. For each flow $f_i$in S,
   {

a) Sample the count of packets for size Lin times lotsttoobtainx_1, x_2, …$x_L$in slots t_1, t_2, …$t_L$respectively;

b) Calculate probability distribution,

$$\frac{x_k^i}{\sum_{x=1}^{L} x^i} \quad \text{for k=1,2 ,…L}$$

}

2. For each pair of flows (f_i,f_j) in S,

{

a) Calculate Sibson distance,

$$D_S = D[Pi(Pi+ Pj)] + D[Pj(Pi+ Pj)]$$

b) If $D_S \leq \delta$, then add f_i and f_j to attack flow set A;

}

3.2 Trace back scheme

Trace back requests are sent to the upstream routers of the flows classified as attack flows by the flow classifier. The upstream routers perform the same process and this scheme continues until the edge routers of the zombies are identified. Then the edge routers identify the zombies and deliver the confirmed zombies information to the victim. Thus, the trace back process is performed in a parallel and distributed manner.

4. Experimental Results

To analyze the effectiveness of our proposed work, we have first evaluated the effectiveness of the metrics used in this work. We tested the stability of the entropy metric under non-attack cases. We simulated the Poisson distribution and measured the entropy for different number of flows [9]. The entropy value is computed for different number of flows and is plotted in a graph. The results are shown in the Fig.2, which infers the entropy value smoothly increases with the increase in number of flows and is stable against traffic fluctuations.

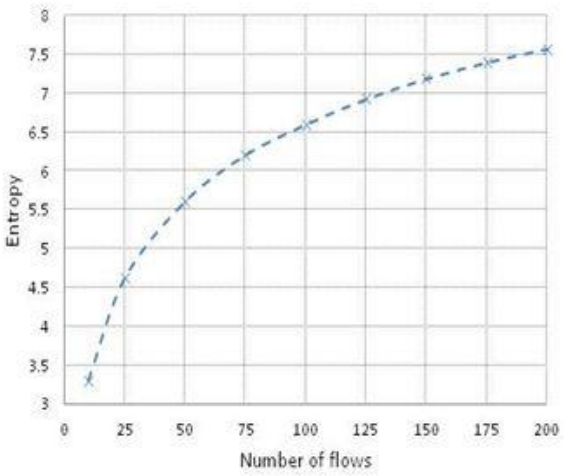

Fig.2. Entropy against normal flows

Next, efficiency of entropy metric under attack caseshas evaluated and simulation done for 100 traffic flows using Poisson distribution. Among them one is attack flow and the remaining are normal flows. The packet rates of normal flows are maintained at the same level. From Fig.3, we can see that the entropy value drops almost linearly with the increase in attack strength.

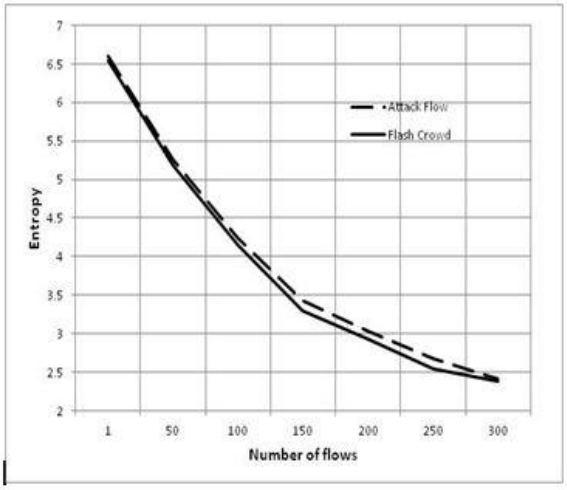

Fig.3. Entropy against attack flow

The efficiency of Sibson distance metric is then evaluated. Two flows are simulated with Poisson distribution and the distance between the mis measured using Sibson distance metric. FromFig.4, we can infer that the Sibson distance metric is stable and smooth for different distributions.

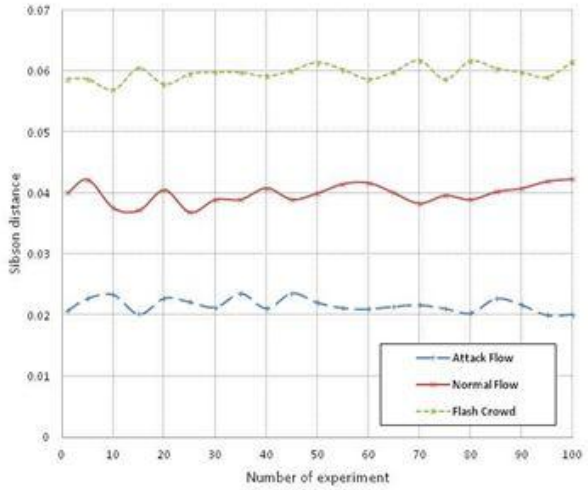

Fig.4. Sibson distance between normal, attack and flash crowd flows

5. Conclusion

In this paper, an entropy- based trace back system for distributed denial-of-service attacks that can differentiate legitimate flash crowd traffic from attack traffic by using the Sibson distance metric has been proposed. Sibson distance is used to find out the similarity between packet flows. We have analyzed the effectiveness of the proposed system using simulations and the results are discussed. By this method used in this work, flash crowd flows can be efficiently discriminated from DDoS attack flows even when both occur simultaneously in IPv6 network. The limitation of our work is that it cannot detect the low rate DDoS attacks. We encourage future works in this aspect.

References

[1]. L Feinstein, D Schnackenberg R Balupari and D Kindred, Statistical Approaches to DDoS Attack Detection and Response. In Proceedings of the *DARPA Information Survivability Conference and Exposition, vol.1, IEEECS Press*, 22-24 April 2003, pp. 303–314.

[2]. S Yu, W Zhou, R Doss, and W Jia, Trace back of DDoS Attacks Using Entropy Variations. *IEEE Trans. Parallel and Distributed Systems*, vol. 22, no. 3, pp. 412-425, Mar. 2011.

[3]. T Baba and s Matsuda, Tracing Network Attacks to Their Sources. *IEEE Internet Computing*, vol.6, no.2, pp.20-26, Mar. 2002.

[4]. S Savage, D Wetherall, A Karlin, and T Anderson, Network Support for IP Trace back. *IEEE/ACM Trans. Networking*, vol. 9, no. 3, pp. 226-237, June 2001.

[5]. MT Goodrich, Probabilistic Packet Marking for Large-Scale IP Trace back. *IEEE/ACM Trans. Networking*, vol.16, no.1, pp. 15-24, Feb. 2008.

[6]. P. Jain, J. Jain, and Z. Gupta, Mitigation of Denial of Service (DoS) Attack. *International Journal of Computational Engineering & Management (IJCEM)*, vol. 11, pp. 38-44, January 2011

[7]. Craig Labovitz, *The Internet goes to war*. Arbor Networks, Dec. 2010.

[8]. S Yu, T Thapngam, J Liu, S Wei and W Zhou, Discriminating DDoS Flows from Flash Crowds Using Information Distance. in Proceedings of the 3rd *IEEE International Conference on Network and System Security (NSS'09),* 18-21 October 2009.

[9]. A Belenky and N Ansari, IP Trace back with Deterministic Packet Marking. *IEEE Comm. Letters*, vol.7, no.4, pp.162-164, Apr. 2003

