Dynamic Detection and Protection Mechanism against Distributed Denial of Service Attacks using Fuzzy Logic

M. Parameswari and Dr. S. Sukumaran
Associate Professor, Department of Computer Science
Erode Arts & Science College, Erode-638 009, Tamilnadu, INDIA

Abstract
DDoS (Distributed Denial of Service) is the attack to pollute the network. The attacker creates a large amount of packet to the particular system. The packets are sending by using the compromised computers. It is an effort to make a device or network resource engaged to its intended users. This paper describes training the DDoS attack detection system to recognize possible attacks on a system. The objective of this paper is to develop practical and scalable mechanisms to identify the DoS or DDoS attack sources. This paper includes implementation of traceback system in Single ISP domain, Multi ISP domain, and stateless Internet architecture and enhancing security features. The proposed system is offered to trace-back method based on entropy variations and fuzzy logic method. The system works on different between the normal data flow and attacking data flow. It has a number of advantages such as avoid packet pollution, no need to change the current routing software. Through empirical evaluation it is confirmed that the detection can be completed within improved real time limits and that by using fuzzy estimators instead of crisp statistical descriptors it can be avoided the shortcomings posed by hypothesis on the model distribution of the traffic.

Keyword: DDoS Attacks, ICMP (Internet Control Message Protocol), CP (Cumulative path), Probabilistic Packet Marking (PPM) and Deterministic Packet Marking (DPM)

INTRODUCTION
Distributed Denial of Service attacks in the internet to pollute the network environment. The attackers generate a huge amount of requests to victims through compromised computers (zombies), with the aim of denying normal service. DDoS attacks are targeted to identify the victim resources, such as bandwidth, memory, and buffer. The attacker needs to attack the victim attacker to establish a large number of request send to the victim. So the attacker needs a multiple system to produce a request. The attacker discovers a large number of vulnerable systems in the same network. These discoverable systems are called as Zombies. The attacker installs the new program to the discovered system for attacking the victim in the network. The vulnerable system is identified by out dated antivirus using system or secure less system [3].
IP traceback is a name given to any method for reliably determining the origin of a packet on the Internet. Due to the trusting nature of the IP protocol, the source IP address of a packet is not authenticated. As a result, the source address in an IP packet can be falsified (IP address spoofing) allowing for Denial of Service attacks or one-way attacks (where the response from the victim host is so well known that return packets need not be received to continue the attack). The problem of finding the source of a packet is called the IP traceback problem. IP Traceback is a critical ability for identifying sources of attacks and instituting protection measures for the Internet. Most existing approaches to this problem have been tailored toward Denial of Service attack detection. Such solutions require high numbers of packets to converge on the attack path(s).
The network does not have effective method to locate the attacker such as hash Based IP Trace back, algebraic approach to IP trace back and Enhanced ICMP trace back as Cumulative Path. The attacker used this advantage and attacks the network. The IP traceback scheme is presently used to identify the attackers. IP traceback schemes are considered successful if they can identify the zombies from which the DDoS attack packets entered the internet and that trace back scheme is easily identified the attacker. In the existing system two types of packet marking are used viz, Probabilistic Packet Marking (PPM) and Deterministic Packet Marking (DPM).
Both of these packets marking require router by the marking to the individual packets. The PPM packet marking can only operate in a small network. So, the detector or protector does not trace out the attacker. The network is very small it does not identify the attacker location, when that person present away from the network and the method is called DPM. The packet marking is require to all routers are updated by the packet marking technique. This DPM only 25 spare bits is available to the single IP packet with the scalability of the DPM is very huge problem [22]. IP traceback mechanism should be independent of the packet pollution and different types of attack patterns.
DPM mechanism poses the extraordinary challenge on storage for packet. Once the attack started is found the different between the attack flow and the normal flow. In this IP traceback mechanism the entropy variation or different flows of the network and there is no packet marking are used. It avoids the problem of the packet marking method. The packets are passing through the router into flows, which are defined by the router where the packet came from the destination address of the packet. During the non attack period the router observes the normal flows. When the DDoS attacks occur the victim send request to the upstream routers, then the router identifies the attack flow path and the process was repeated again because to find the number of attackers in the network.

PACKET MARKING SYSTEM
In packet marking method traceback data is inserted into the IP packet by the routers on the path to the destination node.
The packets are marked as they traverse routers through the internet either probabilistically or deterministically. Packet marking information is stored in the identification field of the IP header. The routers mark the packet with either the router’s IP address or the edges of the path that the packet traversed to reach the Destination / Victim. The victim uses the information in the marked packets to trace an attack back to its source. For the first alternative, marking packets with the router’s IP address, analysis shows that in order to gain the correct attack path with 95% accuracy as many as 294, 000 packets are required. The second approach, edge marking, requires that the two nodes that make up an edge mark the path with their IP addresses along with the distance between them. This approach would require more state information in each packet than simple node marking but would converge much faster. Types of packet marking are Probabilistic Packet Marking (PPM) and Deterministic Packet Marking DPM. Deterministic Packet Marking approach focuses on determining the source of the attack packet and is not concerned with the actual path traversed by the attack packet, while the Probabilistic Packet Marking approach focuses on reconstructing the entire attack path through which the malicious packets have traversed. The DPM mark the spare space with initial router information. The receiver can identify the source location of the packets when it has sufficient information of the marks. The big problem in DPM is changing the current routing software. The problem of the DPM is due to the requirement of large number of packet reconstruction. Every routers mark the own IP address of the packets header. The node sampling algorithm is used to record the router address of the packet. There large number of marked packets can do the reconstruction of the attack path. The edge sampling algorithm is used to mark the start router address to the end router address of the attack path and the distance is fixed in between the two ends [24]. It provider based deterministic packet marking models-to characterize DDoS attack streams-used to make filtering near the victim more effective. The proposed FUZZY BASED DETECTION mechanism has a rate control scheme that protects destination domains by limiting the amount of traffic during an attack, while leaving a large percentage of legitimate traffic unaffected. The above features enable service providers to offer enhanced security protection against DDoS attacks as a value-added service to their customers, hence offer positive incentives for them to deploy the proposed models. The Scheme proposes and evaluates two providers-based packet marking models: Source-End Provider Marking and Source and Destination-End Provider Marking. Both models are based on deterministic packet marking, and aim to give the victim’s provider stable and secure information about the path incoming traffic streams. The Scheme also proposes a rate control system that protects destination domains by limiting the amount of traffic during an attack, while leaving a large percentage of legitimate traffic unaffected. These facilitate providers to offer increased protection to their customers as a value-added service, improving the available throughput for legitimate users during such attacks. Path Identifier (PI) a packet marking approach in which a path fingerprint is embedded in each packet, enabling a victim to identify packets traversing the same paths through the Internet on a per packet basis, regardless of source IP address spoofing. In this approach an identifier is embedded in each packet based on the router path that a packet traverses. The victim needs only to classify a single packet as malicious to be able to filter out all subsequent packets with the same marking. What makes this possible is that our packet marking is a per-packet deterministic mechanism: each packet traveling along the same path carries the same identifier. This allows the victim to take a proactive role in defending against a DDoS attack by using the PI mark to filter out packets matching the attackers’ identifiers on a per packet basis. Most marking schemes are probabilistic in nature, in which the victim needs to collect a large number of packets to reconstruct the path. In this approach, a path identifier fits within a single packet so that the victim can immediately filter traffic after receiving just one attack packet. The scheme is extremely lightweight, both on the routers for marking, and on the victims for decoding. The router marking is also robust to the presence of legacy routers and shows strong incremental deployment properties.

**TRACE BACK APPROACH USING FUZZY SYSTEM**

The proposed method is better than the PPM and DPM packet marking, because there is no packet marking technique and no need to change the current routing software. The proposed mechanism as follows:

- Scalability (the size of attack network that can be handled)
- Storage (the storage space on routers or victims to conduct IP traceback)
- Traceback time (the overall time we need from the start time until the end of tracing process)
- The operation workload (the operations on possible routers or victims).

**Variation**

The different types of data flows are used to find the attack. The difference between the normal flow and attacking flow is called as Entropy Variation. The process as follows:

- Difference between the non-attack and attack period is called Entropy Variation. It is true that the network traffic for a router may dynamically change a lot from peak to off-peak service times. However, this kind of change lasts for a relatively long time interval, e.g., at least at the level of minutes. These changes can be brought town into seconds.
- The number of attack packets is at least an order of magnitude higher than that of normal flows. During a DDoS flooding attack, the number of attack packets increases dramatically and the attack packets are generated by thousands of zombies.
- The number of attack packets is much higher than that of legitimate flows. Therefore, this assumption is reasonable. Of course, for the non-flooding attacks.
- Only one DDoS attack is ongoing at a given time. It could be true that a number of attacks are ongoing.
concurrently in the Internet, the attack paths may overlap as well, but we only consider the one attack scenario to make it simple and clear.

The number of flows for a given router is stable at both the attack cases and nonattack cases.

In this DDoS detection and protection system two types of algorithm, when the entropy variation is different the trace back process is started.

The special flow monitoring algorithm is running at the non-attack period, accumulating information from normal network flows, and progressing the mean and the standard variation of flows. The progressing suspends when a DDoS attack is ongoing. Once a DDoS attack has been confirmed by any of the existing DDoS detection algorithms, then the victim starts the IP traceback algorithm. This continuously monitoring the http request from the internet. When the request is coming, it identifies the IP address and stored in cache and start counting the request from the same IP address and also maintain the timer. More than 20 requests within one second from same IP address is considered as DDoS attack. Then the IP address is blocked for certain time periods prevention that means the suspicious IP address is blocked for certain time periods. That's like a monitoring process are very effective to monitoring the network and this monitoring is used to find out the attacker easily. The monitoring process is used to pushback when the attack is occurring. The traceback process is find out the attacker from the network when the attack traffic is present in the network.

Initiate the local parameter X, U, D. U={u_i} be set of upstream routers, D={d_j} be set a destination address of the packet and the victim is V.

The attack flow as, \( f_i = u_i, V, i=1,2,...n \). That’s like a data flow as \( f_1, f_2, \ldots, f_n \).

For i=1 to n

- Calculate H(F(x))
- If H(F(x))<x
- Upstream router of \( f_i \) to set A
- else break;
- end if;
- end for;

The sample algorithm for tracing the source of the attack can be as follows:

 FixVal:= allocated_mamory(G.tot_space);
 For (i := 0; i <= FixVal ; i :=i+1) 
    For (j := 0; j <= FixVal ; j :=j+1) 
     Routervall[i][j] := 0;
     end For
     Flag:=0;
     end For
     Routerval [ ][ ] := G.subspace;
     For(i := 0; i <= FixVal ; i :=i+1) 
     For(j := 0; j <= FixVal ; j :=j+1) 
     If(Routervall[i][j]==space_A) 
      Path_set[i][j] := Search_path(flag, i, j);
      Flag:=Flag+2;
     Else If(Routervall[i][j]==space_B)
      Path_set[i][j] := Search_path(flag, i, j);
      Flag:=Flag+2;
     Else If(Routervall[i][j]==space_C)
      Path_set[i][j] := Search_path(flag, i, j);
      Flag:=Flag+2;
     Else If(Routervall[i][j]==space_D)
      Path_set[i][j] := Search_path(flag, i, j);
      Flag:=Flag+2;
     End If
     End If
     Return Path_set;
     Break;
     end For
     end For

**Fuzzy Clustering Model**

Further the proposed work has implemented the Fuzzy based clustering technique for a detection system to enhance the purity. The fundamental insufficiency of clustering is used to initiate the value of K, a number which computes the count of clusters to be formed. This classification is done by applying the fuzzy clustering technique to deal with two segregations as normal and abnormal, in which K is assigned to two values. By using this feature of DDoS attack as a criterion, this work considers the partition model. It can consider the volume of attack packets in such a way that it is less than that of normal packets. The investigation shows the inherent nature of these attack packets and provides a clear picture of the factors that signify the abnormality. Subsequent to these partitions being made, every field in the normal and abnormal clusters is investigated to recognize its characteristics. This knowledge helps to distinguish the regular from the irregular ones.

Fuzzy based clustering technique is measured for the minimization of the following objective function, with respect to ‘Q’, a fuzzy partition of the data traffic and to ‘R’, a set of ‘L’ prototypes as shown in Equation 1.

\[
D_Y (Q, R) = \sum_{z=1}^{L} \sum_{y=1}^{Q} Q_{y} \cdot \| N_y - R_y \|_2
\]

where ‘Q’ is a membership of ‘Ny’ in the cluster ‘Dy’ between 0 and 1; where ‘o’ is any real number greater than 1; ‘Ny’ is the Yth d-dimensional measured input data; ‘DDy’ is the centre of fuzzy cluster ‘Y’; \( f_{e_{yz}} = \| N_y - R_y \|_2 \) is the \( z \)th point and yt cluster centre are the measured the Euclidean distance and X(1, ∞) is a weighting exponent.

There are two necessary conditions for ‘C’ to reach a minimum as shown in Equations (2) and (3).

\[
D = \sum_{y=1}^{Y} N_y \cdot Q_{yz} \cdot (x-1)
\]

\[
N_y = \sum_{z=1}^{L} Q_{yz}
\]

\[
Q_{yz} = \frac{1}{\sum_{x=1}^{L} \left( \frac{f_{e_{yz}}}{f_{e_{z}}^{x}} \right)^{2}}
\]
The measures in this iteration will stop when \( \max_{x,z} \frac{Q_{yz}}{Q_{yz}} < \epsilon \), where \( \epsilon \) is a termination measure between 0 and 1. The maximum number of iteration routines shown in Equation 4.4 can be used as a termination measure. An illustration of the pseudo-code and the flow chart for the aforementioned process are shown in Fig. 5.

**Fuzzy based Dynamic Detection System for DDoS**

Initialize Increment = 1
WHILE Increment < Total iterations
  FOR each DataPoint y
    Determine \( N_y \)
  ENDM FOR
  FOR each \( N_y \)
    Calculate Cluster Center
    \( DD_y = \sum Q^0_{yz} N_z \)
  ENDM FOR
  FOR each \( Q_{yz} \)
    Compute \( D_y \)
  ENDM FOR
  IF \( D_y > \text{ThresholdLevel} \)
    Determine new \( N_y \)
  ENDIF
  Add 1 to Increment
ENDWHILE

**PERFORMANCE RESULT AND DISCUSSION**

This study helps to analyze the packet information and filter it based on the available information. It feeds the information in the packet only once when it enters into the first router in the network. The computational burden and comparison of scalability with different techniques is shown below in Figure 1.

As a result, the fuzzy based detection technique stands best among the various other existing techniques. It utilizes the available path information for tracing the source system and hence the traceability is improved. It enables the router to reduce the overhead in packet forwarding and hence the tracing is easy. The performance comparison based on the number of routers traced is shown in Figure 1. The result shows that the performance of the other existing techniques reduces as the number of routers, when the packet crossed increases.

This detection system principle is capable of handling large scale attacks with several advantages as follows:

- Easy to detect the attacker with the single packet information
- Does not involve complex calculation
- Reduces the router overhead and network traffic.
- Easy to mitigate and prevent further attacks

The performance graph as in Figure 21 shows that the neuro-fuzzy based technique detection system has achieved the higher response time for the rate of detection both in the wired and wireless networks. The proposed neuro-fuzzy based clustering technique detection system is higher in performance than the traceback mechanism for DDoS attack detection method. The improvement of the response time is 20 to 35% of the wireless network of the detection in the network traffic data.

**Table 1:** Comparison of performance of DDoS attack detection

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Memory requirements at routers</th>
<th>Computational Burden</th>
<th>Scalability</th>
<th>Time Requirement to Traceback</th>
<th>No. of packet required for traceback</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPM</td>
<td>NIL</td>
<td>High</td>
<td>Good</td>
<td>Fair</td>
<td>High</td>
</tr>
<tr>
<td>PPM for IP Traceback</td>
<td>NIL</td>
<td>High</td>
<td>Good</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>ERPPM</td>
<td>NIL</td>
<td>High</td>
<td>Good</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Flexible Deterministic Packet</td>
<td>NIL</td>
<td>Medium</td>
<td>Good</td>
<td>Fair</td>
<td>Low</td>
</tr>
</tbody>
</table>
Due to the totally different nature of fuzzy based detection and other well known traceback schemes, involving packet marking or packet logging techniques, quantitative comparison of the various schemes is not possible. Hence in this section, we first present a qualitative comparison between fuzzy based detection and other well known traceback schemes. Success of any traceback scheme is determined by four key factors-computational overhead involved for packet marking, memory requirement for packet logging, scalability of the proposed scheme and the need for cooperation between other domains. The overhead of the fuzzy based detection presented here is very light; The fuzzy based detection scheme is also scalable. No Cooperation between different ISPs is required. Furthermore unlike PPM and SPIE, the scheme can be used to mitigate the effect of the attack while the attack is raging on.

CONCLUSION
In this paper a new method was proposed an effective and efficient method of time and accuracy based detection mechanism IP traceback scheme against DDoS attacks based on entropy variations using fuzzy logic. It is a fundamentally different traceback mechanism from the currently adopted packet marking strategies. Because of the vulnerability of the Internet, the packet marking mechanism suffers a number of serious drawbacks: lack of scalability; vulnerability to packet pollution from hackers and extraordinary challenge on storage space at victims or intermediate routers. On the other hand, the proposed method needs no marking on packets, and therefore, avoids the inherent shortcomings of packet marking mechanisms. It employs the features that are out of the control of hackers to conduct IP traceback.

Store the short-term information of flow entropy variations at routers. Once a DDoS attack has been identified by the victim via detection algorithms, the victim then initiates the pushback traceback procedure. The traceback algorithm first identifies its upstream routers where the attack flows came from, and then submits the traceback requests to the related upstream routers. This procedure continues until the most far away zombies are identified or when it reaches the discrimination limitation of DDoS attack flows. Compared with previous works, the proposed strategy can traceback fast in larger scale attack networks. It can traceback to the most far away zombies within 25 seconds in the worst case under the condition of thousands of zombies. Moreover, the proposed model can work as an independent software module with current routing software. This makes it a feasible and easy to be implemented solution for the current Internet.

REFERENCES


