Journal of Computer Science 8 (2): 216-221, 2012
ISSN 1549-3636

Geographical Division Traceback for Distributed Denial of Service

[1]Viswanathan, A., [2]V.P. Arunachalam and [3]S. Karthik
[1]Department of CSE, K.S.R. College of Engineering, Tiruchengode, Namakkal
[2]SNS College of Technology, Coimbatore, Tamilnadu, India
[3]Department of CSE, SNS College of Technology, Coimbatore, Tamilnadu, India

Abstract: Problem statement: Distributed Denial of Service (DDoS) was a serious threat to the internet world that denies the legitimate users from being access the internet by blocking the service. **Approach:** In this study, we proposed a novel approach, Geographical Division Traceback (GDT) for efficient IP traceback and DDoS defense methodology. DDoS attack was one of the most serious and threatening issue in the modern world web because of its notorious harmfulness and it causes the delay in the availability of services to the intended users. **Results:** Unless like a traditional traceback methodology, GDT proposes a quick mechanism to identify the attacker with the help of single packet which imposes very less computational overhead on the routers and also victim can avoid receiving data from the same machine in future. This mechanism for IP Traceback utilizes the geographical information for finding out the machine which was responsible for making the delay was proposed. The IP packet consists of the subspaces details in which the path denotes. It helps to make sure whether the packet travels in the network and falls within any one of the subspaces. The division of subspaces leads to the source of attack system. **Conclusion/Recommendations:** This method possesses several advantageous features such as easy traversing to the attacker and improves the efficiency of tracing the attacker system.

Key words: network security, distributed denial of service, IP traceback, packet marking, Geographical Division Traceback (GDT)

INTRODUCTION

Internet is highly used in the most of our day to day life applications. It is a very essential resource for every one of us. It is our duty to secure this important resource from all the threats. One way of ensuring internet security is making the internet service available to the end user all the time. But DDoS is an attack, which makes the web server incapable of providing normal services to the authenticated user. So the web server will not available even for the legitimate users. Hence it is very critical to identify a standard mechanism for IP traceback and defending against the DDoS attack. Traditional IP traceback mechanism will identify the attacker with the help of source address field of the IP header. But possibilities are there for the attacker to change the address resides in the source address field.

To overcome the above mentioned problem, many traceback mechanisms were proposed. One such principle was "Attack Diagnosis (AD) and Parallel Attack Diagnosis (PAD)" (Chen R *et al*., 2007). This AD and PAD principle determines the attacker and throttles the incoming traffic in the divide and conquers fashion. It applies a "divide and conquer" strategy to separate attacking hosts and filter their traffic. The AD/PAD integrates the concepts of pushback and packet marking (Al-Duwairi, 2006). AD/PAD's framework is in line with the ideal framework of DDoS mitigation schemes in which the attack detection module is placed at the victim end and the filtering module is placed close to the attack sources. It is capable of tracing back and mitigating attack traffic from multiple attackers simultaneously, thus enabling it to handle large scale attacks (Ghazali and Hassan, 2011). It performs both the IP traceback and defends against the attacker. But this approach adds additional overhead on the routers.

Another approach for DDoS attack mitigation and attack detection is the Directed Geographical Traceback (DGT) (Gao and Ansari, 2005). This DGT principle requires a direction field in the IP header that consists of 8 sub fields that denote 8 possible geographical

Corresponding author: Viswanathan, A., Department of CSE, K.S.R. College of Engineering, Tiruchengode, Namakkal

directions. The internet path length can't be more than 32 bits and is required to encode each direction. Therefore 40 bits are required for all directions. Therefore, when a router forwards a packet using DGT, it first decides the next hop then decreases TTL by 1 and adds 1 to the corresponding direction sub field. Irrespective of the source IP address which can be spoofed, victim can locate the relative location of the attacker from the direction field when a packet arrives towards it.

But the DGT approach has some limitations. This principle will not work well if the router has more than 8 interfaces.DGT cannot handle spoofed marking. Not only the DGT, spoofed marking cannot be handled by the IP traceback mechanisms. Hence a very efficient traceback and attack mitigation mechanism is required to overcome the limitations in the above mentioned two approaches. Hence we propose a fast convergence IP traceback mechanism called "Geographical Division Traceback" (GDT).

MATERIALS AND METHODS

The Distributed Denial of Service attack is a major threat to the modern network community. In the present generation, a DDoS attack poses more threat to large number of organizations. The reason is, the number of systems involved in accessing the internet is increasing day to day in a rapid manner. Due to this, the traffic and the information access become difficult. The preventive measures against this attack is also a major difficult task due to various reasons like increase in traffic, availability of latest technologies for packet transmission and increasing the usage of internet among the people. To overcome these drawbacks, we have four basic countermeasures against the attacks namely detection, mitigation, prevention and Traceback. Most of the researches are carried out mainly in the two areas namely Attack Traceback and Attack Mitigation. This study focuses mainly on the Traceback of IP addresses used for the information transmission and attack process. It will be carried out after an attack has been launched and it will prevent the forthcoming attacks to the system.

Attack Traceback addresses the problem of collecting information about individual packet forwarding agents and collating this data to obtain an approximate Internet router-level graph (attack tree rooted at the victim); whereby tracing the routing path that any packet has taken, provides sufficient basis for attack attribution (attack tree leaves). The Attack traceback is necessary for cleansing zombie attackers, while also being of critical forensic value to law enforcement. The major sources of attacks are due to the increasing in the accessing of the network resources by an outside unauthorized user. These users are from different geographical regions and different countries. This makes the traceback process difficult in the real time situation. The information transmitted from one router consists of the source address and the destination address along with the information content.

The advantage of our method is, it will split the entire network into various sub-networks helps to identify the attacker in an easy manner with the use of existing geographic information. The geographical information helps us to trace the system of source of attack resides in the network. In this study, we propose the assumptions should be made before preceding the traceback process followed by the method which is used for tracing the source of attack system. Finally the result is compared with the other existing traceback schemes.

Assumptions: We assume the entire network path as similar to the continents of the earth. We can split the earth into four equal pieces like dividing the major problem into sub problems using the divide and conquer problem solving principle. This idea leads to the system to follow the divide and conquer approach for identifying the system in the network. This method follows the geographical traceback approach for the victim identification process. This scheme brings the fact that the path from one node to another is more associated with their geographical locations. Our traceback mechanism is related to the geographical information mainly based on the direction of attack paths. We can detect the attack source even when one of the routers is quiet. This information is very useful for IP traceback since one can counteract accordingly to the address used by the victim machine.

Earlier traceback works count on the target's ISP to preserve against the DDoS attacks. As our method implants marking information in each single packet and it can be made used to discard the attack packets which have similar marking. Thus the victim can actively filter packets rather than inactively obtain traceback. The network and traffic are assumed as follows:

- An attacker may produce any number of attack packets
- Attacker may be conscious that they are being traced
- The attack packets may be small
- Routing is not circuitous
- A router recognizes its direction with respect to its adjacent routers in any of the coordinates involved in it
- Routing behavior may be unstable
- Router should be static in nature

Most of these assumptions are got from (Xiang *et al*., 2008) and (Savage *et al*., 2000). The first two assumptions are associated with the abilities of the attackers like Packets with spoofed source IP address that can be generated to decoy the operation of traceback. The third assumption makes us to implement mechanism that can deal with not only the flood based attack but also the single packet attack. The fourth is to exhibit the internet measurement and is significant to the design of the IP traceback mechanism. The traceback mechanism should have the power to differentiate if two packets from same source to same destination traverse in different paths. The fifth assumption is associated with the observation of Subramanian (Burch and Cheswick, 2000; Snoeren *et al*., 2001) which depicts that as networks get richer connections, Internet routes are less likely to be circuitous. Finally we consider the routers. It is vital to the feasibility of GT, the comparative direction of a router with respect to its adjacent routers can be known through network configuration. The routing information should be updated periodically.

GDT principle: By combining AD/PAD (Chen *et al*., 2007) and DGT (Gao and Ansari, 2005) Principles, we propose a new scheme called Geographical Division Traceback (GDT). As per our GDT Principle, the attacker will be identified with the fast convergence using single packet. Because whenever a malicious program originates from a location, its geographical information will be encoded in the IP header. To perform traceback, victim does not have to use source IP address to locate the attacker. Because IP address can be spoofed. The steps involved in the proposed Traceback scheme are as follows:

- Consider the Entire world as a grid
- It is divided into various subspaces
- Every subspace has two bit identifier
- Repeat the steps 2 and 3 until the attack source region is identified

The value is assigned to the geographical quadrant as mentioned below in Fig. 1.

The entire geographical location can be covered by assuming the total area including the land and water. The details of the Geographical locations are as follows.

The information's mentioned in the below Table 1 are extracted from (Keromytis *et al*., 2002). The entire world of 510072000 sq km will be initially divided into 4 equal quadrants. As a result, the world will be divided in to four equal quadrants each of 127518000 sq km. Once it got divided, each quadrant will be assigned with two bit identifier as mentioned in Fig. 2.

In the next iteration, each quadrant of 12751800 sq.km will be further divided into 4 quadrants of each 31879500 sq km and those four portions will be again given identifiers 00, 01, 10 and 11.

If the packet originates from any machine, its geographical information will be encoded in the IP header when it enters into the first router. Geographical information is nothing more than the bit value associated with the corresponding quadrant. Consider the following example information in the IP Header.

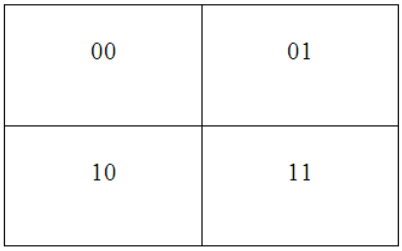

Fig. 1: Assigning values to the Quadrant

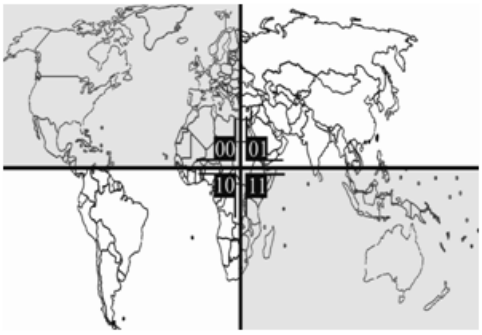

Fig. 2: Division of geographical locations

Table 1: Geographical information

Reference Variable	Total area of the world in sqr.km (water+land)	Total area of the world in sqr. km (only land)
A	510072000.0	148940000.0
B=A/4	127518000.0	37235000.0
C=B/4	31879500.0	9308750.0
D=C/4	7969875.0	2327187.5
E=D/4	1992468.8	581796.9
F=E/4	498117.2	145449.2
G=F/4	124529.3	36362.3
H=G/4	31132.3	9090.6
I=H/4	7783.1	2272.6
J=I/4	1945.8	568.2
K=J/4	486.4	142.0
L=K/4	121.6	35.5
M=L/4	30.4	8.9
N=M/2	15.2	4.4

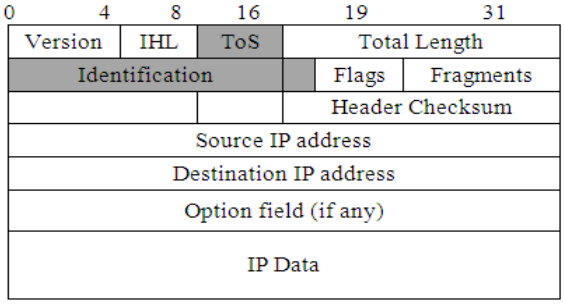

Fig. 3: The IP header fields (darkened) utilized in GDT

Fig. 4: Memory representation of bit identifier

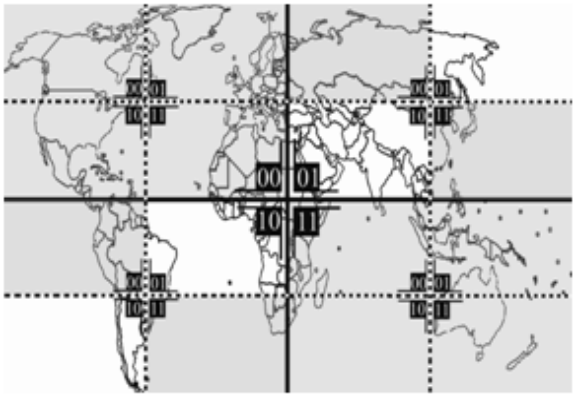

Fig. 5: Subdivision of the Earth

Table 2: Performance evaluation of GDT with other traceback techniques

Techniques	Scalability	Capability to mitigate the effects of attack
PPM	Good	Poor
DGT	Good	Good
GDT	Good	Best

In order to traceback the packet with the above information, victim can directly trace the attacker in the first quadrant (00) of the world. In that first quadrant, tracing can be further confined by searching the packet in the fourth quadrant (11) of the first quadrant. Tracing can be further confined based on the information available in the next coming fields. Hence tracing the attacker can be done quickly.

Identification field of the IP header is usually used to conduct packet marking (Xiang *et al*., 2008) and (Al-Duwairi and Govindarasu, 2006). It is only used for reassembling fragments. Since the fragmented packets are very rare, ID field of 16 bits can be used for storing the information (Gao and Ansari, 2005).

But these 16 bits are not sufficient to encode the information about all the 13 times of division process and its quadrant value. Hence in addition to that, type of service field of the IP header (Xiang *et al*., 2008) which is of 8 bits and a flag bit can also be used to encode the information. So in total we have 25 bits (16+8+1) as shown in Fig. 3.

The utilization of these bits for storing the geographical location leads to the identification of the location to which the source of attack belongs to. The information can be stored based on the order of the sub division of the geographical region denoted in Fig. 4.

Assume that the entire world got divided into four quadrants. First quadrant is given with the identifier 00; second one with the identifier 01, third with the 10 and fourth quadrant is 11. In the next iteration, each quadrant is further divided in to four other quadrants and the same process got repeated four times as shown in Fig. 5. Attacker's location is depicted in the below picture. The packet started from that location will have the following information in the ID field of the IP header.

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

```
/* Number of elements in the "FixVal" array to the
number of subspace of Geographic"G" */
    FixVal:= allocated_mamory(G.tot_space);
    For (i := 0; i <= FixVal ; i :=i+1)
    For (j := 0; j <= FixVal ; j :=i+1)
            Routerval[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 :=i+1)
        If(Routerval[i][j]==space_A)
        Path_set[i][j] := Search_path(flag,i,j);
        Flag:=Flag+2;
        Else If(Routerval[i][j]==space_B)
        Path_set[i][j] := Search_path(flag,i,j);
        Flag:=Flag+2;
      Else If(Routerval[i][j]==space_C)
        Path_set[i][j] := Search_path(flag,i,j);
            Flag:=Flag+2;
      Else If(Routerval[i][j]==space_D)
        Path_set[i][j] := Search_path(flag,i,j);
        Flag:=Flag+2;
        Else
        Return Path_set;
```

```
    Break;
  end For
  end For
```

As per the GDT Principle not only the four times of division as that of in the above example, the division of space into the 4 equal subspaces will be repeated 12 times. Because of this divide and conquer approach victim will be end up with the location of 4.4 sq km. In that small area, only one router will be there using which attacker can be easily traced. The performance evaluation of GDT with other techniques is listed in Table 2.

RESULTS 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 scalability comparison with different techniques is shown below in Fig. 6.

As a result, the GDT 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 easily.

The Performance comparison based on the number of router traced is shown below in Fig. 7. The result shows that the performance of the other existing tehcniques reduces as the number of routers the packet crossed increases.

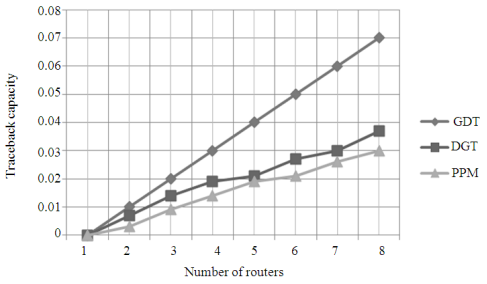

Fig. 7: Traceback capabality in GDT

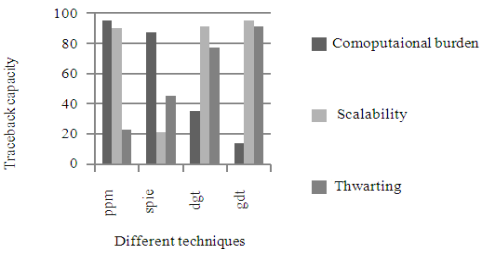

Fig. 6: Comparision among diferent Techniques

CONCLUSION

In this study, a new countermeasure GDT is proposed to defend an attack from any part of the world with a single packet. “Divide and conquer” approach is implemented to detect the attacker and throttle the incoming traffic. This GDT 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
- Easy to mitigate and prevent the further attacks

The future work involves the Thwarting of DDoS attacks and reduces the traffic by routing the information packet as early as possible.

REFERENCES

Al-Duwairi, B. and M. Govindarasu, 2006. Novel hybrid schemes employing packet marking and logging for IP traceback. IEEE Trans. Parallel Distributed Syst., 17: 403-418. DOI: 10.1109/TPDS.2006.63

Al-Duwairi, B., 2006. Novel hybrid schemes employing packet marking and logging for IP traceback. IEEE Trans. Parallel Distribution Syst., 17: 403-418. DOI: 10.1109/TPDS.2006.63

Burch, H. and B. Cheswick, 2000. Tracing anonymous packets to their approximate source. Carnegie Mellon University.

Chen, R., J.M. Park and R. Marchany, 2007. A divide-and-conquer strategy for thwarting distributed denial-of-service attacks. IEEE Trans. Parallel Distribution Syst., 18: 577-588. DOI: 10.1109/TPDS.2007.1014

Gao, Z. and N. Ansari, 2005. Directed geographical traceback. Proceedings of the 3rd International Conference on Information Technology, Research and Education, June, 27-30, IEEE Xplore Press, Newark, NJ, USA, pp: 221-224. DOI: 10.1109/ITRE.2005.1503108

Ghazali, K.W.M. and R. Hassan 2011. Flooding distributed denial of service attacks-a review. J. Comput. Sci. 7: 1218-1223. DOI: 10.3844/jcssp.2011.1218.1223

Keromytis A.D., K.V. Misra and D. Rubenstein, 2002. SOS: Secure overlay services. Proceedings of the Conference on Applications, Technologies, Architectures and Protocols for Computer Communications, (ATAPCC’02), ACM New York, NY, USA., pp: 61-72. DOI: 10.1145/633025.633032

Savage, S., D. Wetherall, A. Karlin and T. Anderson, 2000. Practical network support for IP traceback. Comput. Commun. Rev. 30: 295-306.

Snoeren, A.C., C. Partridge, L.A. Sanchez, C.E. Jones, F. Tchakountio *et al*., 2001. Hash-Based IP traceback. Proceedings of the Conference on Applications, Technologies, Architectures and Protocols for Computer Communications, (ATAPCC' 01), ACM New York, NY, USA., pp: 3-14. DOI: 10.1145/383059.383060

Xiang, Y., W. Zhou and M. Guo, 2008. Flexible deterministic packet marking: An IP traceback system to find the real source of attacks. IEEE Trans. Parall Distribut. Syst., 20: 567-580. DOI: 10.1109/TPDS.2008.132