ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

Survey of IP Traceback Methods in Distributed Denial of Service (DDoS) Attacks

G.Florance

Assistant Professor, Department of Computer Science Engineering, Panimalar Engineering College, Chennai, India

ABSTRACT: Tracing DoS attack that employ source address spoofing is an important and challenging problem. It is important to protect the resource and trace from the distributed denial of service (DDoS) attack, but it is difficult to distinguish normal traffic and DoS attack traffic because the DoS generally hide their identities/origins. Especially the attackers often use incorrect or spoofed source IP address, so tracing the source of the denial of service is hardest in internet. So many techniques and methodologies are used to trace the DDoS attacks. This paper briefly describes different IP traceback techniques to solve the problem. The main goal of this paper is, to describe the different IP traceback techniques of the DDoS.

KEYWORDS: IP spoofing, DoS, DDoS, zombies, botnet, IP Traceback methods.

I. INTRODUCTION

Attacks that employ source address spoofing represent a growing threat to the Internet infrastructure. Denial of Service (DoS) attacks and a more complicated version known as Distributed DoS (DDoS) are the most common to take advantage of source address spoofing. These attacks deny regular Internet services from being accessed by legitimate users either by blocking service completely or by disturbing it such that users become not interested in the service anymore. In fact, IP traceback schemes are considered successful if they can identify the zombies from which the DDoS attack packets entered the Internet.

In denial of service attacks, the packets are routed correctly but the destination becomes the target of the attackers. DoS attacks are very easy to generate and are very difficult to detect. In a typical DoS attack the attackers' nodes spoofs its IP address and uses multiple intermediate nodes to overwhelm other nodes with traffic. DoS classified into two main types, ordinary and distributed. In an ordinary network based denial of service attack, an attacker uses a tool to send packets to the target system. In the DDoS attacks, there might still be a single packets, but the effect of the attacks is multiple by use of attack servers. The attack not only disables that server but denies access to legitimate user.

To launch a DDoS attack, the attacker(s) first establishes a network of computers that will be used to generate the huge volume of traffic needed to deny services to legitimate users of the victim. To create this attack network, attackers discover vulnerable hosts on the network. Vulnerable hosts are those that are either running no antivirus or out-of-date antivirus software, or those that have not been properly patched. These are exploited by the attackers who use the vulnerability to gain access to these hosts. The next step for the attacker is to install new programs (known as attack tools) on the compromised hosts of the attack network. The hosts running these attack tools are known as zombies, and they can be used to carry out any attack under the control of the attacker.Numerous zombies together form an army or botnet. In a typical DDoS attack, the master computer orders the zombies to run the attack tools to send huge volume of packets to the victim, to exhaust the victim's resources.

IP traceback is the process of identifying the actual source(s) of attack packets. This has the benefit of holding attackers accountable for abusing the Internet. Also, it helps in mitigating DoS attacks either by isolating the identified attack sources or by filtering attack packets far away from the victim as proposed in the IP tracebackbased intelligent packet filtering technique [2]. IP traceback is a challenging problem because of the distributed anonymous nature of

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

DDoS attacks, the stateless nature of the Internet, the destination oriented IP routing, and the fact of having millions of hosts connected to the Internet (implying huge search space). All these factors help attackers to stay behind the scenes and hence complicate the process of traceback. To allow for effective attack origin tracing in the future Internet, a large number of IP traceback schemes [1],[3] have been proposed. These schemes all require intermediate routers to inscribe a mark into the packet header that encodes the identity of an intermediate router or edge. After collecting a sufficient number of such marks, the victim will be able to reconstruct the network paths leading to the attackers. In fact, IP traceback schemes are considered successful if they can identify the zombies from which the DDoS attack packets entered the Internet.

II. DDOS ATTACKS

Distributed Denial of Service(DDoS) is an attack which denies authorized user access to the service provider. The recent report shows that the most expensive computer crime over the past year was due to denial of service. DoS attack target different layers of OSI model [16]

- **Physical layer**: by accidentally cutting a communication cable to take down network services.
- **Data link layer**: to disable the ability of hosts to access the local network.
- **Network layer**: by sending a large amount of IP data to a network.
- **Transport layer**: by sending many TCP connection requests to a host
- **Application layer**: by sending large amount of legitimate requests to an application.

The goal of the attacker is to disrupt or shut down an organization's business critical services such as ecommerce transactions, financial trading, email or web site access. By overwhelming network infrastructure, servers or applications with excessive communications requests, an attack means services are unavailable to legitimate users. A Distributed Denial of Service (DDoS) attack is an attack to prevent the users from using the resources of a victim's computer. It is a large scale attack in a co-ordinated fashion, which is typically launched indirectly with the help of other computers in the Internet.

DDOS attack occurs when multiple system resources is not available to network users. A DOS attack floods the remote system with so much traffic that it cannot handle normal, valid requests made from others network systems[16].DDoS attack affects both the victim and the networklink. The victim is suffering from a DDoS attack with two major impacts. Firstly, the victim has limited resourcesfor processing the incoming packets. The victim'sresources, such as CPU and memory, will be exhausted,and then the victim will be unable to serve for normaltraffic. Secondly, consumption of network bandwidth willresult in legitimate flows being blocked.

DDoS attacks with source IP address spoofing is of two types. First is direct attack in which the attackers sendmalformed packets with fake source IP address and second is reflector attack in which attacker send large number ofattack packets through large number of compromised hosts in the network. Network resources are thus wasted inprocessing such attack packets causing denial of service to legitimate clients.There are two main classes of attacks: (1) bandwidth depletion and (2) resource depletion attacks. In case of bandwidth depletion attack, the victim network is flooded with unwanted traffic that prevents legitimate traffic from reaching the victim computer. In the other case of resource depletion attacks, the attack is targeted to tie up the resources of the victim computer [4] [5].

While forwarding packets towards destination only destination address is often used and source address is neververified. Attackers take advantage of this fact for launching attack using spoofing of source IP address in order to hidetheir identity and avoiding the possibility of getting caught.

III. CLASSIFICATION OF IP TRACEBACKMETHODS

IP traceback methods provide the victim'snetwork administrators with the ability to identifythe address of the true source of the packets causinga DoS. IP traceback is vital for restoring normalnetwork functionality as quickly as possible, preventingreoccurrences, and ultimately, holding the attackersaccountable.A number of IP traceback

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

approaches have been suggested to identify attackers.There are two major methods for IP traceback such asreactive orproactive method.

Reactive method initiate the tracebackprocess in response to an attack. They must be completedwhile the attack is active; they're ineffective onceit ceases. Link testing(Input debugging, controlled flooding) ,ingress filtering/egress filtering and hop count filtering areexamples of reactive method. This method is also referred as source based mechanism. Mostreactive methodrequire a large degree of ISP cooperation,which leads to extensive administrative burden anddifficult legal and policy issues, thus, effective IP tracebackmethods should require minimal onISP territory. Essentially, reactive methodare more effective for controlled networks than forthe Internet.

Proactive method also referred as destination based mechanism, which record tracing informationas packets are routed through the network. Thevictim can use the resulting traceback data for attackpath reconstruction and subsequent attacker identification.Examples of proactive method include logging,ICMP, log based traceback, hase based traceback, FDPM traceback, TBPM traceback, traffic filtering, packet marking and filtering, distributed link list traceback(DLLT), probabilistic pipelined packet marking(PPPM), Deterministic packet marking(DPM) and entropy variation.

IV. DIFFERENT IP TRACEBACK METHODS

A. Reactive methods

Link testing: As the name implies, link-testing method(sometimesreferred to as hop-by-hop tracing) work by testing networklinks between routers to determine the origin ofthe attacker's traffic. Most techniques start from therouter closest to the victim and interactively test its incoming(upstream) links to determine which one carriesthe attack traffic. Link testing is a reactive method and requiresthe attack to remain active until the trace is completed.

Input debugging[23]: It is one implementation of the linktesting approach.This feature lets the administrator determine incomingnetwork links for specific packets. If the routeroperator knows the attack traffic's specific characteristics(called the attack signature), then it's possible to determinethe incoming network link on the router. TheISP must then apply the same process to the upstreamrouter connected to the network link and so on, untilthe traffic's source is identified.

Controlled flooding: This technique[24] works by generatinga burst of network traffic from the victim's networkto the upstream network segments and observing howthis intentionally generated flood affects the attack traffic'sintensity. Using a map of the known Internet topologyaround the victim, these packet floods are targetedspecifically at certain hosts upstream from the victim'snetwork; they iteratively flood each incoming networklink on the routers closest to the victim's network.

Ingress/Egress filtering: Ingress Filtering, proposed by Ferguson et al. [25], is a restrictive mechanism to drop traffic with IP addresses that do not match a domain prefix connected to the ingress router. Egress filtering is an outbound filter, which ensures that only assigned or allocated IP address space leaves the network. A key requirement for ingress or egress filtering is knowledge of the expected IP addresses at a particular port. For some networks with complicated topologies, it is not easy to obtain this knowledge. One technique known as reverse path filtering [17] can help to build this knowledge. This technique works as follows. Generally, a router always knows which networks are reachable via any of its interfaces. By looking up source addresses of the incoming traffic, it is possible to check whether the return path to that address would flow out the same interface as the packet arrived upon. If they do, these packets are allowed. Otherwise, they are dropped. Unfortunately, this technique cannot operate effectively in real networks where asymmetric Internet routes are not uncommon. More importantly, both ingress and egress filtering can be applied not only to IP addresses, but also protocol type, port number, or any other criteria of importance. Both ingress and egress filtering provide some opportunities to throttle the attack power of DoS attacks. However, it is difficult to deploy ingress/egress filtering universally.

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

Hop Count Filtering: It is based on packet processing approach for identifying attackers usingspoofed source IP address. In this method the packets from the systems at the same hop count passing through the samerouter are marked with the same identification number which is the combination of 32 bits IP address of the router pathand the encrypted value of the hop count. This value is matched with already stored value at receiving router. Thus,attack packets are identified early and spoofing threats are reduced.

B. Proactive methods

Logging: It is to log the packets at key routersthroughout the Internet and then use data-mining techniquesto extract information about the attack traffic'ssource. The most significant drawbacksinclude the amount of processing and storagepower needed to save andshare this information among ISPs poses logistical andlegal problems as well as privacy concerns. Given today'slink speeds, packet logs can grow quickly to unmanageablesizes, even over short timeframes.

Alex Snoeren and colleagues[26] proposed a novelapproach to logging IP traceback called SPIE(Source Path Isolation Engine). Instead of storing thewhole packet, they suggested storing only a hash digestof its relevant invariant portions in an efficient memorystructure called a Bloom filter. To complete an IP tracebackrequest, a network of data collection and analysis agents spanning the different networks could use thismethod to extract significant packet data and generateappropriate attack graphs, thus identifying the attacktraffic's origin.

Tatsuya Baba and Shigeyuki Matsuda[27] proposed an alternateand innovative logging approach. It entailed anoverlay network built of sensors that could detect potentialattack traffic, tracing agents (tracers) that could log theattack packets on request, and managing agents that couldcoordinate the sensors and tracers and communicate witheach other. It also allows for increased speed andrequires less storage.

ICMP based traceback: One of the iTrace scheme's weaknesses becomes apparentin a DDoS attack in which each zombie contributesonly a small amount of the total attack traffic. Insuch cases, the probability of choosing an attack packet ismuch smaller than the sampling rate used. To overcome this drawback, researchers proposed anenhancement to iTrace called intension-driven ICMPtraceback[28]. This technique separates the messagingfunction between the decision module and the iTrace generationmodule. A recipient network supplies specific informationto the routing table to indicate it requests ICMPtraceback message and the decision modulewould select which kind of packet to use next to generatean iTrace message. Based on this decision, the decisionmodule will set one special bit in the packet-forwardingtable and this bit indicates that thevery next packet corresponding to that particular forwardingentry will be chosen to generate an iTrace message.

Log-Based Traceback: The basic idea of log-basedtraceback is that each router stores the information(digests, signature, or even the packet itself) of networktraffic through it. Once an attack is detected, the victim queries the upstream routers by checking whether theyhave logged the attack packet in question or not. If theattack packet's information is found in a given router'smemory, then that router is deemed to be part of theattack path. Obviously, the major challenge in log-basedtraceback schemes is the storage space requirement at theintermediate routers [9].

Hash Based IP Traceback: Hash based approach [10]is called as Source Path Isolation Engine (SPIE). In thesemethods the forwarding path of a single packet can bereconstructed by querying such routers soon after thepacket is observed. More recent work (privatecommunication) moves the processing from the router toa specialized machine observing traffic on a link. Thismethod can be viewed as a special case of RemoteMonitors. Attacks on SPIE: Attackers can attack thequery/response communication, either the traffic or theendpoints. For that reason access to traceback data willnormally be restricted to the administrative domainowning the routers and possibly a few other trustedplaces.

FDPM Traceback: Flexible Deterministic PacketMarking (FDPM) is falls into the packet markingcategories [11]. It is optimized version of DPM. Thisscheme provide more flexible features to trace the IPpackets and can obtain better tracing capabilities overother previous IP traceback mechanisms, such as Linktesting, logging, ICMP traceback, probability packetmarking (PPM) and Deterministic packet marking(DPM).In FDPM schemes, the Types of Services

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

(ToS)fields will be used to store the mark under somecircumferences. The two fields in the IP header areexploited, one is fragment ID and other is Reversed flag.An identifying value is assigned to the ID field by thesender to aid in assembling the fragments of a datagram.Given that less than 0.25% of all internet traffic isfragments, this field can be safely overloadedwithout causing serious compatibility problems. DPMreconstruction process includes two steps: markrecognition and address recovery. Compared to DPM, thereconstruction process is simpler and more flexible.When each packet that is used to reconstruct the sourceIP address arrives at the victim, it is put into a cache,because in some cases the processing speed is lower thanthe arrival speed of the incoming packets.

TBPM Traceback: Topology aware single packet IPtraceback system is namely TOPO [12].It based on thebloom filter [13], which utilizes router's local topologyinformation, i.e., its immediate predecessor information,to traceback. TOPO can significantly reduce thenumber and scope of unnecessary queries and thus,significantly decrease the false attributions to innocentnodes. The main goals of TOPO as follows: 1) todesign a single packet IP traceback system, this hasfewer unnecessary query messages and fewer falseattributions to innocent nodes.2) to design a singlepacket IP traceback system this needs not to be fullydeployed in the entire network.3) to design amechanism which helps achieve the best performanceof Bloom filters by adaptively adjust using parameter.Topology Based Packet Marking (TBPM) has been anew approach in Anti-IP spoofing techniques [14].TBPM builds on the strengths of the packet markingprincipal; however it focuses not merely on the source,but also the path traversed by a datagram. We havepointed out how a route discovery method can be moreeffective, especially during DoS attacks where edgerouters that mark packets may themselves beunavailable as a result of the attack.

Traffic filtering: It is another method to prevent DoS attacks to define a limit for the AP(Access Points) to process the management frames in per second. AP will count the number of management frames per second receiving from any particular MAC address and if that are exceed from an already decided limit then next all frames will be ignored at that second for that particular MAC address. A problem can occurs only in a case if an intruder is sending continuously management frames by changing the MAC address for every frame per second then AP will process all the frames understanding that a large number of clients want to associated simultaneously.

Packet marking and filtering[15]: Two markingschemes are used namely stack based marking and write ahead marking for improving performance of Pi markingscheme. Optimal threshold strategy is used for flooding based spoof source IP address attacks. Stack marking is basedon TTL field for identifying path from source to destination and in write ahead marking routers mark for its next hoprouter. The results showed that method is efficient even when only 20% routers are involved in marking process.

Distributed link list trace back (DLLT):DLLT is based on the idea of preserving the marking information at intermediate routers and it can be collected using a link list based approach. It combines the features of probabilistic packet marking (PPM) [3], [18] and Hash-based trace back (HBT) schemes ie, distributed link list trace back (DLLT) =PPM+HBT.

In probabilistic packet marking(PPM), routers are far away from the victim have a very low chance of passing their marking information to the victim because intermediate routers overwrite this information, which leads to the loss of valuable marking information written by routers near the attacker. So this is contradiction to our goal.Hash based trace back(HBT) [19] method employed to collect packet information from network routers is inefficient and requires special resources; the scheme assumes that a central management unit is available in each domain to download and search all packet digests looking for specific packet footprints. This results in a tedious process and an unnecessary overhead. The deterministic storage algorithm increases the memory requirement of the scheme. The major drawback of hash based trace back is, to log information about every forwarded packet.

DLLT is based on a "store, mark, and forward" approach. A fixed-size marking field is allocated in each packet. Any router that decides to mark the packet, stores the current content of the marking field (which was written by the previous marking router) in a special data structure, called the Marking Table, maintained at the router. The router generates an ID for that packet to index its marking information in the marking table. The router marks the

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

packet by overwriting the marking field by its own IP address and then forwards the packet. Any router that decides not to mark the packet just forwards it.

Probabilistic Pipelined Packet Marking (PPPM):Probabilistic pipelined packet marking (PPPM) [3], based on the pipelined marking concept. The objective of PPPM is, the destination know about all routers that were involved in marking certain packet P, using constant space in the IP packet header without incurring long term storage overhead at intermediate routers. The marking information field allocated in each packet consists of two parts: the IP address of the marking router, MR, and an ID used to link marking done for a given packet by different routers.

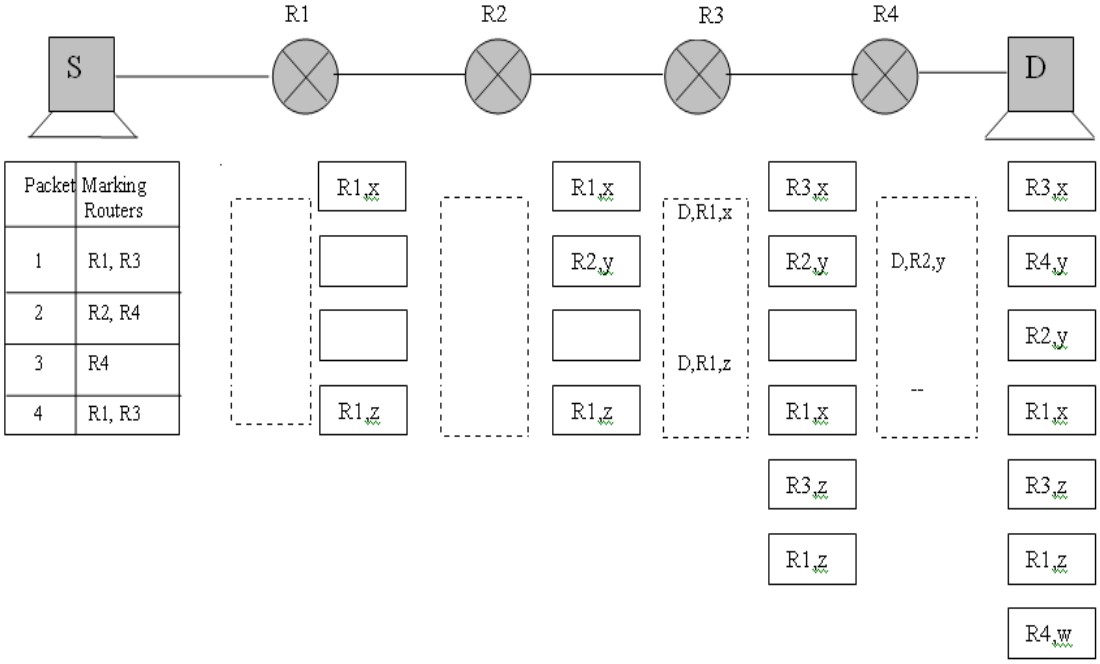

Fig 3.Example of pipelined based packet marking. Dashed rectangular boxes represent buffers associated with each router. Marking information associated with each packet is shown in a solid rectangular box, for each packet at each router along its path.

Fig 3 shows an example of pipelined packet marking. In this example, routers R_1 and R_3 decide to mark the first packet. Since R_1 is the first marking router of packet 1, it puts its own mark into the packet. This consists of its IP address, R_1, and a randomly chosen ID, x. Notice that nothing is buffered at R_1. The same packet is remarked by R_3. Since the packet is already marked, R_3 has to buffer the marking information for the packet's destination before writing its own IP address in the packet. Notice that R_3 keeps the same ID, x, in the marked packet. The buffered information is transferred to D when R_3 decides to mark another packet destined to D. This happens to be the fourth packet. Since both packets (i.e., the first and the fourth) hold the same ID, x by the time they arrive at D, it is easy for D to conclude that packet x has been marked by R_1 and R_3 and so on.

Deterministic Packet Marking (DPM): Deterministic Packet Marking (DPM) in which edge routers only perform packet marking. In DPM 16-bit Packet ID field and 1-bit Reserved Flag (RF) in the IP header will be used to mark packets. Each packet is marked when it enters the network. This mark remains unchanged for as long as the packet traverses the network. The packet is marked by the interface closest to the source of the packet on an edge ingress router. The mark is a partial address information of this interface. The interface makes a distinction between incoming and outgoing packets. Incoming packets are marked; outgoing packets are not marked. This ensures that the egress router will not overwrite the mark in a packet placed by an ingress router. A 32-bit IP address needs to be passed

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

to the victim. A total of 17 bits are available to pass this information. A single packet would not be enough to carry the whole IP address in the available 17 bits. Therefore, it will take at least two packets to transport the whole IP address. An IP address will be split into two segments, 16 bits each: segment 0- bits 0 through 15, and segment 1- bits 16 through 31. DPM Mark [20], [21] can be used to not only transfer the bits of the ingress address but also some other information. This additional information should enable the destination to determine which ingress address segments belong to which ingress address. The destination would first put the ingress address segments in RecTbl, and then only after correctly identifying the ingress address, out of many possible address segments permutations, would transfer it to IngressTbl.

Entropy Variation: In this paper, we use flow entropy variation or entropy variation interchangeably. Once a DDoS attack has been identified, the victim initiates the pushback process to identify the locations of zombies: the victim first identifies which of its upstream routers are in the attack tree based on the flow entropy variations[22] it has accumulated, and then submits requests to the related immediate upstream routers. The upstream routers identify where the attack flows came from based on their local entropy variations that they have monitored. Once the immediate upstream routers have identified the attack flows, they will forward the requests to their immediate upstream routers, respectively, to identify the attacker sources, this procedure is repeated in a parallel and distributed fashion until it reaches the attack source(s) or the discrimination limit between attack flows and legitimate flows is satisfied.

V. CONCLUSION

This paper describes the detailed survey of different Distributed Denial of Service traceback mechanisms. It is also describes about two ways of classification such as reactive and proactive traceback methods. Possible attacks in a collaborative environment and their impacts are identified of which denial of service is a serious threat. There are many traceback methods are available to identify the attacks. The real challenge in security provisioning is to identify the source of unknown attack at the earliest possible which motivated us to work on novel fast traceback mechanism with less computation and storage costs, scalability to high attacker population, and providing best network performance.

REFERENCES

[1]A. El-Atawy et al., "Adaptive Early Packet Filtering for Protecting Firewalls against DoS Attacks," Proc. IEEE INFOCOM, 2009.
[2]A. Belenky and N. Ansari, "On IP Traceback," IEEE Comm. Magazine, pp. 142-153, July 2003.
[3] B. Al-Duwairi and M. Govindarasu, "Novel Hybrid Schemes Employing Packet Marking and Logging for IP Traceback," IEEE Trans. Parallel and Distributed Systems, vol. 17, no. 5, pp. 403- 418, May 2006.
[4] Thomas Dubendorfer, Matthias Bossardt, Bernhard Plattner; Adaptive Distributed Traffic Control Service for DDoS Attack Mitigation; Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05) - Workshop 17 - Volume 18, 2005.
[5]Stephen Specht, Ruby Lee; Taxonomies of Distributed Denial of Service Networks, Attacks, Tools, and Countermeasures; Department of Electrical Engineering, Princeton Architecture Laboratory for Multimedia and Security, Technical Report CE-L2003-03, May 16, 2003
[6] P. Ferguson, and D. Senie, "Network ingress filtering: Defeating denial of ser-vice attacks which employ IP source address spoofing," RFC 2267, the Internet Engineering Task Force (IETF), 1998.
[7]Baker, F. "Requirements for IP version 4 routers," RFC 1812, Internet Engineering Task Force (IETF).Go online to www.ietf.org.