Study on Classification and Characteristics of Source Address Spoofing Attacks in the Internet

Jun Bi, Ping Hu, Peiguo Li
Tsinghua National Laboratory for Information Science and Technology
Network Research Center of Tsinghua University
Beijing 100084, China

***Abstract*—Source address spoofing is one of the most dangerous means of attacks in today's Internet. In this paper, we study and classification and characteristics of the source address spoofing. Source address spoofing can be classified in six categories. The characteristics of first category of source address spoofing are analyzed by the statistical analysis of packets collected by the CAIDA network telescope, because we found that partial information of the original source address spoofing attack packets could be found in the ICMP error packets returned to the network telescope.**

Keywords- source address spoofing; network attacks; backscatter packets

I. INTRODUCTION

In the current Internet architecture, the IP address has dual important characteristics, namely, the location and identity. By design, the Internet forwards data packets solely based on the destination IP address. The source IP address is not validated during the forwarding process in most cases. This makes it easy for malicious hosts to spoof the source address of the IP packet [1]. Source address spoofing attack is an attacker sends forged packets in the Internet and takes over the identity of a trusted host in order to subvert the security of the target host [2], [3]. It can also attack receiver through any innocent third parties reflection [4].

Although not all the source address spoofing action launch attacks, there were tremendous harms with IP spoofing. Therefore, to study the classification and characteristics of IP spoofing will help people to understand the problems and find the solutions.

This paper is organized as follows. We first briefly describe relates works about source address spoofing in Section 2. Then we give a detailed description of the classification of source address spoofing in Section 3. In Section 4 we analyze H0 category attack by backscatter data capture by CAIDA network telescope. Finally, Section 5 will conclude the paper.

II. RELATED WORK

Source address spoofing technology was discovered in an academic conference in 1985. In this meeting, Robert T. Morris proposed the TCP/IP protocol sequence of predictability and source address spoofing technical bug [5]. This discovery indicates that the application of TCP/IP protocol is not fully secure. However, the vast majority of research work [1], [4], [6][7][8][9][10][11][12][13], and [14] were based on the DoS [15] or DDoS [16] attacks using IP spoofing, from the viewpoint of the format of some concrete attacks. This is few study are from the viewpoint of the format of source address spoofing itself, which will be addressed in this paper.

There is some work on the measurement of source address spoofing. According to MIT's *spoofer* project [17], in the Internet, the net-blocks spoofable accounted for about 16%, the IP addresses spoofable accounts for about 20%, and the Autonomous Systems spoofable accounting for about 24%. This paper will analyze the traffic of one category of source address spoofing attacks.

III. CLASSIFICATION OF SOURCE ADDRESS SPOOFING ATTACKS

The IP spoofing attacks can be described by three elements: the attacker A, the victim V, and the host H which is being spoofing.

The total set of all the IP addresses defined as the set I, in which all the announced valid IP addresses are defined as the set I'. As set I for the total set, the complement set of I' is denotes as set I'', where $I' \cup I'' = I$ and $I' \cap I'' = \Phi$.

The set of source addresses of attackers is defined as I_A, and the set of addresses of other hosts on the same sub-network of attackers is defined as I_{AN}.

The set of source addresses of victims as I_V, and the set of source and other hosts on the same sub-network of victims is defined as I_{VN}.

The set of source addresses of hosts/routers on the path between the attacker A and victim V is defined as I_K.

We define set of source addresses of hosts/routers which are not in the same sub-network of attackers or victims, nor on the path between attackers and victims, as set I_O, where $I_O \cup I_A \cup I_V \cup I_{AN} \cup I_{VN} \cup I_K = I'$.

In the more general cases, attackers and victims are not in the same sub-network. Then the source address spoofing can be classified in six categories form H0 to H5 by the position of the host being spoofed:

H0 category means that the source address of being spoofed is non-existent or inactivated in the Internet, $H0 = (i \mid i \in I'')$;

978-0-7695-3979-9/10 $26.00 © 2010 IEEE
DOI 10.1109/ICN.2010.43

H1 category means that the source address of being spoofed is the victim, $H1 = (i \mid i \in I_V)$;

H2 category means that the source address of being spoofed is within the same sub-network of the victim, $H2 = (i \mid i \in I_{VN})$;

H3 category means that the source address of being spoofed is within the same sub-network of the attacker, $H3 = (i \mid i \in I_{VN})$;

H4 category means that the source address of being spoofed is on the path between the attacker and the victim, $H4 = (i \mid i \in I_K)$;

H5 category means that the source address of being spoofed is neither in the same sub-network of the attacker or the victim, nor on the path between the attacker and the victim, $H3 = (i \mid i \in I_O)$.

The relations of those sets are shown in figure 1.

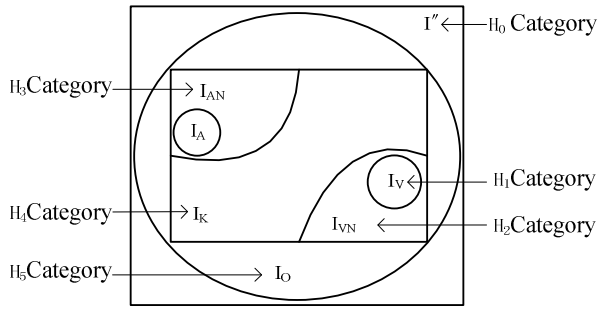

Figure 1. Classification of source address spoofing attacks set sketch

When the attacker and the victim are in the same sub-network, then H2 and H3 category will be equivalent. In this case, the attackers could launch non-blind attacks [19] because they can obtain TCP and ACK serial number by packets sniffer [18].

A. H_0 category

H0 category of attacks' characteristic is an attacker fakes non-existent or inactivated IP address to disable victim's normal services to other normal users. These addresess include the unallocated address, not visible in public or the private Internet address, that is, RFC1918 [20] definition of IP, RFC333 defined automatic allocation of address, loop address and address of network tests. Their forms or manifestations do not exist [21] in the Internet.

SYN flooding [2] [12] [13] is a typical way in the category of attacks. This will lead to depletion of victim of the CPU and memory space resources and prevent legitimate user from creating connection. A small number of SYN flooding attack could enable a remote host collapse if there are no other protection methods. However, in such attacks, attack can be achieved only by H0 category addresses because an activated host will return RST packets to release of holding half-open connection when it received a SYN-ACK packet from the victim.

B. H_1 category

The attacker fakes victim's IP address as source address. It can achieve reflective attacks, direct attacks and trap attacks. Basically, there are three forms:

The first, an attacker sends forged ICMP packets to the network broadcast address within victim host. Victim was submerged by the ACK floods, weakened the normal service and even collapsed.

The second, an attacker directly sents packets with the same source as destination IP address to a victim, that tried to send its response to itself, resulting in victim subject to interference and paralyze or restart.

The third, attacker forges the IP address of the victim in a number of unlawful actions, such as seizing bandwidth and so on. The host which is being spoofed then will be actually punished, because the source address is suppose to be the host identity in most cases.

Smurf attack [3] is an H1 category typical reflection attack. Attackers use ICMP echo request packets directed to IP broadcast addresses from remote locations to generate denial-of-service attacks. Attacker fakes H1 category source address so the host of destination address as the "signal amplifier", resulting in DoS in the same sub-network of victim.

The famous DrDoS [22] is another the typical examples of such attacks, which overflow victim or overload bandwidth.

TFN [23] and Land attacks [24] are also H1 category typical direct attacks. An attacker sends packets to a victim with the source host/port the same as the destination host/port and with SYN flag set. It will lock up victim or crash protocol stack.

C. H_2 category

The H2 category of attacks fakes source address within the same sub-network of the victim. In fact, such source address spoofing attacks often rely on the trust relationship between the victim and the host being spoofed.

The blind IP spoofing attacks [19] based on the TCP connection is a typical example of this category.

TFN2K [25] is also an example, that is next generation version of TFN [23].

D. H_3 category

The H2 category of attacks fakes source address within the same sub-network of the attacker, which could easily pass the ingress filtering, because the granularity of such filtering is not fine.

Bounce Scan [24] is a typical example of this category. In order to get response packet from victim, attackers fake the source address of a neighbor in the same sub-network and sniff the network traffic back to neighbor. This attack could be used on port scanning. If a port of the victim is closed, the victim will reply RST packet [26].

It makes the things even worse is that such attacks can effectively escape from uRPF [27].

E. H_4 category

In H4 category of attack, the attacker usually force the source address of the network equipment on the path from the attacker to the victim.

Under this type of attack, the attacker could disseminate false DNS or routing information, and redirect the network traffic [28]. This would be considered as one of the most dangerous attacks.

F. H_5 category

H5 category of attack is that the attacker does not rely on the special topological relationship between the victim and the source address being spoofed.

This category of attack is often combined, such as the MITM attack [29] is a sort of combination of two typical H5 category attacks. If the original communication is between the V1 and V2, then A forge the source address of V1 when communicates with V2, while forge the source address of V2 when communicates with V1.

IV. BACKSCATTER DATA ANALYSIS OF H_0 CATEGORY OF ATTACK

A. Backscatter ICMP technology

The SYN-ACK traffic will be replied to the forged host rather than attackers when the target received source address spoofing packet. The deployment of network Telescopes can capture the respond packets of fake source addresses of the attack traffic. Those packets are *Backscatter* packet [30]. Now, this is a very popular method to detect hosts that infected by attacks using not yet allocated IP addresses.

CAIDA deployed the 0./8 network Telescopes [31] to capture 224 IP address block traffic. The dark network address block is IPv4 address space of 1/256 and they found the daily traffic is usually above the 10G.

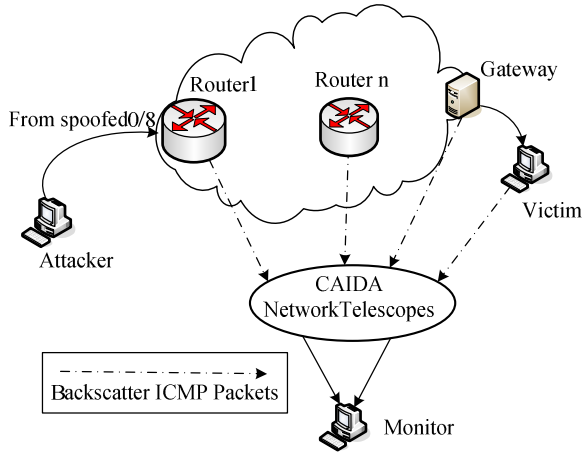

Figure 2. Backscatter ICMP principle

According to the principle of ICMP [32], [33], a packet reply immediately ICMP packet to the source address of host when the gateway, router or host forward and check transmission errors. ICMP is used to handle errors and exchange control messages by containing part of the original packets. Therefore, we found that partial information of the original source address spoofing packets could be found in the ICMP error packets returned to the telescope. The information will help us to analyze the characteristics of source address spoofing. Figure 2 shows the principle of the backscatter data analysis of ICMP packets captured by the network telescope.

B. Source address spoofing attack statistics and analysis

With data mining techniques, ICMP packets are extracted from samples after a series of processes, such as choosing and collating, checking the integrity and transforming.

During our research period, CAIDA released the latest data in 2008 (from 7:00am on May 21 to 7:00am on the May 29, 2008) [34]. We collect the ICMP messages during this period, as the distribution shown in figure 3 (21-07 means 7am on May 21).

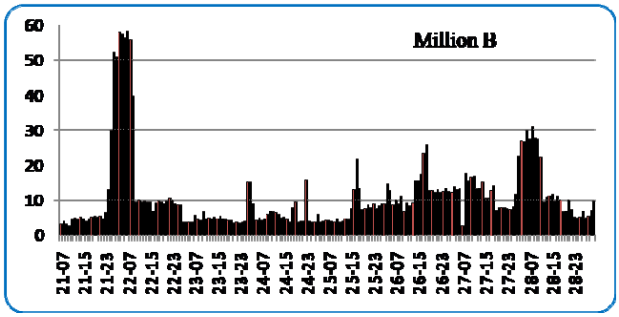

Figure 3. Traffic of collected ICMP packet which contains the information of orginial source address spoofing packets

As show in Figure 3, even the smallest hourly traffic is more than 2.5 million bytes, up to 59 million bytes. Huge volume of samples reflected the source address spoofing attack characteristics. In this paper, we analyze the target TCP/UDP port of the original source address spoofing packets. The numbers of attacks on major TCP/UDP ports are shown in figure 4.

As show in figure 4, we can draw a conclusion that in the H0 category of source address spoofing attacks targeted mainly at the TCP port 80, 443, 23, followed by the UDP port 53, 137, 138, 1900, 67 and TCP port 6667.

TCP port 80 (HTTP) and port 443 (HTTPS) [23][35-36]: TCP port 80 is opened for HTTP, mainly used for Web server. It can be seen from figure 4 that HTTP is the number one target of the source address spoofing attacks. Anti-spam organization Spamhaus [37] found that HTTP-based DDoS attacks is a great threat. The attackers often launches thousands threads to DDoS the target web server by controlling the BotNet. TCP port 443 is used as a secure web connection port of HTTPS. Form our analysis, HTTPS is the second frequent target of source address spoofing attacks, as shown in figure 4.

TCP port 23 (Telnet): Telent is the third frequent target of source address spoofing attacks in figure 4. We have to make some assumptions that the attacker could sniff the response from the victim server or blindly send message to the victim server.

UDP port 137 and 138 (NetBIOS service) [38]: UDP port 137 138 are mainly used for NetBIOS services and LAN to provide computer name or IP address for service of sharing of resources. Through NetBIOS protocol, the two ports will automatically be opened. We guess that some attackers sends service requests to targets by spoofing none-existence source address, then cause the targets continue to return broadcasting packet to respond to the request, resulting in a large number of invalid response packets and network congestion.

UDP 53 (DNS query): UDP port 53 is mainly used for DNS service. An attacker flood a huge number of DNS query packets with spoofed source address to disturb the normal service of DNS servers.

UDP 1900 (simple service discovery protocol) [39]: UDP port 1900 is mainly used for Window XP or later version to exchange information with equipment. However, this protocol has a fatal Bug to cause the infinite loop.

UDP port 67 (Bootp services): DHCP is extended from Bootp. DHCP server opens UDP port 67 and broadcast UDP port 68 responding to client requests. Since the protocol itself doesn't have access control mechanism, the attacker can spoof any client or DHCP server to steal the client's information or redirect the user traffic to a faked default gateway.

TCP port 6667 (IRC) [40-41]: TCP port 6667 is the default communications port in IRC, which is now being used between attacker and Zombies/Bots host to send the BotNet control instructions.

In summary, the source address spoofing attacks mainly targeted the following well-known ports and services: HTTP(TCP port 80), HTTPS(TCP port 443), Telnet(TCP port 23), NetBIOS(UDP port 137 and 138), DNS(UDP port 53), Simple service discovery protocol (UDP port 1900), DHCP(UDP port 66 and 67), and IRC(TCP port 6667). Form figure 4, we also found that TCP port 21, 22, 3389, 2119, 4600, and 15100 appeared frequently as the target of attacks in some period. For example, the attacks targets at TCP port 2119 remains high traffic volume from 23:00 May 25th to 14:00 May 26th, 2008.

V. CONCLUSION

Source address spoofing has become a widely used mechanism for attack s because it can be easily launched and is difficult to defend against and trace. Therefore, the classification of network attack s based on source address spoofing is very importance. This paper p resents a survey of the modes of source address spoofing and their effect on the Internet. The source address spoofing attack s are related to the attacker, the victim and the spoofed host. Source address spoofing attack s can be classified into six categories based on the position of the host being spoofed. This classification clarifies the problems of source address spoofing which w ill lead to imp roved prevention methods as a foundation of a trustworthy Internet architecture.

This paper also presents a new method to find the information of original source address attack packets from Backscatter ICMP packets captured by network telescope, and analyzes the latest Backscatter packets captured by CAIDA network telescopes. The traffic of ICMP packet is extracted and gathered by data mining and statistical analysis techniques. The distribution of TCP/UDP ports as the targets of source address spoofing attacks is presented. We discovered the attacks mainly concentrated on some well-known ports, in particular, TCP port 80 and 443 (HTTP and HTTPS).

ACKNOWLEDGMENT

This paper was supported by China Science and Technology Supporting Project under Grant 2008BAH37B02 and by the program of Ministry of Education of China for New Century Excellent Talents in University.

REFERENCES

[1] P. Ferguson, D. Senie. (2000,May) Network ingress filtering: Defeating Denial of Service Attacks Which Employ IP Source Address Spoofing. [Online]. Available: http://www.ietf.org/rfc/rfc2827.txt?number=2827.

[2] Computer Emergency Response Team (CERT). (2000, November 29). TCP SYN flooding and IP spoofing attacks. [Online]. Available: http://www.cert.org/advisories/CA-1996-21.html .

[3] Computer Emergency Response Team(CERT). (2000, March). Smurf IP Denial-of-Service Attacks. [Online]. Available: http://www/cert.org/advisories/CA-1998-01.html.

[4] D. Moore, C. Shannon, D. Brown, *et al*. (2006). Inferring Internet denial-of-service activity. [J]. IEEE/ACM Transactions on Computer System (TOCS).

[5] Tanase M. (2003, March 11). IP Spoofing an Introduction. [Online]. Available: http://www.securityfocus.com/infocus/1674 .

[6] V. Santiraveewan, Y. Permpoontanalarp. (2004,March) A Graph-based Methodology for Analyzing IP Spoofing Attack. [C]. 18th AINA, pp227-231.

[7] M. Bailey, E. Cooke, F. Jahanian, et al. (2006,March). Practical Darknet Measurement [C]. Proc. 40th Conf. on Information Sciences and Systems, pp1496-1501.

[8] G. Jin, H. Wang, K. G. Shin. (2003). Hop-count filtering: an effective defense against spoofed DDoS traffic. [C]. In Proceedings of the 10th ACM conference on Computer and communication security. Washington, D.C., USA.

[9] K. Park, H. Lee. (2001, August) On the Effectiveness of Route-Based Packet Filtering for Distributed DoS Attack Prevention in Power-Law Internets. [C]. ACM SIGCOMM 2001.

[10] J. Mirkovic, A. Hussain, B. Wilson, et al. (2007, June). Towards User-Centric Metrics for Denial of Service Measurement [C]. In Proceedings of the Workshop on Experimental Computer Science (part of ACM FCRC).

[11] K. Lee, J. Kim, K.H. Kwon, et al. (2008). DDoS attack detection method using cluster analysis. [J]. Expert Systems with Applications 34, pp1659–1665.

[12] W. Eddy. (2006, Dec.). Defenses Against TCP SYN Flooding Attacks. [J]. The Internet Protocol Journal, Volume 9, No. 4.

[13] W. Chen, D. Y. Yeung. Defending Against TCP SYN Flooding Attacks Under Different Types of IP Spoofing [C]. ICN/ICONS/MCL Jan 2006, pp38-43.

[14] H. Wang, G. Jin, K. G. Shin. Defense Against Spoofed IP Traffic

Using Hop-Count Filtering. [J]. IEEE/ACM Transactions on Networking, Vol 15, No 1, February 2007.
[15] L. Garber. Denial-of-service attacks rip the Internet. [J]. Computer, Apr. 2000, pp12–17.
[16] J. Elliott. Distributed denial of service attack and the zombie ant effect. [J]. IT Professional, March/April 2000, pp55–57.
[17] MIT IP Spoofer Project website. [Online]. Available: http://spoofer.csail.mit.edu/summary.php.
[18] J. Howe. (1993, June 21). An Environment for "Sniffing" DCE-RPC Traffic, CITI Technical Report 93–4. [Online]. Available: http://www.citi.umich.edu/techreports/reports/citi-tr-93-4.pdf.
[19] F. Ali. (2007, Dec.) IP Spoofing. [Online]. Available: http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_10-4/104_ip-spoofing.html.
[20] Y. Rekhter, B. Moskowitz, D. Karrenberg, et al. (1996, Feb.)Address Allocation for Private Internets. [Online]. Available: http://www.ietf.org/rfc/rfc1918.txt?number=1918.
[21] Bogon list sources. (Oct 18, 2003). [Online]. Available: http://www.completewhois.com/bogons/.
[22] S. Gibson. (2002, Feb.). Distributed reflection denial of service: description and analysis of a potent, increasing prevalent, and worrisome internet attack. [Online]. Available: Http://grc.com/dos/drdos.htm.
[23] D. Dittrich. (1999, Oct.) The "Tribe Flood Network" distributed denial of service attack tool. [Online]. Available:http://staff.washington.edu/dittrich/misc/tfn.analysis.txt.
[24] CERT Advisory CA-1997-28 IP Denial-of-Service Attacks. [Online]. Available: http://www.cert.org/advisories/CA-1997-28.html.
[25] J. Barlow, W. Thrower. (2000, Mar 7). TFN2K-An Analysis. [Online]. Available: http://www.packetstormsecurity.org/distributed/TFN2k_Analysis-1.3.txt.
[26] M. D. Ray. (1981). Transmission Control Protocol DARPA Internet Program Protocol Specification. [Online]. Available: http://www.ietf.org/rfc/rfc793.txt?number=793 .
[27] Cisco IOS, (2005). Unicast reverse path forwarding. [Online]. Available: http://www.cisco.com/en/US/docs/ios/12_2t/12_2t13/feature/guide/ft_urpf.html#wp1048141.
[28] D. Huang, Q. Cao, A. Sinha, et al. New Architecture For Intra-Domain Network Security Issues. [J]. Communications of the ACM, Nov 2006, pp64-72.
[29] Definition of man-in-the-middle. (2007, Jun 5). [Online]. Available: http://searchsecurity.techtarget.com/sDefinition/0,,sid14_gci499492,00.html.
[30] K. E. Giles, D. J. Marchette, C. E. Priebe. On the spectral analysis of backscatter data. [C]. in Hawaii International Conference on Statistics and Related Fields, 2004.
[31] D. Moore, C. Shannon, G. M. Voelker, *et al*. Network telescopes, CAIDA, Tech. Rep., 2003.
[32] J. Postel. (1981, Sep.). Internet control message protocol, DARPA Internet program protocol specification. [Online]. Available: http://www.ietf.org/rfc/rfc792.txt?number=792 .
[33] Tcpipguide. [Online]. Available: http://www.tcpipguide.com/
[34] C. Shannon, D. Moore, and E. Aben, The CAIDA Backscatter-2008 Dataset-2008-05. [Online]. Available: http://www.caida.org/data/passive/backscatter_2008_dataset.xml.
[35] F. Kargl, J. Maier, S. Schlott, et al. Protecting Web Servers from Distributed Denial of Service Attacks. [C]. WWW10, May 1, 2001.
[36] Microsoft Windows 2000 WebDAV buffer overflow vulnerability (MS03-007). [Online]. Available: http://www.securityfocus.com/bid/7116.
[37] The Spamhaus Project [Online]. Available: http://www.spamhaus.org/.
[38] An analysis of TCP/IP NetBIOS file-sharing protocols. [Online]. Available: http://www.windowsecurity.com/whitepapers/An_analysis_of_TCPIP_NetBIOS_filesharing_protocols_.html
[39] Microsoft Windows Locator Service buffer overflow vulnerability (MS03-001). [Online]. Available: http://www.securityfocus.com/bid/6666.
[40] Analysis of Network Denial of Service. (2004, Nov.) [Online]. Available: http://www.ufsdump.org/papers/uuasc-november-ddos.pdf.
[41] C. Kalt. (2000, Apr.). Internet relay chat: Architecture. [Online]. Available: http://www.ietf.org/rfc/rfc2810.txt?number=2810.

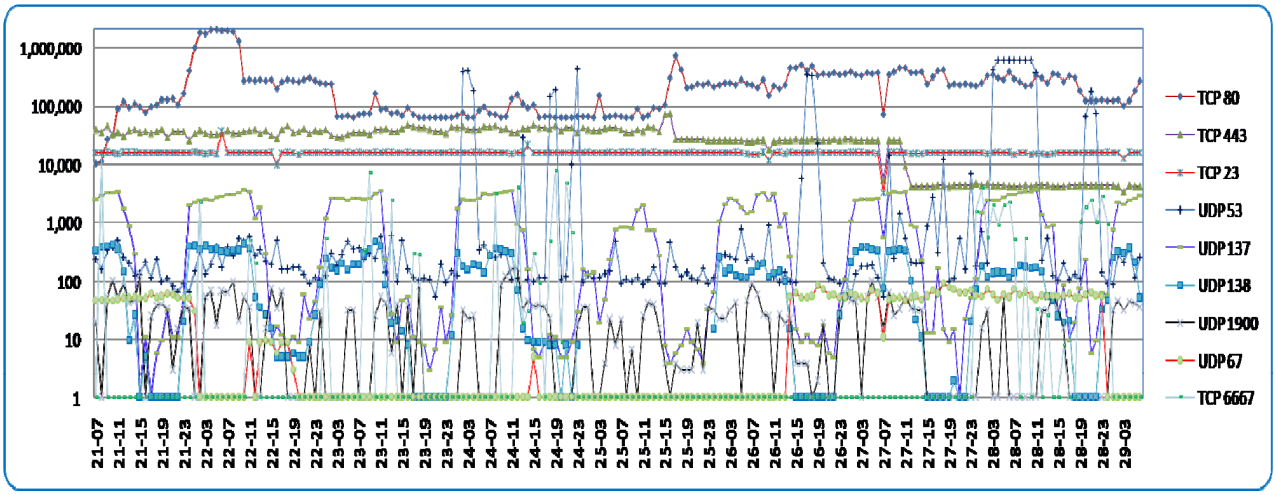

Figure 4. The distribution of TCP/UDP ports attacked by IP spoofing