Amplification DDoS Attacks: Emerging Threats and Defense Strategies

Antonio Colella, Clara Maria Colombini

Amplification DDoS Attacks: Emerging Threats and Defense Strategies

Antonio Colella[1⋆], Clara Maria Colombini[2]

[1] Italian Army and Italian Atlantic Committee, Piazza di Firenze 27, Roma, Italy
antonio.colella.it@ieee.org
[2] University of Milan, External Researcher, Milano, Italy
cmcolombini@email.it

Abstract. There are too many servers on the Internet that have already been used, or that are vulnerable and can potentially be used to launch DDoS attacks. Even though awareness increases and organizations begin to lock down those systems, there are plenty of other protocols that can be exploited to be used instead of them. One example is the Simple Network Management Protocol (SNMP), which is a common UDP protocol used for network management. Several types of network devices actually come with SNMP "on" by default. A request sent to an SNMP server returns a response that is larger than the query that came in.
The main aim of this paper is to investigate on the increasing prevalence and destructive power of amplification-based distributed denial of service (DDoS) attacks in order to present a solution based on a profiling methodology. The paper encompass three aspects: amplification DDoS attacks and main port used, the profiling methodology as a mean of identifying the threat and shape it. Finally, a proposal solution is given by considering both strategic and technical aspects.

Keywords: DDoS Attack; Amplification of DDoS Attacks; DNS; Digital Profiling

1 Introduction

A Denial of Service (DoS) attack has the main goal of preventing legitimate usage of a specific resource available on the Internet such as a web portal or any kind of network-based service. A Distributed Denial of Service (DDoS) attack is a coordinated attack against the availability of services belonging to a given target system or network. It is launched indirectly through many compromised computing systems (see Figure 1). The services under attack are those of the "primary victim", while the compromised systems used to launch the attack are often called the intermediate attack vectors or "secondary victims". The use of secondary victims in a DDoS attack provides the attacker with the ability to

⋆ Corresponding author: Antonio Colella, Italian Army and Italian Atlantic Committee advisor, Piazza di Firenze 27, I-00186, Rome, Italy. E-mail: antonio.colella.it@ieee.org

wage a much larger and more disruptive attack while remaining anonymous since the secondary victims actually perform the attack making it more difficult or almost impossible for network forensics experts to track down the real attacker (see Figure 2).

In the first quarter of year 2014, Verisign DDoS Protection Services saw an 83 percent jump in average attack size over Q4 2013, which was primarily attributed to NTP-based attacks. While DNS amplification was the most common vector in 2013 and continues to be seen, the NTP attack type is the largest attack vector seen in 2014. Verisign said that they mitigated multiple amplification attacks – commonly ranging from 50 to 75 Gbps – for their customers. Directly related to the popularity of amplification attacks was the sharp decline in more complex application-layer attacks. With the presence of so many vulnerable NTP servers and reflection vectors readily available on the Internet, attackers were able to cause maximum disruption with minimum effort, by ditching smarter application-layer attacks in favor of volume-based amplification attacks (for a more detailed study please refer to the "Verisign's Q1 2014 DDoS Trends Report" in [1]).

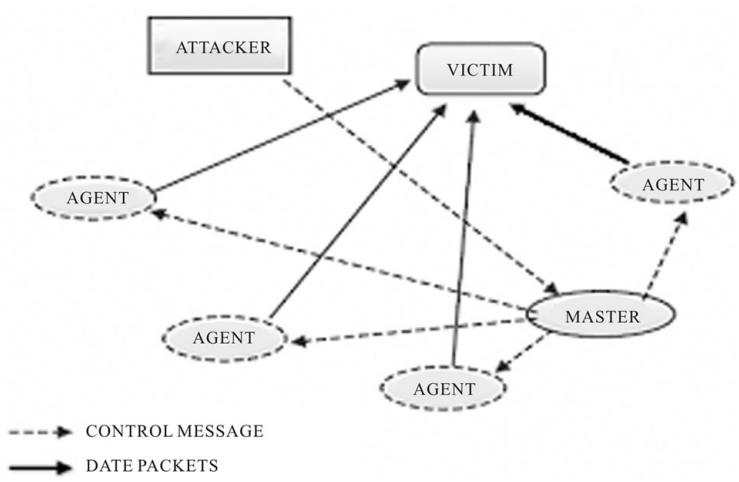

Fig. 1. DDoS Attack components

2 Amplification DDoS Attacks

Amplification DDoS attacks are based on a third-party reflection paradigm in which the real *source* of the attacks originates a significant number of service requests, implying the transmission of a certain quantity of small-sized solicitation packets, towards several completely unaware third-party entities, providing some kind of service on the Internet, and assuming the role of *reflectors*. The

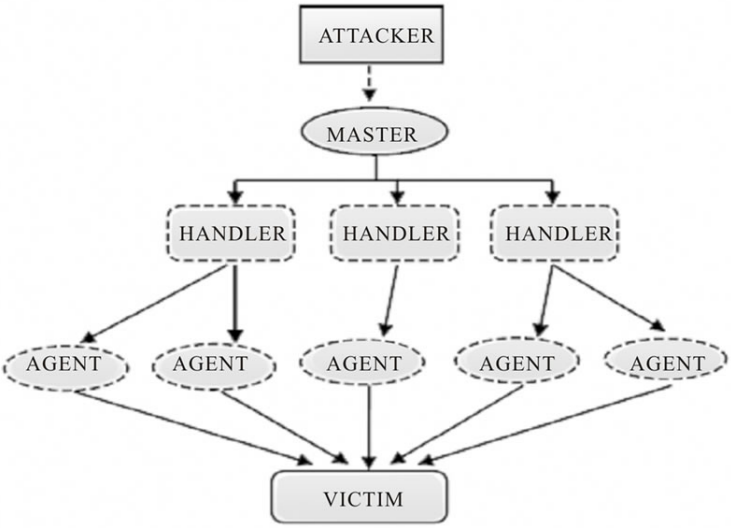

Fig. 2. DDoS Architecture

attacker properly forges the source IP address of all the requests (IP spoofing) so that they seem to come from the victim rather than from the real attack source. Furthermore, the requests are crafted to: (i) solicit a large number of individual responses (amplification through flow or reflection multiplication attacks) or (ii) obtaining response messages whose size is much larger that te one of the corresponding request (amplification through payload magnification). Reflection and Magnification can be considered as two separate attributes that can coexist in the same attack, resulting in combined amplification strategies (see Figure 4). Clearly, an attack does not get magnified unless the sizes of response messages are bigger than the request ones and analogously a flow multiplication implies that the number of attack flows received by the victim must be an integer multiple of the number of flows originally sent by the source.

In all the cases, the traffic volume generated by the third-party reflector servers in response to the individual queries received, results to be several order of magnitude higher than the one originated by the attacker alone, so that the above amplification/boosting effects greatly enhance the attack power, by overwhelming the victim's own resources both in terms of available bandwidth and (often) computing power. More formally, the bandwidth waste (in bits/s) associated to an amplification DDoS attack can be expressed as follows:

$$B = 8a \cdot S \cdot N \cdot r \tag{1}$$

where S is the request packet size, a the magnification factor characterizing the response size, N the number of simultaneous reflecting entities (flow multiplier) and r is the attack flow transmission rate (in pakets/s).

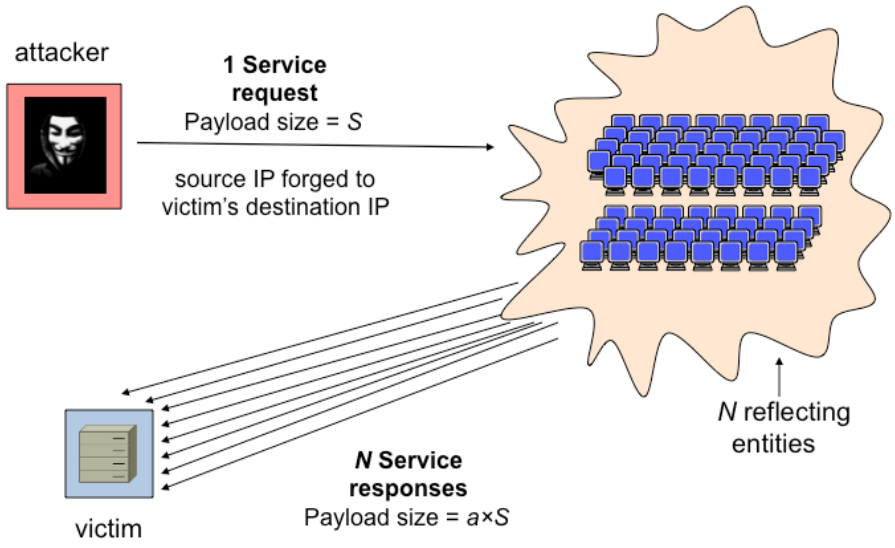

Fig. 3. Amplification DDoS attacks dynamics

One of the first attacks deploying amplification techniques over the internet is known as Smurf [2][3], exploiting in a coordinated way the ICMP replay and the direct broadcast propagation mechanisms, together with address spoofing practices, in order to completely saturate the connectivity of the target victims with huge volumes of ICMP reply messages. The great strength of this kind of attack is that it did not required the availability of a large number of compromised hosts to be successful, but only a single source machine, originating the attack flow (made of individual ICMP requests) by spoofing its own address with the one of the victim, as well as one or more third party local area networks exposing a large number of hosts on the Internet and allowing directed broadcast propagation at the gateway level. Clearly, by directly transforming the network-level broadcast packets into link-layer ones, these network provided the amplification effect by forcing each of their host to reflect the ICMP reply messages towards the attack target (whose address has been spoofed as the source of the ICMP requests). In such a way, the greater the number hosts available on the network, the higher the number of flows reflected toward the victim, with obvious consequences on its connectivity. Thus the power of the attack is proportional to the size of the intermediate "reflecting" networks, by operating according to a pure flow multiplication model ($a \approx 1 \text{and} N \gg 1$).

A very close variant of the Smurf attack, known as Fraggle [4], used UDP ECHO packets and the "Chargen" service, instead of ICMP, to implement the request/response mechanism characterizing the attack flow, but was substantially identical in all the attack dynamics based on direct broadcast propagation.

While devastating in their denial of service effects, both these attacks have been rapidly defeated by inhibiting direct broadcast propagation at the gateway level, however, the idea of relying on amplification effects offered by unaware third party systems opened new perspectives in the large-scale denial of service arena.

In particular, two widely used network services, carrying out a fundamental role in the modern Internet, have been recently exploited in order to implement successful amplification DDoS attacks: the Domain name system (DNS) protocol and the Network Time Protocol (NTP).

A DNS reflection attack is another kind of amplification-based denial of service relying on the Domain Name System query/response mechanism, and specifically exploiting situations in which the size (in terms of amount of data to be sent throughout the network) of the response to a specific DNS query, delivered to multiple servers, is much larger than the request payload ($N = 1 \text{and} a \gg 1$). Also in this case the attacker, originating the individual queries, forges (through spoofing) its source IP address in order to divert the responses incoming from all the solicited servers towards the target victim, that in turn is not able to detect the real originator of the attack. We can distinguish the reflection effects, characterized by the number of DNS servers receiving the query, from the amplification ones, whose success is proportional to the ratio between the query and response payloads. Amplification is typically achieved by soliciting specific responses involving the use of previously individuated `TXT` records. For example, specific queries sent to an authoritative DNS server can result in responses with a 10-20x amplification factor (e.g., a 40-byte size DNS query can result in a 400-byte response or greater). Attackers are able to take advantage of a huge number of "open" DNS servers available in the Internet that respond to any request sent to them. This is a quite easy technique that has proven successful in launching very large-scale attacks (i.e., several hundred Gbps in size).

Also the NTP-based attacks rely on the UDP protocol, by taking advantage of some mechanisms commonly used to synchronize the electronic clocks of computers connected to the Internet. The attack dynamics are very similar to those characterizing DNS reflection, based on soliciting the generation of very large response messages. One particularly damaging variant of the NTP attack uses the `MONLIST` command, supported in older NTP implementation, that returns the last 600 clients that an NTP server has talked to, and hence resulting in responses with an amplification factor of 10-200x with just a single NTP server. Accordingly, large scale attacks that simultaneously solicit thousands of NTP servers can produce incredible damages while requiring a very limited amount of resources on the attacker's side. For example, on February 10, 2014 about 1300 NTP servers on different networks were involved in an unprecedented cyber attack, where each server generated at peak hours approximately 90 Mb/s of traffic towards particular targets located on the Internet.

A taxonomy on the most common amplification DDoS attacks is reported in Figure 3.

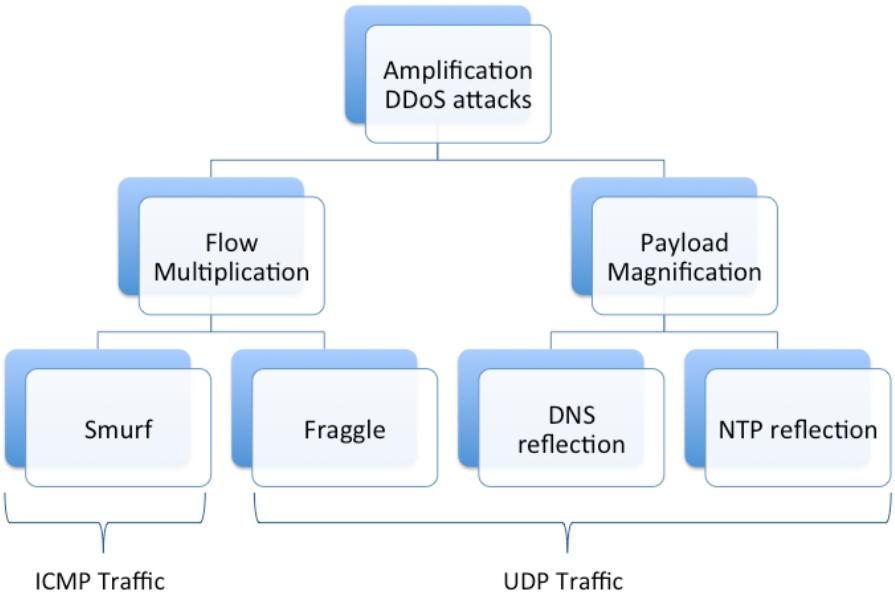

Fig. 4. Amplification DDoS attacks taxonomy

3 Profiling a DDoS Attack

One of the biggest problems of cyberdefense in general is represented by anonymity and the resulting non-imputability that cyberspace can offer to the authors of a cyber attack, as it becomes difficult, if not impossible, to identify them [5]. Chief among these is the challenge of *situational awareness* [6–8] which is defined as "the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future event" [9].

It is therefore essential to gain this view, which allows to acquire those information, more detailed and updated, useful to reach the solution. This is achieved by the method of analysis developed by Colombini and Colella in [10], where authors give a new point of view to common ICT protections, in order to recognize and prevent a DDoS attack.

3.1 The Method

The method consists of four steps:

1. threat analysis;
2. target analysis;
3. motivation analysis;

4. attack results analysis.

The study of these aspects allows to obtain a full profile of the attack, which provides new information useful for a real-time view of the presence of possible situations of DDoS attacks. It is also very important for the implementation an effective dynamic cyber defense system, that means that it can be pre-configured from time to time on specific threats that are to be fought [11].

3.2 Threat Analysis

The analysis of the behavior of the most known DDoS attacks reveals that they have the same properties of any other cyber-weapon:

- every attack is specially customized to the characteristics of the systems to hit, with the aim to reach a specific advantage;
- the implementation of the attack will be different for each attack;
- the impact of the caused damage is publicly revealed with a lag: as with all crimes, the victim is not willing to reveal his vulnerability;
- source and path are difficult to find, because their authors can take advantage of the anonymity offered by the cyberspace;
- it is often used as part of a larger conventional attack in support of it within a conflict, to gain more advantage over the enemy [12].

3.3 Target Analysis

The design of a DDoS attack takes place purely in a strategic way, in which decisions are made to define and guide the course of the attack. In the first place, the choice of the targets, that is, the type of the enemy's critical structure to hit, closely linked to the motivations of the attack, ranging from the threat to use as a deterrent to retaliation or to counterattack as a result of a suffered attack, until the attainment of some strategic advantage: economic, political, military, social [13].

It can thus describe the target with the answer to four questions:

- WHERE - physical location of the target: nation, region, town, building, up to determine whether it is a government institution, an industrial or military installation;
- WHAT - target functions: for example, a military or industrial specific process control equipment, a database that contains sensitive data, the project of a new instrumentation, etc.;
- WHO - owner and users of the target: a person, a company, a government or other groups;
- WHY - motivation of the attack: type of damage and expected results.

In this respect, it determines the type of damage to cause, which can be *digital* (exclusively aimed for delay or interruption of service) or *physical* (such as a consequently material destruction of a control system). It is measured in terms of severity of the indirect effects caused and persistence of the effects.

3.4 Motivation Analysis

The study of known DDoS attacks [14], [15], [16] confirmed that they are most often used as part of a larger conventional attack in support of it, or replacement as invisible as part of a conflict or at least one situation of tension/antagonism in place.

This observation leads to the creation of a monitoring system of political, economic and social, to extrapolate those indicators that show the possibility of an attack on available critical infrastructures, through the analysis of available information from different type of sources:

- *open* (national reports of companies of antivirus production sites, national and international news, analysis sites political, economic, social);
- *semi-open* (sites hacker circles, antagonists, extremists, fundamentalists);
- *closed* (documentation strategic-military).

3.5 Attack Effects Analysis

This step analyses the effects of a possible attack on a specific target, in relation to the information obtained from the previous steps, especially related to the time of the attack, the type of damage, the time duration of the damage on the system and the cost of the damage on the critical infrastructure.

Any impact in the short, medium and long term is evaluated in terms of:

- side effects of damage on systems/people in the environment of the target system;
- impact of the damage out of the environment of the affected system;
- side effects of damage on structures/people outside the affected system;
- social impact of the publicity of the attack;
- immediately and over time effect (military-political-social).

The information obtained by the analysis performed in the previous steps must allow to answer to the seven questions presented in Table 1.

4 The Proposed Solution

Here we propose a multi-dimensional solution, composed by the analysis of the two aspects of the problem: the technical one and the strategic one. From the technical side, analyzing the specific tool that attackers have to use, can prevent attacks first of all configuring client systems and using antivirus protection so that the attacker is unable to recruit his botnet arm, and finally configuring name servers to reduce the attacker's ability to corrupt a zone file with the amplification record. Another important action is to disable open recursion on name servers and accept only recursive DNS from trusted sources. Unfortunately, to prevent the impersonation attack fundamental is the possibility to perform source IP address validation: the botnet hosts cannot generate DNS request

WHO	Possible attackers	Type identification: - external (opposite nations, international terrorism, international organized crime) - internal (domestic terrorism, antagonist groups, organized crime)
WHY	Possible reasons	- Identification of topics of tension / crisis / antagonism in the field: -political -social -economic -military - (internal or external)
WHERE	Possible objectives	- Identification of critical infrastructures. - Identification of supersensitive data and their location.
WHAT	Damage type	- Detection of damage type: physical, digital, interruption of service, data theft.
WHEN	Attack time	- Detection of the moment of maximum vulnerability of critical infrastructure in relation to the type of attack.
RESULTS	Damage extent	- Detection of damage in relation to critical infrastructure type and damage type. - Detection of the kind disadvantage that could be caused by specific damage on specific infrastructure.
REACTION	Response actions	- Detection of the moment of maximum vulnerability of critical infrastructure in relation to the type of attack.

Table 1. The seven questions and related answers useful for the attack analysis

messages posing as the targeted name server, which stems the attack at the outset.

A new preventive way is the use of honeypots, that are systems intentionally set up with limited security to be an enticement for an attacker's hostile action. Honeypots serve to deflect attacks from hitting the systems they are protecting as well as serving as a means for gaining information about attackers by storing a record of their activity and learning what types of attacks and software tools the attacker is using. The goal of this type of honeypots is to lure an attacker to install either handler or agent code within the honeypot, thereby allowing the honeypot's [17] owner to track the handler or agent behavior and better understand how to defend against future DDoS installation attacks. Honeypots are also helpful because they can store event logfiles during a DDoS attack.

Data can be analyzed post-attack to look for specific characteristics within the attacking traffic. To help identify the attackers, tracing Internet traffic back to it's source helps to identify the authors. Additionally, when the attacker sends different types of attacking traffic, this method assists in providing the victim system with information that might help develop filters to block the attack. An example of countermeasure in which network profiling can be effective has been reported in the Table 2 .

Countermeasures *(what)*	**Actions** *(in what way)*	**Aim** *(why)*
detect and neutralize handlers	analyzing the content of source code and detected feature in the malware behavior	comprehension of communication protocol and traffic among handlers, clients, and agents.
detect/prevent secondary victims	heightened awareness of security issues and prevention techniques from all Internet users	to prevent themselves from participating in the attack.
detect/prevent potential attacks	egress filtering	scanning of IP packet headers leaving a network and checking to see if they meet certain criteria.
mitigate/stop attacks	load balancing	network providers can increase bandwidth on critical connections to prevent them from going down in an attack.
mitigate/stop attacks	throttling	the Max-min Fair server-centric router throttle method. This can prevent flood damage to servers.
deflect attacks	honeypots	gaining information about attackers by storing a record of their activity.
post-attack forensics	traffic pattern analysis	these techniques help identify the attackers tracing Internet traffic back to it's source.
post-attack forensics	packet traceback	these techniques help identify the attackers tracing Internet traffic back to it's source.

Table 2. Countermeasures

To work as better, this solution must be based on information obtained by the application of the digital profiling method, explained in Section 3. The result of the study provides information on:

- critical infrastructure as possible target;
- list of exploitable vulnerabilities;
- time of the attack;
- extent of expected damage;
- origin of attack;
- the identity of the attacker/s.

The obtained information can then be used to implement a series of effective countermeasures, for protection and prevention of a DDoS attack, by improving the physical protection of critical infrastructure, as well as the resolution of digital vulnerability of critical systems and implementation of new security policies for access to critical data [18]. From the point of view of offensive defense, the

real-time view of the presence of possible situations of attack allows the implementation of an effective system of cyber defense in a dynamic way, every time capable of pre-self-configure depending on the eventual threats from time to time detected, and to implement countermeasures specifically proportionated to the threat in act [11]. In the analysis of log files, for instance, depending on the purpose it is intended, are used to highlight relationships between data and build a result of behavioral models that describe two types of approach:

- Top down: search for confirmation of facts already known or assumed (e.g., an action resulting from an intrusion has already occurred)
- bottom-up: useful to find information useful to construct hypotheses (e.g., the most likely causes that produce a particular result).

As we have explained in a previous paper [17] the analysis cycle takes place in 6 steps (see Figure 5)

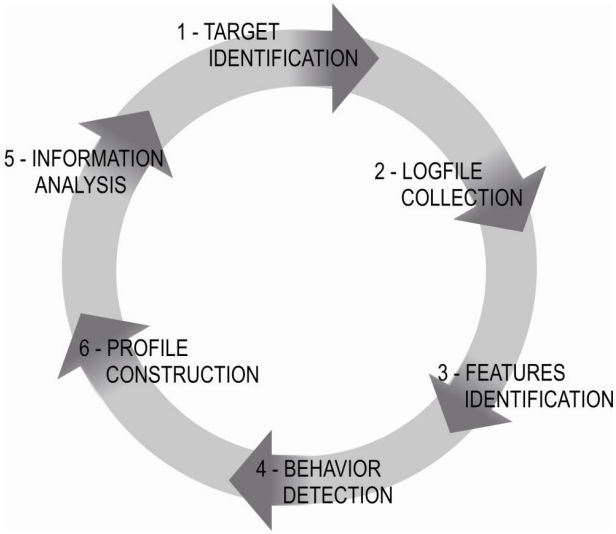

Fig. 5. network profiling cycle of analysis

Step 1 - Identification of the target.
Step 2 - Collection of data log files.
Step 3 - Identification of characteristic properties (features) from the mass of data collected from log files and collect this information (indicators) contained the features detected.
Step 4 - Detection of possible subjects to which it is possible to attribute behavior Digital.

Step 5 - Analysis of information and construction of the behavior of digital accesses.

Step 6 - Construction of the user profile and usage of digital information obtained, depending on the objective.

5 Conclusions and Future Work

Countering amplification DDoS attacks is an important security issue to be faced with in modern network-empowered organizations. In this work we have presented an alternative defense strategy, obtaining information from profiling techniques that can be used to implement a series of effective protection and prevention countermeasures against such attacks. The intent of the authors for the future is to deeply analyze this approach trying to find a full set of indicators able to foster a complete and effective prevention methodology.

References

1. Verisign Inc.: Verisign Distributed Denial of Service Trends Report, issue 1 ? 1st quarter 2014. http://www.verisigninc.com/en_US/website-availability/ddos-protection/ddos-report/index.xhtml
2. Kumar, S.: Smurf-based distributed denial of service (ddos) attack amplification in internet. In: Internet Monitoring and Protection, 2007. ICIMP 2007. Second International Conference on, IEEE (2007) 25–25
3. CC, C.: Smurf IP Denial-Of-Service Attacks - CERT ADVISORY CA-1998-01. http://www.cert.org/advisories/CA-1998-01.html (2000)
4. Specht, S.M., Lee, R.B.: Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures. In: ISCA PDCS. (2004) 543–550
5. Schreier, F.: On Cyberwarfare, DCAF Horizon 2015 Working Paper Series (7). http://www.dcaf.ch/Publications/On-Cyberwarfare (2012)
6. Fenza, G., Furno, D., Loia, V., Veniero, M.: Agent-based Cognitive approach to Airport Security Situation Awareness. 2010 International Conference on Complex, Intelligent and Software Intensive Systems **0** (2010) 1057–1062
7. Furno, D., Loia, V., Veniero, M., Anisetti, M., Bellandi, V., Ceravolo, P., Damiani, E.: Towards an agent-based architecture for managing uncertainty in situation awareness. In: Intelligent Agent (IA), 2011 IEEE Symposium on. (April 2011) 1–6
8. De Maio, C., Fenza, G., Furno, D., Loia, V.: Swarm-based semantic fuzzy reasoning for situation awareness computing. In: Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on. (June 2012) 1–7
9. Vidulich, M., Dominguez, C., Vogel, E., McMillan, G.: Situation awareness: papers and annotated bibliography - Armstrong Laboratory, Human System Center, ref. AL/CF-TR-1994-0085. http://www.dtic.mil/dtic/tr/fulltext/u2/a284752.pdf (1994)
10. Colombini, C., Colella, A.: Digital Profiling: A Computer Forensics Approach. In Tjoa, A., Quirchmayr, G., You, I., Xu, L., eds.: Availability, Reliability and Security for Business, Enterprise and Health Information Systems. Volume 6908 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2011) 330–343

11. Colella, A., Colombini, C.M.: Cyber-space, Cyberware, Cyber-weapons. In Attanasio, A., Costabile, G., eds.: IISFA MEMBERBOOK 2012 DIGITAL FORENSICS. Experta Edizioni, (in Italian) (2012)
12. Colombini, C., Colella, A., Mattiucci, M., Castiglione, A.: Network Profiling: Content Analysis of Users Behavior in Digital Communication Channel. In Quirchmayr, G., Basl, J., You, I., Xu, L., Weippl, E., eds.: Multidisciplinary Research and Practice for Information Systems. Volume 7465 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2012) 416–429
13. Colombini, C., Colella, A.: Digital scene of crime: technique of profiling users. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA) **3**(3) (9 2012) 50–73
14. Palmieri, F., Fiore, U.: Network anomaly detection through nonlinear analysis. Computers & Security **29**(7) (2010) 737 – 755
15. Palmieri, F., Fiore, U., Castiglione, A.: A distributed approach to network anomaly detection based on independent component analysis. Concurrency and Computation: Practice and Experience **26**(5) (2014) 1113–1129
16. Fiore, U., Palmieri, F., Castiglione, A., De Santis, A.: Network anomaly detection with the restricted Boltzmann machine. Neurocomputing **122**(0) (2013) 13 – 23
17. Colombini, C., Colella, A., Mattiucci, M., Castiglione, A.: Cyber Threats Monitoring: Experimental Analysis of Malware Behavior in Cyberspace. In Cuzzocrea, A., Kittl, C., Simos, D., Weippl, E., Xu, L., eds.: Security Engineering and Intelligence Informatics. Volume 8128 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2013) 236–252
18. Colella, A., Colombini, C.M.: La rete e le informazioni, raccolta e uso illecito dei dati. In Attanasio, A., Costabile, G., eds.: IISFA MEMBERBOOK 2011 DIGITAL FORENSICS. Experta Edizioni, (in Italian) (2012) 201–220