Swarm and Evolutionary Computation xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Swarm and Evolutionary Computation

journal homepage: www.elsevier.com/locate/swevo

Adaptive artificial immune networks for mitigating DoS flooding attacks

Jorge Maestre Vidal*, Ana Lucila Sandoval Orozco, Luis Javier García Villalba

Group of Analysis, Security and Systems (GASS), Department of Software Engineering and Artificial Intelligence (DISIA), School of Computer Science, Office 431, Universidad Complutense de Madrid (UCM), Calle Profesor José García Santesmases s/n, Ciudad Universitaria, 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Anomalies
Artificial immune system
Denial of Service
Forecasting
Intrusion detection system
Network

ABSTRACT

Denial of service attacks pose a threat in constant growth. This is mainly due to their tendency to gain in sophistication, ease of implementation, obfuscation and the recent improvements in occultation of fingerprints. On the other hand, progress towards self-organizing networks, and the different techniques involved in their development, such as software-defined networking, network-function virtualization, artificial intelligence or cloud computing, facilitates the design of new defensive strategies, more complete, consistent and able to adapt the defensive deployment to the current status of the network. In order to contribute to their development, in this paper, the use of artificial immune systems to mitigate denial of service attacks is proposed. The approach is based on building networks of distributed sensors suited to the requirements of the monitored environment. These components are capable of identifying threats and reacting according to the behavior of the biological defense mechanisms in human beings. It is accomplished by emulating the different immune reactions, the establishment of quarantine areas and the construction of immune memory. For their assessment, experiments with public domain datasets (KDD'99, CAIDA'07 and CAIDA'08) and simulations on various network configurations based on traffic samples gathered by the University Complutense of Madrid and flooding attacks generated by the tool DDoSIM were performed.

1. Introduction

By definition, Denial of Service (DoS) has the objective of disabling computer systems or networks. The DoS attacks with origin in multiple sources are referred as Distributed Denial of Service (DDoS) attacks. In recent years, the number of incidents related with these threats reported by the various organizations for cyber defense shows an alarming growth. According to the European Network and Information Security Agency (ENISA), between 2013 and 2014 an increase of 70% was observed [48]. In addition, they pose a threat that has begun to be used in order to achieve other objectives. These include disguising activities in relationship with malware spreading, concealment of fraudulent money transfers [61] or compromising anonymous networks, such as Tor or Freenet [35]. This growth is attributed to various reasons: the first of them is that DDoS attacks are usually triggered by previously infected systems, which in most of the cases are part of botnets. The botnets have been adapted to be resilient against the classical detection schemes, thus allowing the construction and maintenance of larger collections of zombies and increasing their difficulty to be identified. [60]. Another important reason is that attackers are able to take advantage of amplifying elements, in this way enhancing their potential to be harmful. To do this, they exploit vulnerabilities in protocol implementations on the intermediate network devices, particularly at DNS, NTP and SNMP. On the other hand, as the European Police Office (Europol) warns [25], the DDoS is becoming increasingly linked with the organized crime. Rent botnets for execution of these attacks is a very profitable business on the black market, often supplied as Crimeware-as-a-Service (CaaS). Finally, offenders with lack of formation have a wide variety of tools for easily configuration and deployment of flooding attacks. The black market also offers technical support, a situation that expands the range of user profiles which are able to attack with success.

The most common DDoS methods are based on flooding. Because of this, they are the principal object of study in this research. The flooding attacks modus operandi involves the injection of large volumes of traffic in order to saturate the victim systems [70]. The popularity of this group of DoS attacks is mostly due to cheap price and simple implementation, in comparison with the good results they provide. Thus, there are lots of proposals aimed at their detection and mitigation. However few of them meet all the requirements to be effective in real use cases, emphasizing the needs of high true positive rates, unrepresentative false positive rates, low consumption of computational resources and real time performance. In addition it is noteworthy that proposals of the literature seldom consider advantages

* Corresponding author.
E-mail addresses: jmaestre@ucm.es (J.M. Vidal), asandoval@fdi.ucm.es (A.L.S. Orozco), javiergv@fdi.ucm.es (L.J.G. Villalba).

http://dx.doi.org/10.1016/j.swevo.2017.07.002
Received 23 November 2016; Received in revised form 5 May 2017; Accepted 3 July 2017

of the new trends on networking. It is expected that the emerging networks, taking the example of 5G, increasingly move towards self-management. The use of novel technologies such as Software-defined networking (SDN), Network-Function Virtualization (NFV), Artificial Intelligence or Cloud computing, facilitates the design of Self-Organizing Networks (SON). This should encourage the appearance of more interesting proposals that are capable of reacting with a much more comprehensive view of the problem being treated.

To serve this cause, a strategy for detection and mitigation of DDoS flooding attacks is proposed. Therein the deployment of a sensor network that integrates an Artificial Immune System (AIS) inspired by the biological defense mechanisms of human beings is introduced. Unlike similar proposals, conventional bio-inspired methods for pattern recognition were not applied. Instead a combination of strategies for DDoS detection based on the study of variations of the entropy on the network traffic by thresholding, with the adaptation of the biological immune reactions is proposed. This makes it possible to apply real time countermeasures, building an immune memory and establishment of quarantine areas, all in accordance with the current state of the protected network.

In view of this, it is important to highlight the two major contributions of this paper: firstly, a new method for detecting DDoS that is able to forecast anomalies on the entropy of the traffic analyzed, and thereby recognition of flooding attacks is introduced. This is performed by representation of the entropy in time series and the definition of prediction intervals. It has been evaluated considering public domain datasets (KDD'99 [37], CAIDA'07 [15] with CAIDA'08 [16]) samples, and traffic monitored in the University Complutense of Madrid (UCM) and flooding attacks generated by the tool DDoSIM [22]. The first two allow its comparison with previous detection approaches, and the latter offers a more realistic view of its behavior. The preliminary experiments showed promising results, which motivates the development of a cooperative deployment strategy. This is the second main contribution, where a method for management of immune agents that implements the previously described detection system is proposed. Within this, the decisions are made as to when and how they will act and in what level of restriction, all this depending on the status of the network and orchestrated by an artificial immune approach. To evaluate the effectiveness of the deployment of the AIS, a simulator capable of generating traffic distributions and different networks with various locations of the sensors acting as immune agents has been implemented. In the generation of new networks, several parameters had been taken into account: number of nodes, legitimate traffic volume, branching component, and cyclic component; demonstrating in all of them the improvements offered by the cooperation and the emulation of the defensive capabilities of the biological immune systems.

The paper is divided into seven sections, and the first of them is the present introduction. The background necessary for a better understanding of the approach is described in Section 2. The proposed AIS is introduced in Section 3. The novel DDoS detection method implemented in the various agents of the AIS is detailed in Section 4. Experiments, datasets and methodology are described in Section 5. Results are discussed in Section 6. Finally, conclusions and future work are presented in Section 7.

2. Background

The following describes all aspects necessary for understanding the proposal. Among them it is important to highlight those involving the characteristics of the DDoS flooding attacks and their countermeasures, a general review of the human being immune system and the different approaches with this in mind, in order to provide defenses against cybercrime.

2.1. Flooding threats: Attacks and countermeasures

According to [67], there are two types of traffic injection able to compromise a system or network by flooding. The first one is based on the constant and continuous generation of large volumes of information, and is well known as high rate flooding. This is a method which is usually very visible that easily overflows the computing capacity of the victim. On the other hand, the victim may be compromised by less noisy attacks, which are able to exploit vulnerabilities in the various communication protocols. They are known as low rate flooding attacks, using as a typical example, the attacks with On/Off patterns addressed against the TCP [62,42]. In both cases, the malicious traffic may be sent to the victim in a direct or reflected way [12,6]. There are different ways of exploiting the capacity of the flooding attacks, as is the case of the link-flooding and target link-flooding attacks [66,28]. They are based on depleting the bandwidth of the victim by aiming to certain network links or target regions. Since they concentrate the flood at specific points of the network and are able to reuse legitimate traffic to congest the victim, they are often much more difficult to detect than the conventional threats. Most efforts of the community to deal with denial of service assume these behaviors, and from them different methods for detection, mitigation and identification of sources are proposed. Their most important features and evaluation schemes are described below.

The detection of the attacks is often necessary to conduct either of the other defensive reactions. This is based on the analysis of traffic that flows through the protected environment, looking for signatures of previously known attacks or anomalous behaviors. For that purpose various techniques have been proposed, such as probabilistic models based on Markov [58], Genetic Algorithms (GA) [41], Chaos theory [20], CUSUM statistical analysis with wavelet transforms [17], forensic methods based on visualization [14], SVM (Support Vector Machines) [3,5], k-means [31], decision trees [52,53], artificial neural networks [43,59], fuzzy logic [40] or the study of variations on the traffic entropy [51,12]. In approaches like [72], the problem of the similarity of the nature of the DDoS attacks in comparison with non-malicious events is studied. This is the case of the well-known situations such as flash crowds, which often occur when a large amount of legitimate users converge on a certain service in a short time interval.

The total or partial reduction of the damage inflicted by the attacks is defined as mitigation. To do this, it is common to use honeypots [33], puzzles that recognize non-human users [73], bandwidth enlargement [38], reputation-based schemes [45], filtering or the adoption of security protocols such as IPsec [25]. As can be observed, the set of mitigation actions may also contain prevention strategies. These are characterized by not having direct dependence on the attack detection [46].

To find the compromised systems from which the malicious traffic is originated is referred to as the identification of their sources. Ideally, its objective is to track the cybercriminal. However, given the administrative difficulties that this process involves (different Internet Service Providers (ISPs), proxies, data privacy legislations, etc.), and the recent advances on footprint occultation, this goal is often very hard to carry out successfully. Consequently, many of the proposals in the literature just focus on getting as close as possible to the attacker, in order to sanitize the largest amount of regions within the protected network. On the identification of sources, the packet traceback is the most frequent approach. In [4] this issue is discussed, a lot of current approaches are collected, and a new scheme for uniform tracking is proposed. The Passive IP Traceback (PIT) that bypasses the deployment difficulties of the conventional IP traceback techniques by investigation of ICMP error messages is proposed in [69]. Finally, in [36,69] the influence of the characteristics of the network topology on the effectiveness of the strategies for packet marking is studied.

The evaluation of defense systems against DoS/DDoS is a controversial issue today. Over the years, various collections of traffic samples and methodologies for assessment of these tools were pro-