Fingerprinting Internet DNS Amplification DDoS Activities

Claude Fachkha, Elias Bou-Harb, Mourad Debbabi
Computer Security Laboratory, CIISE, Concordia University & NCFTA Canada
Montreal, QC, Canada
{c_fachkh, e_bouh, debbabi}@encs.concordia.ca

***Abstract*—This work proposes a novel approach to infer and characterize Internet-scale DNS amplification DDoS attacks by leveraging the darknet space. Complementary to the pioneer work on inferring Distributed Denial of Service (DDoS) activities using darknet, this work shows that we can extract DDoS activities without relying on backscattered analysis. The aim of this work is to extract cyber security intelligence related to DNS Amplification DDoS activities such as detection period, attack duration, intensity, packet size, rate and geo-location in addition to various network-layer and flow-based insights. To achieve this task, the proposed approach exploits certain DDoS parameters to detect the attacks. We empirically evaluate the proposed approach using 720 GB of real darknet data collected from a /13 address space during a recent three months period. Our analysis reveals that the approach was successful in inferring significant DNS amplification DDoS activities including the recent prominent attack that targeted one of the largest anti-spam organizations. Moreover, the analysis disclosed the mechanism of such DNS amplification DDoS attacks. Further, the results uncover high-speed and stealthy attempts that were never previously documented. The case study of the largest DDoS attack in history lead to a better understanding of the nature and scale of this threat and can generate inferences that could contribute in detecting, preventing, assessing, mitigating and even attributing of DNS amplification DDoS activities.**

I. Introduction

A DDoS attack is one of the major cyber attacks that attempts to make a computer or network resources unavailable. DDoS activities, indeed, dominate today's attack landscape. In a recent report by Arbor Networks [1], it was concluded that 48% of all cyber threats are DDoS. A DNS amplification attack is a well known practice of a DDoS, in which malicious users abuse open DNS servers to bombard a victim with DNS reply traffic [2]. A recent event demonstrated that even a cyber security organization became a victim of the largest (i.e., 300 Gbps) DNS amplification DDoS attack in history [3]. The above facts concur that DDoS attacks in general, and DNS amplification in particular, are and will continue to be a significant cyber security issue, causing momentous damage to a targeted victim as well as negatively affecting, by means of collateral damage, the network infrastructure (i.e., routers, links, etc.), the finance, the trust in, and the reputation of the organization under attack. In this work, we tackle the following questions: 1) How to infer large-scale DNS amplification DDoS activities? 2) what are the characteristics of DNS amplification DDoS attacks? and 3) what inferences can we extract from analyzing DNS amplification DDoS traces?

In this context, we frame this paper's contributions as follows: 1) Proposing a systematic flow-based approach for inferring DNS amplification DDoS activities by leveraging DNS queries to darknets; 2) Characterizing the inferred DNS amplification DDoS threats during a recent 3 months period; and 3) Analyzing traces from the largest DNS amplification in history [4] and uncovering the mechanism behind them.

The remainder of this paper is organized as follows: In Section II, we survey the related work. In Section III, we provide an overview and background information on DNS amplification attacks and darknet space. In Section IV, we present the proposed approach and elaborate on various aspects of its components. In Section V, we empirically evaluate the approach and disclose a case study on the largest DNS amplification DDoS attack. Finally, in Section VI, we summarize the paper, discuss the lessons learned and list the future work.

II. Related work

Cyber security experts and researchers employ darknet analysis for several purposes such as monitoring of large-scale Internet events, including, DDoS [5, 6] and probing activities [7, 8]. Since our work deals with cyber threats in general and DNS amplification DDoS in particular, we subsequently pinpoint the major related work in the areas of backscattered traffic analysis and DNS traffic investigation.

First, the use of darknet to infer DDoS activities owes much to the pioneer work carried out by Moore et al. in [5] that was revisited in [9]. The key observation behind the authors' technique is that attackers, before executing a DDoS attack, spoof their addresses using random IPs. Hence, once the attack is executed, all the victims' replies (i.e., backscattered packets) are bounced back to the fake IP addresses, which could be in the monitored darknet space. Their work is operated by CAIDA [10], which provide backscattered data for researchers. Numerous research work has been performed on such data to analyze DDoS activities. The majority focus on implementing new detection techniques to infer DDoS attacks [11], tracing-back the sources of attacks [12], investigating spoofed attacks [13] and visualizing attacks [14]. Our work is different from this category as their dataset is only based on reply packets and do not include request packets such as DNS queries. Hence, DNS amplified activities may not be inferred using their approach.

Second, in the area of DNS traffic analysis, the most related work in this area is rendered by Oberheide et al. [15] who analyze DNS queries that target darknet sensors. The authors characterize these traces and propose a mechanism to implement a secure DNS service on darknet sensors. Moreover, Paxson [16] is among the first to pinpoint the threats of DNS reflectors on making DDoS attacks harder to defend. In another work, Dagon et al. [17] analyze corrupted DNS resolution paths and pinpoint an increase in malware that modified these paths and threatened DNS authorities. In comparison to our

978-1-4799-3223-8/14/$31.00 ©2014 IEEE

work, Oberheide et al. have not linked or investigated any DNS DDoS traces through their analysis but solely focused on analyzing DNS traffic. On the other hand, Paxson and Dagon et al. did not investigate darknet data. Therefore, all DNS amplification traces destined to unused IP addresses (darknet) cannot be detected through their analysis. However, darknet and other sources of data (i.e., Pasive DNS) could be associated to extract further intelligence on DNS amplification DDoS activities such as the approximate number of infections. Future work could consider the latter task.

III. Background

In this section, we provide some background information related to the mechanism of DNS amplification attacks, the darknet space and DNS queries targeting the darknet.

A. DNS Amplification DDoS Attacks

A DNS amplification attack is a well known practice of a DDoS, in which malicious users abuse open DNS servers to bombard a victim with DNS reply traffic [2]. The technique consists of an invader directing a DNS name lookup query to an open DNS server having the source IP spoofed to be the victim's address. Subsequently, all DNS server responses will be sent to the targeted victim. In general, malicious users will request domains that cover a large zone to increase the amplification factor. In order to have a high impact on the victim, the attackers use DNS requests of type `ANY` to returns all possible known information to the victim, and hence increase the amplification of the attack. Moreover, in order to increase the size of the attack with little effort, attackers use botnets to synchronize an army of bots and order them to send the DNS requests.

B. Darknet Space

In a nutshell, darknet traffic is Internet traffic destined to routable but unused Internet addresses (i.e., dark sensors). Since these addresses are unallocated, any traffic targeting such space is suspicious. Darknet analysis has shown to be an effective method to generate cyber threat intelligence [18, 19]. Darknet traffic is typically composed of three types of traffic, namely, scanning [20], backscattered and misconfiguration [21]. Scanning arises from bots and worms while backscattered traffic commonly refers to unsolicited traffic that is the result of responses to DDoS attacks with spoofed source IP addresses. On the other hand, misconfiguration traffic is due to network/routing or hardware/software faults causing such traffic to be sent to the darknet sensors.

C. DNS Queries on Darknet

On the darknet space, we observe a significant number of DNS queries that could be sent by the following sources: 1) Victim of Spoofed IP when the attacker sends spoofed DNS queries on the Internet address space using the victim's IP address. All replies from the open DNS resolvers will bounce back towards the victim; 2) Compromised victim when the attacker uses the victim's machine to send DNS queries. The attacker might use several techniques to control the victim's machine, including malware infection and/or vulnerability exploitation. This scenario do not involve spoofed DNS queries; 3) Scanner when the attacker scans the Internet to infer the locations of open DNS resolvers. This task requires collecting information from the reply packets and hence, a non-spoofed address is used by the scanners.

In our work, we assert that high speed `ANY` DNS queries will be sent from a victim of spoofed IP or/and compromised victim but not from a scanner. In other words, scanners might send `ANY` DNS queries to the Internet but with low-speed rate to avoid receiving the amplified flood of replies.

IV. Proposed Approach

This section presents and elaborates on our proposed approach that aims at inferring DNS amplification DDoS activities by leveraging darknet data. *The approach exploits the idea of analyzing DNS queries that target the darknet space that were originally intended by the attacker to reach Internet open DNS resolvers*. The approach takes as input darknet traffic and outputs inferred DNS amplification DDoS insights. It is based on 2 components, namely, the detection and the rate classification components. We discuss these components in what follows.

A. Detection Component

The detection component takes as input darknet traffic and outputs DNS amplification DDoS flows. A flow is defined as a series of consecutive packets sharing the same source IP address targeting darknet addresses. To achieve the detection task, we base our detection component on analyzing DNS queries targeting darknet addresses. These DNS queries are attempts towards port 53. In order to detect DNS amplification DDoS, we built our approach in accordance with the parameters of Table I. In this work, we build on top of [9] that

Parameter	**Value**
Packet Count	> 25
Scanned Hosts	> 25
DNS Query Type	ANY
Requested Domain	Found in Root_DNS_DB

TABLE I: DNS amplified DDoS Identification Parameters

inferred DDoS from darknet; we use 25 as a threshold for both the packet count and scanned hosts. This permits the filtering of misconfiguration traffic (i.e., a host sending many packets to only 1 unused IP address). Moreover, this verifies that the inferred DNS amplified attempts involve at least 25 distinct open DNS resolvers. Note that, we could have also added other parameters such as $attack\text{-}duration$ and $packet\text{-}rate$ to our detection component. However, we avoid using time-based constraints; we have detected some flash attempts [22] that targeted thousands of distinct unused IPs within seconds and other stealthy scanning activities [23] that persisted for several weeks.

In summary, our detection component labels a flow of traffic as a DNS amplification DDoS attack if it has sent at least 25 DNS query of type `ANY` to distinct unused dark IP addresses. Further, the flow must have requested domains that exist in our root and TLD database.

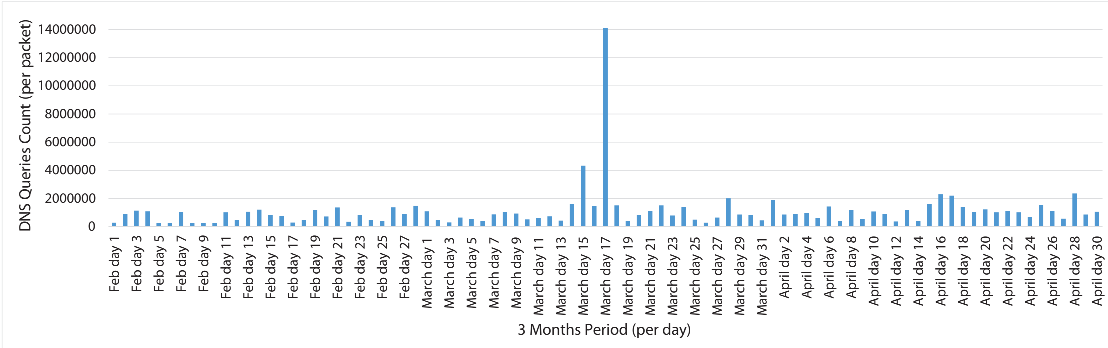

Fig. 1: DNS Queries Distribution of February, March and April 2013

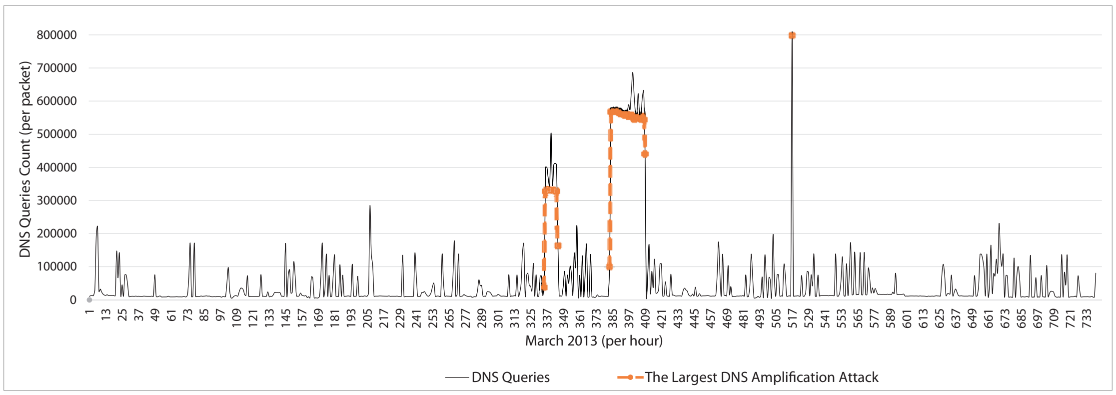

Fig. 2: DNS Queries Distribution of March 2013

B. *Rate Classification Component*

The rate of the attack is one of the major characteristics of DDoS activities [9]. After inferring DNS amplification flows, we noticed the existence of a large deviation among DNS amplification DDoS attack rates. For example, some flow rates reached more than 50 thousand packets per second (pps) whereas others were below 1 pps. Therefore, in order to understand more this large deviation and to group attacks per attack rates, we executed a rate classification exercise based on the values found in Table II. We have chosen a threshold of

Attack Rate Category	Value (pps)
Low	rate ≤ 0.5
Medium	$0.5 <$ rate < 4700
High	rate ≥ 4700

TABLE II: Attack Classification per Rate

0.5 pps to differentiate between low and medium attacks. This value is used in [9]. However, instead of neglecting such low rate attacks, similar to what [9] did, we adopt and exploit this value to detect a sub-category of attacks; stealthy attempts [23], which consists of slow scans that are generally hard to detect through flow-based detection and intrusion detection systems. In regards to high attack rate, this category contains high rate attempts that are commonly referred to as flash attacks [22]. We have chosen a threshold of 4700 pps, which is the average rate of the Slammer worm propagation [22], to differentiate between medium and high rate attacks. In this exercise, we assume that the average rate of the fastest worm propagation in 2003 will have, at least, similar rates as flash attacks in 2013. Please note that in general, worm propagation performs scans for vulnerabilities on hosts in an attempt to exploit or infect the victims. In comparison to the DNS amplification DDoS attempts, the attackers generate a similar portsweep propagation attempts to detect open DNS resolvers and execute the attack in one shot. The latter technique does not aim at searching for a vulnerability to exploit, but instead sends benign DNS `ANY` queries to abuse the open DNS resolver services and amplify the reply on a victim.

V. Empirical Evaluation

The evaluation is based on three real darknet datasets that respectively represent a 3 months period, namely, February, March and April, 2013. In general, we possess real darknet data that we receive on a daily basis from a trusted third party. The darknet sensors monitor /13 address blocks (i.e., $\approx$ half a million dark IPs). The analyzed data consists of an average of 720 GB of one-way communications to unused IPs. In regards to our characterization tasks, we used several network-based monitoring and statistical tools such as TCPdump. Note that our DNS amplification inference approach is capable of processing and inferring attacks in around 90 seconds per 20 GB of darknet traffic. The latter advocates that the proposed approach is practically viable in operational environments. We abide and closely follow the steps of our proposed approach that was discussed in Section IV to elaborate on our analysis, which is based on three main elements, namely, the characterization and a case study. In total, during the 3 months period, our approach identified a total of 134 (29 in February, 54 in March and 51 in April) DNS amplification DDoS attacks including high-speed, medium and stealthy attacks. It is significant to note that the findings that are based on our empirical analysis hold throughout the three datasets. In fact, the findings are coupled with Internet DNS amplification attack events and hence are independent from those specific datasets. Further, in this work, we exclusively focus on one case study; the largest DDoS attack in history as elaborated in Section V-B.

A. DNS Amplification DDoS Characterization

In this section, we present the overall DNS amplification DDoS statistics related to our analyzed dataset. The overall DNS queries distribution is shown in Figure 1. The outcome clearly fingerprints the largest DNS amplified DDoS attack that occurred in March 2013 [4]. On the other hand, in order to have a closer look at this attack, we depict Figure 2 that illustrates the distribution of the queries for the month of March. The average DNS queries arrival time per hour is approximately 58050 packets. Obviously, several large-scale DNS Amplified DDoS attacks caused some peaks at some periods such as at hours 340, 400 and 517 in which the distribution of packets was raised to 503995, 686774 and 798192 packets, respectively. More explanation on these peaks are discussed in Section V-B.

1) Query Type Distribution: In order to understand the types of DNS queries received on our dark space, we list in Table III the DNS query type distribution of the analyzed dataset. As expected, the vast majority of these are ANY queries. Note that the top 4 records are the same for the entire 3 months period. Further, in contrast with the results in 2007 by [15], that found that ANY records scored only 0.0199% of the entire perceived records, we record 52.23% as observed on the darknet space. As a result, we can safely assume that the recent trend of DNS amplification attacks are behind the increase of ANY records found on the darknet in the current year [4].

February Packet Count (%)	**March Packet Count (%)**	**April Packet Count (%)**
10047038 A (49.02%)	27649274 ANY (64.23%)	18378685 ANY (54.60%)
7763817 ANY (37.88%)	11310058 A (26.28%)	11595908 A (34.45%)
2479572 TXT (12.10%)	2459257 TXT (5.71%)	3402073 TXT (10.11%)
100463 MX (0.49%)	500143 MX (1.16%)	180779 MX (0.54%)
29232 PTR (0.14%)	63340 RRSIG (0.15%)	28716 AAAA (0.09%)

TABLE III: Top 5 DNS Query Type Distribution of 3 Months Period

2) Requested Domains: In our analysis, we found that Root is the most requested domain name as perceived by the monitored darknet. Recall that malicious users will request domains that cover a large zone to increase the amplification factor. Note that, from our data, the second top requested domain belongs to one of the largest Internet-scale DNS operators.

B. Case Study

Out of the 134 detected attacks, we discuss in this section one of the major case studies that belong to a medium speed attack. The latter is one of the major inferred DNS amplification DDoS in terms of size and impact. This attack targeted one victim using 2 hosts (ID M1 and M2 of Table IV). This attack scanned around 360000 unique dark IPs (68% of the monitored /13 darknet), and hence could be considered the most comprehensive compared to all other threats. Our analysis linked these traces to the largest DNS amplification DDoS [4] for the following reasons: 1) in addition to the use of the ANY DNS query , the traces of this attack targeted the "'ripe.net"' domain name; this domain was used in the largest DDoS as declared in a blog posted by the victim [4]; 2) the timing of the traces from the host with ID 1 started on March 15^{th}, whereas those of the host with ID 2 started on March 17^{th}. The two mentioned dates could be found in the media [24, 25] and were posted on Twitter on March 17^{th} by one of the victim's support personnel [26]. In order to depict this distributed attack, in Figure 2, we highlighted the threat using a colored dashed-line. The first or/and second peaks are likely performed as testing before actually executing the largest DDoS as demonstrated by the third peak. Our result match the ascending order of peaks as discussed by the victims [4]. This case study is probably sent by an attacker using spoofed IP address of the victims or using compromised machines; we unlikely consider these activities as scanning events that are using legitimate addresses (i.e., the intention is not to DDoS themselves but other targeted victims).

In addition to performing several validation of our results through DShield and the media, we execute a renowned Network Intrusion and Detection System (NIDS) (i.e., Snort) on the whole traces to see if we can detect such malicious activities. The NIDS labeled 129 out of the inferred 134 (96%) DNS amplification DDoS as executing filtered portsweep probes. We have found that the 5 undetected attacks refer to the low-speed (stealthy) attacks, which are, by default, undetectable using a typical NIDS. In summary, we can claim that our approach that aims at inferring DNS amplification DDoS yielded zero false negative in comparison with a leading NIDS. Further, our approach, leveraging the darknet space, can infer DNS amplified DDoS activities while a NIDS is limited to pinpointing scanning attempts.

VI. Concluding Remarks

This work presented a new approach to infer Internet DNS Amplification Denial of Service activities by leveraging the darknet space. The approach corroborated the fact that one can infer DDoS attacks without relying on backscattered analysis. The detection module is based on certain parameters to fingerprint network flows as DNS amplification DDoS related. The classification module amalgamates the attacks based on their possessed rate. The analysis was based on 720 GB of real darknet traffic collected during a recent 3 months period. The results disclose 134 DNS amplified DDoS activities, including flash and stealthy attacks. Moreover, the case study provided significant cyber security intelligence related to the largest DNS amplification attack.

Victim	Requested Domain Name	Detection Period	Analyzed AttackDuration (second)	Intensity (packet)	Contacted UniqueDark IPs	Avg. Packet Size (Bytes)	Avg. Rate (pps)	RateCategory
M1	B	March 15	34605	3176785	360683	68.00	91.80	Medium
M2	B	March 17 to 18	93508	14464427	360705	68.00	154.69	Medium

TABLE IV: DNS Amplification DDoS Traces

A. Lessons Learned & Future Work

From this work, we can extract the following insights related to DNS amplification attacks: First, when compared to previous years, we have found that the DNS amplification attacks are behind the increase of DNS queries of type `ANY` on the Internet. Second, we have pinpointed that the majority of the attacks target the root domain. Third, we have encountered that DNS amplified attack rates can range from very low to high speeds. High speeds attacks pinpoint victims of spoofed attacks and compromised machines whereas the very slow attacks reflects stealthy scans. Last but not least, we have unexpectedly uncover a UDP-based mechanism used by DNS amplification attackers to execute DNS amplification attacks in a highly rapid manner without collecting information about open DNS resolvers. Further, more importantly, we have inferred that unlike typical DDoS attempts that scan for vulnerable machines and then execute the attack, the largest DNS amplification analyzed was executed in only one step; DNS queries are sent to the Internet with the intention to reach open DNS resolvers, which subsequently trigger an amplified reply to the victim. As for future work, we aim to execute our model on a larger data set and implement our proposed approach in a near real-time fashion.

References

[1] Arbor Networks. 2012 Infrastructure Security Report. [Online]. Available: http://tinyurl.com/ag6tht4

[2] US-CERT. DNS Amplification Attacks - Alert (TA13-088A). [Online]. Available: http://tinyurl.com/bp4xud4

[3] Ars Technica. When spammers go to war: Behind the Spamhaus DDoS. [Online]. Available: http://tinyurl.com/d9vkegg

[4] CloudFlare blog. The DDoS That Knocked Spamhaus Offline. [Online]. Available: http://tinyurl.com/d46gpkj

[5] D. Moore *et al.*, "Inferring internet denial-of-service activity," in *In Proceedings of the 10th Usenix Security Symposium*, 2001.

[6] Fachkha *et al.*, "Towards a forecasting model for distributed denial of service activities," in *the 12th IEEE International Symposium on Network Computing and Applications (NCA), 2013*, 2013, pp. 110–117.

[7] Bou-Harb *et al.*, "A statistical approach for fingerprinting probing activities," in *the Eighth International Conference on Availability, Reliability and Security (ARES), 2013*. IEEE, 2013, pp. 21–30.

[8] Dainotti *et al.*, "Analysis of a/0 stealth scan from a botnet," in *Proceedings of the 2012 ACM conference on Internet measurement conference*. ACM, 2012, pp. 1–14.

[9] D. Moore *et al.*, "Inferring internet denial-of-service activity," *ACM Transactions on Computer Systems (TOCS)*, 2006.

[10] CAIDA. Cooperative Association for Internet Data Analysis. [Online]. Available: http://www.caida.org

[11] Fadlullah *et al.*, "Dtrab: combating against attacks on encrypted protocols through traffic-feature analysis," *IEEE/ACM Transactions on Networking (TON)*, vol. 18, no. 4, pp. 1234–1247, 2010.

[12] Yao *et al.*, "Passive ip traceback: capturing the origin of anonymous traffic through network telescopes," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, Aug. 2010.

[13] J. Bi, P. Hu, and P. Li, "Study on classification and characteristics of source address spoofing attacks in the internet," in *Proceedings of the 2010 Ninth International Conference on Networks*. Washington, DC, USA: IEEE Computer Society, 2010.

[14] B. Irwin and N. Pilkington, "High level internet scale traffic visualization using hilbert curve mapping," in *VizSEC 2007*. Springer, 2008.

[15] J. Oberheide, M. Karir, and Z. M. Mao, "Characterizing dark dns behavior," in *In Fourth GI International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA)*, 2007.

[16] V. Paxson, "An analysis of using reflectors for distributed denial-of-service attacks," *SIGCOMM Comput. Commun. Rev.*, 2001.

[17] Dagon *et al.*, "Corrupted dns resolution paths: The rise of a malicious resolution authority." in *NDSS*, 2008.

[18] Bailey *et al.*, "The internet motion sensor," in *Proceedings of the 12th ISOC Symposium on Network and Distributed Systems Security (SNDSS)*, 2005, pp. 167–179.

[19] Fachkha *et al.*, "Investigating the dark cyberspace: Profiling, threat-based analysis and correlation," in *the 7th International Conference on Risk and Security of Internet and Systems (CRiSIS)*, 2012.

[20] Bou-Harb *et al.*, "Cyber scanning: A comprehensive survey," *Communications Surveys and Tutorials*, 2013.

[21] Wustrow *et al.*, "Internet background radiation revisited," in *Proceedings of the 10th annual conference on Internet measurement*. ACM, 2010, pp. 62–74.

[22] Staniford *et al.*, "The top speed of flash worms," in *Proceedings of the 2004 ACM workshop on Rapid malcode*, ser. WORM '04. New York, NY, USA: ACM, 2004, pp. 33–42.

[23] S. Staniford, J. A. Hoagland, and J. M. McAlerney, "Practical automated detection of stealthy portscans," *Journal of Computer Security*, vol. 10, no. 1, pp. 105–136, 2002.

[24] Bloomberg News. Dutch Man Arrested in Spain in Probe of Spamhaus Attack. [Online]. Available: http://tinyurl.com/lzqheb5

[25] The Irish Times. Spam dispute results in biggest ever cyber attack. [Online]. Available: http://tinyurl.com/d2xwv9w

[26] Luc Rossini. Twitter: Spamhaus is currently under a ddos attack. [Online]. Available: http://tinyurl.com/m7dayxr