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*—Recently, there has been a noteworthy shift towards a new phenomenon of Distributed Denial of Service (DDoS) attacks dubbed as DNS Amplification DDoS attempts. Although the latter technique has been identified for several years, it has been seldom used until the debilitating DDoS attack in March, 2013 that peaked at 300 Gbps targeting an anti-spam organization. This work proposes a novel approach to infer and characterize Internet-scale DNS amplification DDoS attacks by leveraging the darknet space. Contrary to the pioneer work on inferring DDoS using darknet, this work proves 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 and the expectation maximization and k-means clustering techniques in an attempt to identify botnets of DNS Amplification DDoS. 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 a 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 that was obscured in recent years. Further, the results uncover high-speed and stealthy attempts that were never previously documented. The extracted insights from various validated DNS amplification DDoS case studies 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.**

Keywords*—*DDoS, DoS, DNS Amplification, DNS Attacks

I. Introduction

Cyber attacks continue to threaten today's information technology. These threats are growing dramatically in terms of size and impact targeting large organizations, Internet service providers, as well as governments. 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. Further, it was stated that the top 4 perceived threats for the next 12 months will be DDoS related, targeting customers, network and service infrastructure. Governmental organizations, corporations as well as critical infrastructure were also recently deemed as DDoS victims [2].

A DNS amplification attack is a form of DDoS that relies on the use of publically accessible open recursive DNS servers to overwhelm a victim system with DNS response traffic [3]. 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 [4]. 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.

When a large scale DNS amplified DDoS attack hits the Internet, it becomes essential for IT security operators and experts to answer the following questions:

- How to infer large-scale DNS amplification DDoS activities?
- What are the characteristics of DNS amplification DDoS attacks?
- What inferences can we extract from analyzing DNS amplification DDoS traces?

Answering those questions would aid organizations, governments and law enforcement agencies to build a central infrastructure to scrutinize DNS amplification traffic in order to contribute in understanding, detecting, preventing, assessing, mitigating and even attributing of DNS amplification DDoS attacks.

In this context, we frame this paper's contributions as follows:

- Proposing a systematic flow-based approach for inferring DNS amplification DDoS activities by leveraging DNS queries to darknets.
- Characterizing the inferred DNS amplification DDoS threats during a recent 3 months period.
- Applying clustering and similarity testing algorithms in an attempt to identify botnets of DNS amplification DDoS attacks.

The remainder of this paper is organized as follows: In Section II, we present the proposed approach and elaborate on various aspects of its components. In Section III, we empirically evaluate the approach and disclose several DNS amplification DDoS case studies. Finally, Section IV summarizes the paper and discusses the future work.

II. 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 3 components, namely, the detection, the rate classification and the clustering 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, as discussed in the previous section, 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.

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

TABLE I: DNS amplified DDoS Identification Parameters

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 [5] that targeted thousands of distinct unused IPs within seconds and other stealthy scanning activities [6] that persisted for several weeks. These 2 types of attacks are discussed in Section III.

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.

B. Rate Classification Component

The rate of the attack is one of the major characteristics of DDoS activities [7]. 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.

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

TABLE II: Attack Classification per Rate

C. Clustering Component

In an attempt to uncover and cluster similar DNS amplification DDoS traces that might be executed by similar authors/code/botnet, we resort to data mining clustering approaches. To achieve this task, we have extracted a number of attributes as shown and described in Table III.

Attribute	Description	Options
ip.flag	IP Flags	0x00 or 0x02
ip.flag.df	Don't fragment	0 or 1
ip.len	Total IP Length	56 or 64 or ip-others
ip.ttl	Time to live	< 60 or > 100
udp.len	UDP Length	36 or 44 or udp-others
dns.count.add.rr	DNS Additional RRs	0 or 1
dns.qry.name	DNS Query Name	11 values
flow.avg.pkt.size	Average Packet Size	70 or 78 or flow-others
flow.attack.duration	Attack Duration	$<$day or btw-day-week or $>$week
high.asn.numb	Autonomous System #	42 values

TABLE III: Chosen Clustering Attributes

Rank	Attribute
1	high.asn.numb
2	dns.qry.name
3	ip.len
4	udp.len
5	flow.avg.pkt.size

TABLE IV: Top 5 Attributes

Note that, we have initially analyzed more than 260 attributes. However, we have leveraged a ranker [8] to evaluate the information gain of all the attributes and have chosen the top 10. This allowed us to filter out those attributes that were not applicable or has no or low information gain such as missing values, noisy or inconsistent data. Further, the top 5 ranked attributes based on the information gain metric are shown in Table IV.

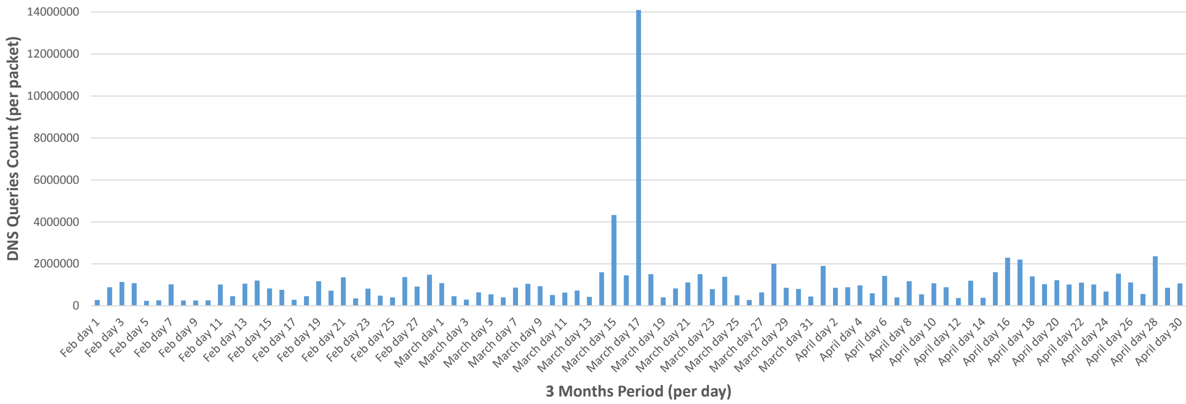

Fig. 1: DNS Queries Distribution of February, March and April 2013

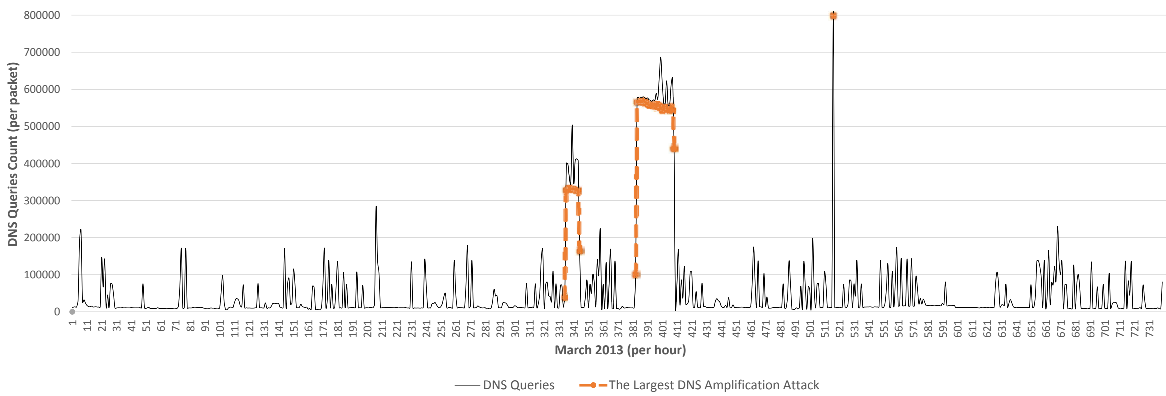

Fig. 2: DNS Queries Distribution of March 2013

We have also employed the generalization technique [9] to improve the results of our clustering approach. Generalization is performed on some attributes in our dataset such as IP length, IP time to live, UDP length and average packet size. For example, in regards to the IP length attribute, we have noticed that the majority of attacks have either 56 or 64 bytes whereas the rest, which are the minority, possess distinct values. Hence, we grouped all the rest under one category called 'ip-others'. We employ the same technique on the above mentioned attributes that share a similar case.

In order to perform the clustering, we have leveraged two algorithms, namely, the k-means and the Expectation Maximization (EM) algorithms. Subsequently, we briefly describe the latter two algorithms. For more information regarding the inner workings of the aforementioned clustering algorithms, we kindly refer the reader to [9].

III. Empirical Evaluation

The evaluation is based on a real darknet dataset during 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 are distributed in many countries and 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, wireshark and tcpstat. In regards to our data mining exercises, our analysis is based on Weka [10], which is a data mining software implemented in Java.

We abide and closely follow the steps of our proposed approach that was discussed in Section II to elaborate on our analysis, which is based on two main elements, namely, the characterization and the insights generation of the DNS amplification DDoS traces. In total, our approach identified a total of 134 DNS amplification DDoS attacks including high-speed, medium and stealthy attacks. Note that the IPs and domains (except root) of the analyzed data are anonymized for privacy and sensitivity issues.

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 [11]. 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 III-D.

1) Query Type Distribution: In order to understand the types of DNS queries received on our dark space, we list in Table V the DNS query type distribution of the analyzed dataset. As expected, the vast majority of these are `ANY` queries. Moreover, it is interesting to find that the top 4 records are the same for the entire 3 months period. Further, in contrast with the results in 2007 by [12], 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 [11].

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 V: Top 5 DNS Query Type Distribution of 3 Months Period

2) Requested Domains: In this section, we illustrate the top requested DNS domains as shown in Figure 3. We anonymize TLDs for sensitivity issues.

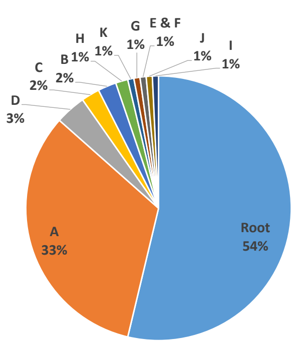

Fig. 3: Top Requested Domains

Figure 3 shows that Root is the most requested domain name as perceived by the monitored darknet. Recall that attackers will typically submit a request for as much zone information as possible to maximize the amplification effect. Hence, the use of Root as the requested domain name. Note that the second top requested domain (labeled as A) is a TLD that belongs to one of the largest Internet-scale DNS operators.

B. Clustering Insights

This section highlights our clustering results. Recall that the aim is to cluster similar DNS amplification DDoS traces that might be executed by similar authors/code/botnet.

Since we had no prior knowledge on the number of clusters, we first run the EM algorithm to infer the number of clusters by cross validation [13]. We executed the algorithm in several cluster modes, using a training set and several percentage split tasks. We compared all the results and chose the model with the highest log likelihood for the best fit. After retrieving the number of clusters, we run the k-means with that number of clusters for further analysis. Again, we run several experiments using the k-means algorithms and chose the model with 60% training data and 40% for testing as it achieved the minimum cluster sum of squared errors. Based on our testing data, Table VI lists our summarized instances per clusters while Figures 4 and 5 visualize the results.

Cluster	**EM Instances**	**k-means Instances**
0	76 (57%)	31 (57%)
1	9 (7%)	4 (7%)
2	32 (24%)	12 (22%)
3	7 (5%)	5 (9%)
4	10 (7%)	2 (4%)

TABLE VI: EM and k-means Clustered Instances

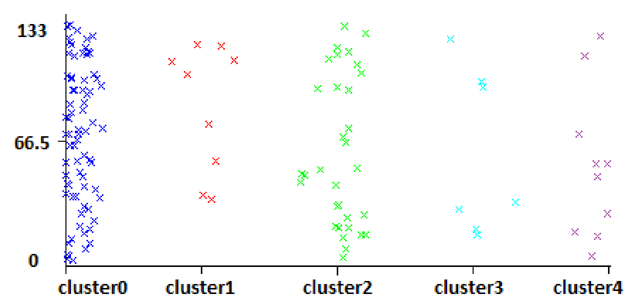

Fig. 4: EM Clustering of DNS Amplified DDoS attacks

It is evident that both algorithms have clustered 56% of the instances in cluster 0 and 7% of them in cluster 1. This relatively validates the quality of our chosen attributes (Recall Section II-C) and the accuracy of our clustering approach. Next, we disclose the attributes that formed the clusters. Table VII shows the cluster centroids of the k-means algorithm.

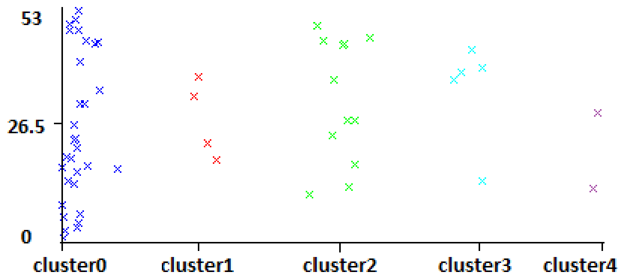

Fig. 5: k-means Clustering of DNS Amplified DDoS attacks

This table is based on the training set of the data.

Attribute	Cluster 0 (49)	Cluster 1 (8)	Cluster 2 (14)	Cluster 3 (5)	Cluster 4 (4)
high.asn.numb	ASN-V	ASN-W	ASN-X	ASN-Y	ASN-Y
ip.flag	0x02	0x00	0x02	0x00	0x02
ip.flags.df	0	1	0	1	0
ip.len	56	others-ip	64	64	64
ip.ttl	<60	<60	<60	>100	<60
udp.length	36	others-udp	44	44	44
dns.qry.name	Root	B	A	A	A
flow.avg.pkt.size	70	others-flow	78	78	78
flow.attack.duration	<1day	<1day	<1day	<1day	btw-day-1week

TABLE VII: k-means Training Cluster Centroids

It is shown that our model clustered the traces based on 4 different ASNs with some specific attributes. For instance, in regards to cluster 0, all the DNS amplification DDoS attacks have source IPs within ASN-X, have the DF flag set in the IP header, and its packets are available for fragmentation. Moreover, the same flow must have an IP length of 56 bytes and a TTL value less than 60. In addition, the UDP length must be 36 bytes while the requested domain is root. Additionally, all the attacks that belong to cluster 0 should be launched within a 1 day period and possess DNS queries of an average size of 70 bytes. Through manual inspection, we found that the majority of IPs that fall within cluster 0 are originating from Netherlands. Similar concept applies for other clusters. Note the similarities between cluster 2, 3 and 4 which could be the result of one botnet using different ASNs from different locations.

After the clustering exercise, in order to evaluate our model, we run the cluster evaluation algorithm in weka [10]. First it ignores the class attribute and generates the clustering. Then it assigns classes to the clusters during the testing mode, based on the majority value of the class attribute within each cluster. Then it calculates the classification error. Based on this technique, we have achieved a 71 % accuracy. In other word, our model incorrectly classified 29% of the traces to their corresponding cluster. We aim, in our future work, to analyze more data and run other clustering algorithms to improve this result.

Please note, that although we do not have a decisive proof of whether each cluster represent a campaign or a botnet of DNS amplification DDoS, we succeeded in this task by pinpointing similarities among the DNS amplification DDoS traces.

C. Similarity Insights

This exercise aims at inferring insights related to the used darknet address space. The rationale behind this task states that since bots in the same botnet typically utilize the same list of IPs when launching their attacks, it would be interesting to capture the similarity of use related to these IP lists. By accomplishing this, we can possibly infer botnet campaigns or at least detect similarities in attack mechanisms. To achieve the intended goal, we executed an experiment to represent attacks that exchange at least 90% of dark IPs. Figure 6 depicts an IP map[1] that satisfies the latter condition.

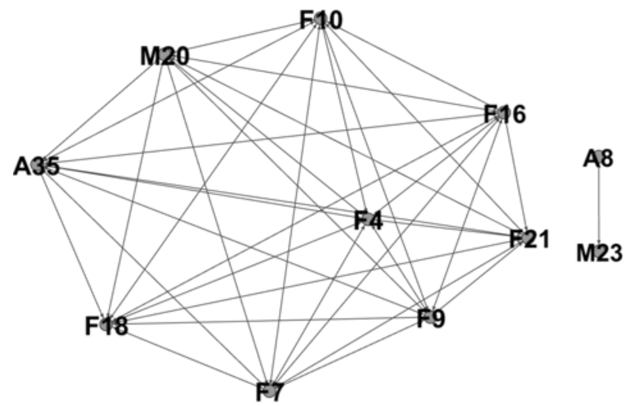

Fig. 6: IPs Sharing at least 90% Darknet Space

It is disclosed that two groups of IPs share at least 90% of dark IPs. The smaller group consists of 2 IPs from different months (March and April). Our analysis identified that these two sources share not not just dark IP usage, but also country, ASN number, speed range, requested domain, and many other attributes as previously identified in Section III-B in cluster 0. As for the second group, 7 out of 8 originate from the same ASN number. All of the attacks in this group are initiated from Europe, specifically from Netherlands. Similar to the first group, these attacks share similarities in clustering attributes and 55.56% of these traces are found also in cluster 0. One of the interesting point uncovered by analyzing this group is that all its members are sharing a specific address space range, possibly highlighting a DNS amplification DDoS botnet campaign.

D. Case Studies

We discuss below some major case studies that belong to three different attack rates.

The first case study represents high-speed (flash) DNS amplification DDoS detected attacks. In our dataset, we have

[1]The map was automatically generated using Gephi [14], an open source visualization tool

Victim/ Scanner ID	Requested Domain Name	Detection Period	Analyzed Attack Duration (second)	Intensity (packet)	Contacted Unique Dark IPs	Avg. Packet Size (Bytes)	Avg. Rate (pps)	Rate Category
F1	A	Feb 19	0	34410	34410	78	79565.67	High
M1	A	March 18	1	50257	50257	78.00	46677.36	High
A1	A	Apr 15	3	61859	61859	78	21672.18	High
M5	B	March 17 to 18	93508	14464427	360705	68.00	154.69	Medium
M10	B	March 15	34605	3176785	360683	68.00	91.80	Medium
M51	D	March 27 to 28	41548	44	44	70.00	0.00	Low
M52	D	March 27 to 28	75803	42	42	70.00	0.00	Low
M53	D	March 27 to 28	90128	39	39	70.00	0.00	Low
M54	D	March 27 to 28	56874	37	37	70.00	0.00	Low

TABLE VIII: DNS Amplification DDoS Traces

found 3 attacks that fall within this category; ID F1, M1 and A1. These are shown in the Table VIII. First, attack F1 is the fastest detected DNS amplification DDoS attack. It was lunched on February 19$^{\text{th}}$. The detected attack has a rate of 79565.67 pps. This propagation speed is 17 times faster than the Slammer worm [5]. This attack scanned 6.5% of our darknet space in less than 1 second. Assuming the intent of the attacker is to send one packet for each IP, a malware with this speed can scan the whole IPv4 Internet address space in less than a week (6 days and few hours). In order to validate the occurrence of this flash DNS amplification DDoS attack, we resorted to publicly accessible Dshield [15] data and inspected port 53 for the 3 days before and after the 19$^{\text{th}}$ of February. We have noticed a significant increase at this specific date. According to Dshield data, the average incident reports measured 14.28% for the whole 7 days. However, on February 19$^{\text{th}}$, the average reached 38.19% with a 10347879 increase in reports from the previous day. Second, attack M1 was launched from Taiwan on March 18$^{\text{th}}$. This date is the same date of the largest DDoS attack as declared in [11]. This flash scan sent probes to 50257 unique dark IP (9.5 % of the our /13 darkspace) within 1 second with an average rate of 46677.36 pps. This speed is almost 10 times faster than the Slammer worm. With this speed, this DNS amplification DDoS can scan 16 millions IPv4 hosts (/8) on the Internet in less than 6 minutes. Assuming that the attacker found the same amount (i.e., 50257) of open DNS resolvers[2] and subsequently he sent just 5 requests towards them, and supposing a packet size of 78 bytes, maximum amplification factor of 100 times, and a 150 Mbps reply, then it can be computed that a botnet of 513 bots will indeed generate a 75 Gbps attack. The latter number refers to the peak speed of the largest DNS amplification DDoS attack as declared in [11]. It is of momentous importance to note that in order to generate a 300 Gbps DNS amplification DDoS, with the same assumptions, a botnet of 2055 hosts is required. Further, if each bot sends 25 requests instead of 5, then the required botnet infrastructure to execute a 75 Gbps and 300 Gpbs DNS amplified DDoS attack will decrease to 81 and 323 hosts, respectively. Third, attack A1 was also launched from the same city as of F1 on April 15$^{\text{th}}$. The attack possesses a rate of 21672.18 pps. This attack scanned 11.7% of our darknet address space.

[2]This is very probable as there is around 33 million open DNS servers on the entire Internet [16]

The second case study, which involve medium speed attacks, is one of the major inferred DNS amplification DDoS in terms of size and impact. This attack targeted only one organization using 2 hosts (ID M5 and M10 of Table VIII). 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 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 [11]; 2) the timing of the traces from the host with ID 10 started on March 15$^{\text{th}}$, whereas those of the host with ID 5 started on March 17$^{\text{th}}$. The two mentioned dates could be found in the media [17, 18] and were posted on Twitter on March 17$^{\text{th}}$ by a company support personnel [19]. 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 [11].

The above two mentioned case studies are probably sent by an attacker using spoofed IP address of the victims or using compromised machines; we unlikely consider these activities as scanning event that are using legitimate addresses (i.e., the intention is not to DDoS themselves but other targeted victims).

The third case study involves slow rate attacks such as hosts with ID M51 to M54 in Table VIII. This analysis targets stealthy DNS amplified DDoS; these attacks have low sending rate and are typically hard to detect [6]. From this Table, all information regarding these 4 hosts appears very similar or the same. Therefore, they are mostly generated by the same author/botnet or using the same malware. Although we cannot claim the orchestration among these hosts, our data highlights some shared characteristics among such stealthy DNS ampfication DDoS threats. Note that the requested domain names within these attacks is a top-notch organization that deals with securing online transactions.

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 [20]) 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 third case study that was previously discussed. After manual inspection, the attacks turned out to be originating from the same source (i.e., scanner) who is executing stealthy scans but in different time periods. Moreover, all these attacks are requesting one organization's domain. 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 detect DNS Amplified DDoS activities while an NIDS is limited to pinpointing scanning activities.

IV. Conclusion

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 while the clustering component attempts to identify flows that share similarity features in an attempt to disclose campaigns or botnets of DNS Amplification DoS. 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. The results also pinpointed some of the largest DNS amplification DDoS traces. The clustering and similarity exercises provided insights and inferences that permit the detection of DNS amplification DDoS botnet activities. Moreover, the discussed case studies elaborated on three attack categories and provided significant cyber security intelligence related to them. As for future work, we aim to run our model on a larger period of data and execute more complex data mining exercises to improve our clustering model. Moreover, we would like to 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] Forbes. Testing The Limits, LulzSec Takes Down CIA's Website. [Online]. Available: http://tinyurl.com/bfhzbta

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

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

[5] S. Staniford, D. Moore, V. Paxson, and N. Weaver, "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.

[6] 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.

[7] D. Moore, C. Shannon, D. J. Brown, G. M. Voelker, and S. Savage, "Inferring internet denial-of-service activity," *ACM Transactions on Computer Systems (TOCS)*, 2006.

[8] The University of Waikato. WEKA Select Attributes: Ranker. [Online]. Available: http://tinyurl.com/kt6rccu

[9] Han, Jiawei and Kamber, Micheline and Pei, Jian, *Data mining: concepts and techniques*, 2006.

[10] The University of Waikato. Weka: Data Mining Software in Java. [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka/

[11] CloudFlare blog. The DDoS That Knocked Spamhaus Offline (And How We Mitigated It). [Online]. Available: http://tinyurl.com/d46gpkj

[12] 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.

[13] P. Smyth, "Model selection for probabilistic clustering using cross-validated likelihood," *Statistics and Computing*, vol. 10, no. 1, pp. 63–72, 2000.

[14] Gephi. Gephi: An open source graph visualization and manipulation software. [Online]. Available: https://gephi.org/

[15] DShield. Dshield port report. [Online]. Available: http://www1.dshield.org/

[16] Open Resolver Project. [Online]. Available: http://openresolverproject.org/

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

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

[19] Luc Rossini. Twitter tweet: Spamhaus is currently under a ddos attack against our website which we are working on mitigating. our dnsbls are not affected. [Online]. Available: http://tinyurl.com/m7dayxr

[20] Sourcefire. A free lightweight network intrusion detection system. [Online]. Available: http://www.snort.org/