Who Gets the Boot? Analyzing Victimization by DDoS-as-a-Service

Arman Noroozian[1(✉)], Maciej Korczyński[1], Carlos Hernandez Gañan[1], Daisuke Makita[2,3], Katsunari Yoshioka[2], and Michel van Eeten[1]

[1] Delft University of Technology, Delft, Netherlands
a.noroozian@tudelft.nl
[2] Yokohama National University, Yokohama, Japan
[3] National Institute of Information and Communications Technology, Koganei, Japan

Abstract. A lot of research has been devoted to understanding the technical properties of amplification DDoS attacks and the emergence of the DDoS-as-a-service economy, especially the so-called *booters*. Much less is known about the consequences for victimization patterns. We profile victims via data from amplification DDoS honeypots. We develop victimization rates and present explanatory models capturing key determinants of these rates. Our analysis demonstrates that the bulk of the attacks are directed at users in access networks, not at hosting, and even less at enterprise networks. We find that victimization in broadband ISPs is highly proportional to the number of ISP subscribers and that certain countries have significantly higher or lower victim rates which are only partially explained by institutional factors such as ICT development. We also find that victimization rate in hosting networks is proportional to the number of hosted domains and number of routed IP addresses and that content popularity has a minor impact on victimization rates. Finally, we reflect on the implications of these findings for the wider trend of commoditization in cybercrime.

1 Introduction

While Distributed Denial-of-Service (DDoS) attacks have been around for a long time, the use of amplification techniques has transformed the criminal ecosystem. These techniques now make up the bulk of the observed attack traffic [1,2]. This shift is intimately related to another trend: the rise of DDoS-as-a-service, also known as *booters*. Booters are a clear example of the so-called commoditization of cybercrime [3]: criminal service providers bundling all the resources and tools needed for an attack and offering them in an accessible way as a commodity service to anyone willing to pay.

Several in-depth studies have illuminated the supply side of the market for DDoS: the technical resources and techniques deployed by the criminal service providers [2,4,5]. We have also learned quite a bit about the economics of booters from publicly-leaked dumps of several operational databases containing information about revenue and customers [6–8].

F. Monrose et al. (Eds.): RAID 2016, LNCS 9854, pp. 368–389, 2016.
DOI: 10.1007/978-3-319-45719-2_17

What is much less understood, however, is how the abundance and affordability of DDoS-as-a-service has impacted victimization patterns. Who is bearing the brunt of the lowered barriers for DDoS attacks? Existing studies have revealed some basic distributions of victims across countries, Regional Internet Registries (RIRs) and Autonomous Systems (ASes). They have pointed to end hosts, gaming servers and hosting providers [1], but they lack a more in-depth investigation and explanation of victimization patterns.

This paper addresses this knowledge gap and profiles the affected networks and victims. It uses a dataset of 1,115,795 victim IP addresses captured over the past two years (2014–2015) via several amplifier-honeypots [2]. From the IP addresses, we infer certain properties of the victims and identify the factors determining their distributions across networks and countries.

Since the existing work on amplifiers and booters has not focused on the victims, the public understanding of them has been shaped by anecdotal news articles and by industry reports compiled by DDoS mitigation providers. The former focus on the more news-worthy cases, i.e., the attacks against high profile targets. The latter are biased towards their own customer base, i.e., enterprises purchasing DDoS protection services, as that is where the data is being collected. As we demonstrate in this paper, neither provide a good understanding of the ecosystem of commoditized DDoS attacks.

We summarize the main contributions of this paper as follows:

- We show that the bulk of the victims (62 %) are users in access networks, rather than in hosting networks (26 %). Only a small fraction resides in enterprise networks;
- We demonstrate that the victimization rate in access networks is highly proportional to the number of broadband subscribers in those networks, suggesting that the commoditization of attacks has led to a democratization of victims;
- We find that certain countries have a significantly higher number of victims per subscriber. This rate is weakly related to institutional factors such as information and communication technologies (ICT) development, suggesting geographical network effects among attackers and victims increasing the uptake of DDoS-as-a-service;
- We demonstrate that victimization in hosting networks is proportional to the number of IP addresses and hosted domains, and also influenced by the popularity of the hosted content.
- Where we were able to specifically identify webhosting victims, we find that they have barely any enterprises among them or other valuable targets. The largest victim group are gaming-related sites, most notably around Minecraft, suggesting that the commoditization of DDoS facilitates crime that is mostly not profit driven.

In what follows we first present some background (Sect. 2) and the data collection method (Sect. 3), we then discuss the distribution of victim IP addresses over access, hosting and other networks (Sect. 4). Next we delve deeper into victimization patterns in access networks (Sect. 5) and hosting networks (Sect. 6).

We briefly explore whether attack duration is different across victim populations (Sect. 7). After comparing our findings to related work, we summarize our conclusions on the consequences of DDoS-as-a-service and discuss the implications for the wider issue of the commoditization of cybercrime.

2 Background

DDoS attacks have been associated with a range of motives. They can be profit-driven – as in the case of extortion, disrupting competitors, or using it as a smoke screen for committing financial fraud – or motivated by other objectives, such as political protest, harassment, or gaining advantage in online gaming [1,3].

Amplification DDoS attacks now make up a considerable fraction of network-layer DDoS incidents [9–11]. Attackers send requests to amplifiers – a.k.a. reflectors – and spoof the source IP address, so that the amplifiers responses are directed to the victim. A whole range of protocols can be abused for amplification and millions of machines run these protocols which enables such attacks [12].

Most of the amplification attacks stem from booter services [2,7]. The price for purchasing an amplified DDoS attack can be as low as $1, as the analysis of some leaked booter databases demonstrates [7,13]. A purchase from a booter would typically entail access to the service for a limited amount of time, tied to different pricing tiers. Most attacks are very short, less than 10 min [7].

On the customer side of booter services, leaked databases have shown that most customers of DDoS-as-a-service use it only once to attack a single target [7] and only a small fraction of them hide their tracks via `TOR` or `VPN`. This might indicate that their technical skills are limited or that they do not perceive a need to hide. The users that do hide their tracks, tend to return for more and also tend to launch more attacks [6]. The databases have also revealed that gamers make up a specific and important customer group [6]. On the victim side, booter databases contain the targeted IP addresses or URLs, but these sets are limited in scope and volume. The top 100 most attacked sites were mostly game servers and game forums [6].

Besides booter databases, `NTP` amplification attacks allow victim IPs to be retrieved from the `NTP` servers. From this data, Czyz et al. [1] point to end hosts and gaming servers to be common victims [1]. Amplification honeypots have also collected victim IP addresses [2]. They have only been superficially analyzed, in terms of the distribution over countries and IP address space. The U.S., China and France were the most attacked countries. In this paper, we significantly extend the analysis of honeypot data.

The only other systematic source of information comes from industry reports by DDoS mitigation providers. Akamai points to gaming, software and the financial industry as the major victims [9], with a small fraction of victims belonging to the telecom industry. Other reports suggest hosting as major victims [14]. These industry reports have specific limitations and biases, which we will return to in Sect. 4.

3 Honeypot Data

The victim data used in this study was gathered via a set of amplifier honeypots – dubbed AmpPots [2] – which have been deployed over the past two years (2014–2015). They run services that are known to be misused for amplification attacks: `QotD` (17/udp), `CharGen` (19/udp), `DNS` (53/udp), `NTP` (123/udp), `SNMP` (161/udp) and `SSDP` (1900/udp). Each AmpPot uses real server software (in 'proxy' mode) to provide the aforementioned services except for `SSDP` in which an emulated script is used instead. The responses of AmpPots are filtered in order to prevent from contributing to actual attacks. More details of AmpPot can be found in the previous study [2].

Table 1. Overview of deployed AmpPots.

AmpPot ID	Deployed on	IP Changes	Notes
H01	`2012-10-07`	19	added `QOTD`, `NTP`, `SNMP`, `SSDP` on `2014-09-25`. Discontinued on `2015-10-09`
H02	`2013-05-13`	25	only `DNS` supported
H03	`2014-05-13`	9	added `SNMP` support on `2014-09-17` and `SSDP` on `2014-10-03` *
H04	`2014-05-13`	10	added `SNMP`, `SSDP` support on `2014-09-17` *
H05	`2014-05-10`	4	added `SNMP`, `SSDP` support on `2014-10-18` *
H06	`2014-05-10`	6	added `SNMP`, `SSDP` support on `2014-10-18` *
H07	`2014-05-10`	8	added `SNMP`, `SSDP` support on `2014-10-18` *
H08	`2015-11-09`	0	–**

Note:* Deployed with `QOTD`, `CharGen`, `DNS` and `NTP` support
Note:** Deployed with support for all protocols

In total 8 AmpPots were deployed on the Internet during the measurement period of 2014–2015. Table 1 shows a summary of the operational timeline and supported protocols of these devices. At the start of the measurement period (2014-01-01), two AmpPots were operational and initially only supported the `CharGen` and `DNS` protocols. With a sustained effort to monitor more amplification attacks, more devices were gradually added with support for additional abused protocols. At the end of the measurement period (2015-12-31) the deployed AmpPots collectively monitored 6 services except for `H02` which only supports `DNS`. All AmpPots are located at ISPs in Japan and their IP addresses are dynamically assigned. Depending on the ISP, the IP addresses changed every 5–30 weeks, on average.

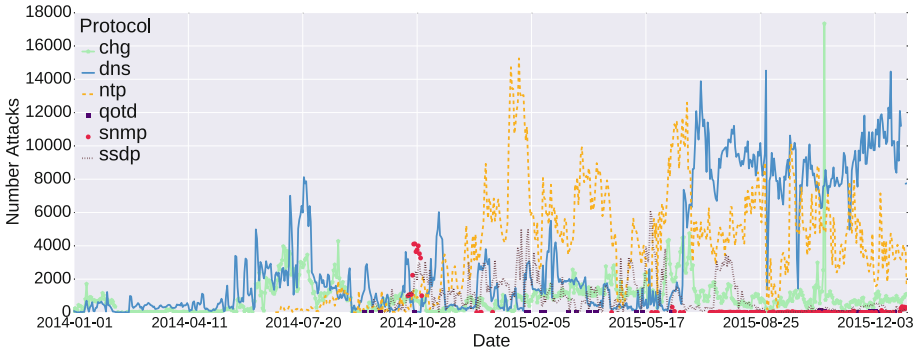


Fig. 1. Number of amplification attacks per protocol

AmpPots observe not only amplification attacks, but also scans from researchers or attackers who search for vulnerable devices. To separate actual attacks from scans, attacks are defined as a series of at least 100 consecutive query packets that a single host sent to an AmpPot, where consecutive means that there was no gap of more than 600 s between two packets. This definition is in concord with the one used in [2]. We did, however, reduce the gap from 3600 s to 600 s, in order to analyze attack duration with a more fine-grained approach.

Collectively, the AmpPots have monitored 1,115,795 unique victim IP addresses from 92 countries and 15,044 unique victim ASes. Figure 1 shows the number of attacks per protocol during 2014 and 2015. As the figure demonstrates, the total number of attacks has increased over time and protocols like `DNS`, `NTP` and `SSDP` have been used more often to launch amplification attacks. During the measurement period, the AmpPots have monitored 5,726,150 amplification DDoS attacks in total: `DNS` (41.26 %), `NTP` (38.73 %), `CharGen` (11.32 %), `SSDP` (8.01 %), `SNMP` (0.65 %), and `QotD` (0.01 %).

4 Victims of Amplification Attacks

Given our amplification attack data the first question we pursue is: *In which type of networks are victims concentrated?*

To avoid confusion, we first define the main concepts. The term *attack* has been defined and operationalized in the previous section. We use the term *target* to refer to the entity (or entities) that the attacker intended to affect. This may be a person, organization, service or machine. Since the data consists of IP addresses, the attacker's intention is not directly observable. For this reason, we use the term *victim* to refer to the targeted IP addresses and the hosts residing there. As DDoS attacks are also a cost to the networks in which the victims reside, we refer to the Autonomous System (AS) that routes the traffic for the victims as *victim AS* or *victim network*. To answer our question we looked up the ASes of the victims and categorized them into three groups: *broadband ISP*s, *hosting* providers, and *other* networks.

To reliably identify the broadband ISPs, we utilize a previously developed mapping that identifies the ASes of broadband ISPs in 82 countries and that has been used to study botnet mitigation in broadband ISPs [15]. The mapping accurately distinguishes between and provides labels for ASNs which have been manually mapped to broadband ISPs, hosting, governmental, mobile ISP, educational and other types of networks. In total, the mapping contains 2,050 labeled Autonomous Systems. The mapping is organized around ground truth data in the form of a highly accurate commercial database; *TeleGeography Globalcomms* [16], containing the broadband subscriber numbers of 211 countries. Compared to machine learning approaches that map AS types [17], our mapping is more accurate since it manually identifies access networks, and the completeness of the mapping is verified with the Telegeography database.

To identify *hosting* providers, we take all the non-broadband ASes in our data and apply a simple heuristic to them. First, we count the number of unique second-level domains (2LDs) hosted within the ASes. For this we used all observed domains in 2014 and 2015 in DNSDB, a large passive DNS (pDNS) database generously provided to us by Farsight Security [18]. DNSDB is sourced from more than 100 sensors located around the world, in addition to authoritative DNS data from various top-level domain (TLD) zone operators. To illustrate: in 2015 DNSDB observed $287M$ unique 2LDs, which map to 69M distinct IP addresses.

We use the accurate AS labels mentioned above to identify a threshold for the number of hosted domains per AS that most accurately separates the ASes labeled as hosting from other types of ASes which may also host domains. Our approach does mean that CDNs and others networks like Cloudflare also get categorized as hosting. Based on the ROC curve constructed we identify this threshold to be 2700 2LDS. Therefore we define as hosting any AS that has not been previously identified as a broadband ISP and that hosts more than $2,700$ 2LDs. This corresponds to a false-positive/true-positive rate of 0.17/0.74. This accuracy is far from perfect, but better than available alternatives. We compared it to machine learning approaches, such as CAIDA's classification of ASes [17]. Using CAIDA's `Content` label as an alternative means for classifying the hosting providers results in a 0.04/0.32 false-positive/true-positive rate of classification. This classification has a better false-positive rate, but this comes at the cost of a highly reduced true-positive rate in comparison to our classification. Alternative methods for identifying hosting providers have also been explored in [19]. They are not directly comparable due to their organizational level classification rather than AS level.

Finally, all ASes that have not been classified as broadband ISP or hosting are labeled as *other*. Our labels and manual inspection show that this group contains governmental and educational networks, mobile and cloud providers, enterprises and more.

Having constructed our network classification, we can now examine the distribution of victims over these networks. Figure 2 plots the results.

It clearly shows that the majority of attacks and victim IPs are located in broadband ISPs, even though they only constitute a small fraction of all ASes

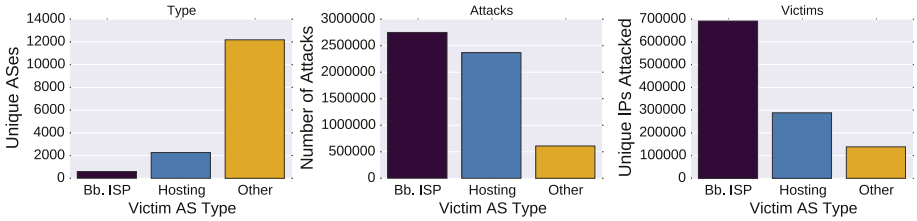


Fig. 2. Distribution of types, attacks and victim IPs

that have been attacked. More precisely, 48 % of the attacks and 62 % of the victims are in access networks. In total, we observe victim IPs from 92 countries in the attack data. We have detailed information on ISPs from 77 of these 92 countries. All identified ISPs in these 77 countries receive attacks, except for 5 countries (GB, US, JO, KE, LV) where at most 2 smaller ISPs are missing from the attack data. This suggests that the whole global broadband market is victimized by these attacks.

The second largest category is hosting: 41 % of attacks and 26 % of victims. The remaining victim networks constitute only a small fraction of the attacks and victims (11 % and 12 %, respectively).

This distribution of victims across broadband and hosting networks is different from earlier work by Czyz et al. [1]. They observed that the top 10 most targeted networks consisted of eight hosting providers and two telecom companies and that access nodes made up around half of all victims. They did observe already a trend that the portion of victims in access networks was increasing, which seems to have continued after their measurement period. Our analysis of the `UDP` ports used for the attacks largely agrees with that of [1]. The most frequently attacked UDP ports by a large margin include ports 80 and 8080, that are more likely to be open and accessible through firewalls. Other application specific ports are also targeted such as (7000) for BitTorrent trackers and CORBA management agent (1050).

We have triangulated our results with CAIDA's mapping of ASes [17], which classifies them as `Content`, `Enterprise` or `Transit/Access`. While these categories are different from ours, which means we cannot directly compare the exact distributions, the CAIDA mapping also locates most victims in `Transit/Access` networks, followed by `Content` and `Enterprise`. This is consistent with our findings.

Networks are not homogeneous, of course. Broadband networks, for example, can also contain hosting services. To probe deeper into the AS-level pattern, we take a closer look at the IP addresses of victims in access and hosting networks. We checked whether the addresses were associated with any domains in our pDNS data. Domains are used for a variety of hosting services; websites, but also for gaming servers, email servers, basically for any service where a human readable name is more convenient than an IP address. The pDNS data found that 95 % of the victims in broadband networks have never been associated with any domains in 2014 and 2015. This suggests that the bulk of the victims in these

networks are access nodes. The remaining 5 % host on average 20.8 domains per IP address (The median domain count is 1 and 75 % of these victims host 3 or less domains).

Since this categorization is dependent on the coverage of our pDNS data, we have cross-checked our domain data with the *Bing.com* search engine. We took a random sample of 1000 broadband victim IP addresses and queried Bing ('IP:<x.x.x.x>') to see if any domains were associated with it. For 9 % of the cases, BING reports observing domains where our pDNS data did not observe any. The opposite was true in 2 % of the cases. This suggests that the pDNS data gives a reasonably accurate picture.

In hosting networks, we found that 46.6 % of the victim IPs have been associated with domains. This confirms earlier work that webhosting is just one among many targets. Figure 3 summarizes the breakdown of the victim types and the subsets which we analyze in more detail in subsequent sections.

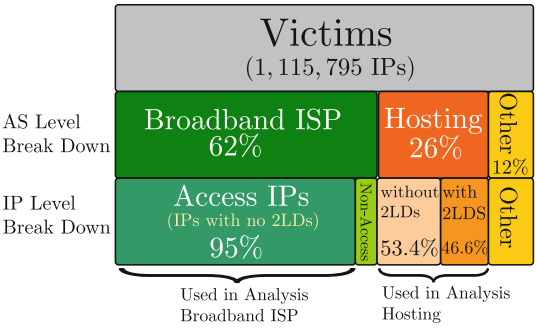


Fig. 3. Breakdown victims

Our results substantially differ from the victimization analysis in the reports of DDoS mitigation providers. There are two types of industry reports: based on traffic data or based on customer surveys. An example of the former is Akamai's State of the Internet report [20]. It identifies the gaming industry as the largest victim of DDoS attacks with 54 % of the attacks, followed by the software and technology industry (23 %) and financial industry (7 %). Only 4 % of attacks map to the Internet and Telecom industry. Another type of industry report is based on surveys among customers of DDoS mitigation providers. A recent example is Arbor Networks' WISR [10], which surveys 287 different organizations of which 24 % are ISPs and 5 % hosting providers. Other industry reports [14] point to hosting as the main victim however, this could be due to a focus on botnet-assisted DDoS attacks.

The mismatch between these reports and our findings is evident. We would argue that when it comes to observing victimization, the industry analyses are more biased than the honeypot data. Industry data is typically collected in the networks of the customers of the DDoS mitigation providers. It is unlikely that users in retail broadband networks are purchasing these kinds of services and thus those victims are severely under-counted by the industry reports. The amplifier data is much less biased towards certain types of victims. This strength does come at the cost of a weakness: missing attacks that are not amplifier-based. Earlier work suggests this is not a significant issue. Czyz et al. compared the data captured by observing NTP amplifiers against industry measurements and victim network data and they found that the patterns observed in the amplifier data were consistent with the industry measurements [1].

The contrast between our findings and industry reports are more than measurement issues. They have significant theoretical implications for our understanding the DDoS ecosystem, a point to which we will return later in the paper. We first turn to a more in-depth look at the victimization patterns in broadband ISPs and hosting.

5 Victims in Broadband Providers

We have now established that the majority of victims reside in broadband providers and that the majority of these victims are access nodes. In other words, home routers are typically the most affected devices. It suggests that the actual target is a regular home user behind that router. This brings us to the next question: *How are victims distributed over broadband networks?*

A simple count of unique victim IP addresses over the whole measurement period, does not give us a decent metric of victimization rates per ISP because of DHCP churn. ISPs re-assign IP addresses to their users at varying rates, where high rates lead to significant over-estimation of the number of victims. One method to reduce the effect of churn is to use short measurement windows [15,21]. For this reason, we count the unique number of IP addresses seen for each day and then average those daily counts to get to victimization rates over larger time frames. This results in a more accurate representation of the relative victimization rate per ISP.

In Fig. 4, we have plotted the average daily number of victims against the number of subscribers of those ISPs. The subscriber data is drawn from the TeleGeography database discussed in the previous section [16]. The database provides accurate worldwide subscriber numbers for ISPs from 77 countries that appear in our attack data. It provides a more precise proxy for the number of users in a network than technical network properties, like the number of advertised IP addresses, can provide.

As we can see, victimization rates differ by several orders of magnitude across ISPs, but these difference are highly correlated with the size of the customer base: $R^2 = 0.60$. As an aside, the correlation with the number of IP addresses advertised by each ISP also shows a firm linear relation, though a bit weaker ($R^2 = 0.56$).

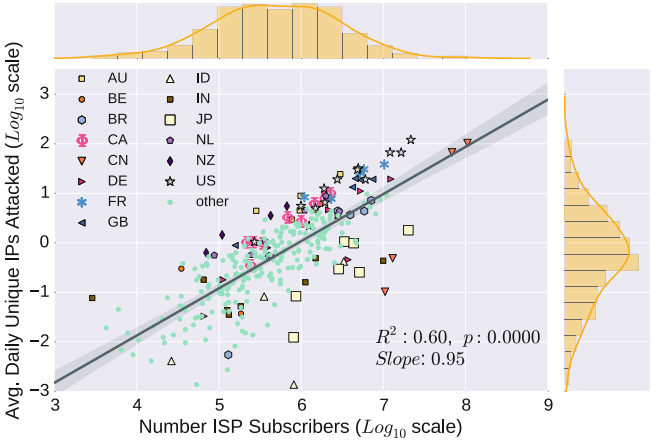


Fig. 4. Correlation access victims with ISP subscribers

In other words, the number of users is a strong predictor for the number of observed victims. This is consistent with the earlier speculation that it is individual users that are being attacked, rather than services or devices. It also means that, to some extent, victimization rates are uniform across ISPs. Whatever motivations attackers may have, it seems they select targets somewhat evenly across broadband networks.

Notwithstanding the effect of the size of the subscriber base, as captured by the regression line, the figure also clearly shows that there is significant variation around that line. That raises a new question: *why do some ISPs have disproportionately more or fewer victims?* We can use the victim rates of ISPs (i.e., the daily average number of victim IP addresses divided by the number of ISP subscribers) to further explain the variance among them. From the size-corrected victim rates we can see that there are several orders of magnitude differences among the most and least attacked ISPs. How can these differences be explained?

In Fig. 4, we have color coded ISPs by the country in which they operate. To better highlight between and within country relations, Fig. 5 plots the distribution of ISP victims per subscriber in relation to the country in which they operate. Two things become apparent. First, in many countries, ISP victimization rates are remarkably clustered, compared to the overall variance across countries. Second, ISPs in some countries are victimized less, according to our metrics. In other words, there seem to be country-level effects at work, in addition to network- and user-level effects. The plot shows that ISPs in New Zealand, Australia, U.S., U.K. and France have disproportionately more victims, while ISPs within countries such as China, Japan and Indonesia have disproportionately fewer. It is important to note that almost all ISPs in the 77 countries are present in the data, so there is no selection bias at work in these patterns.

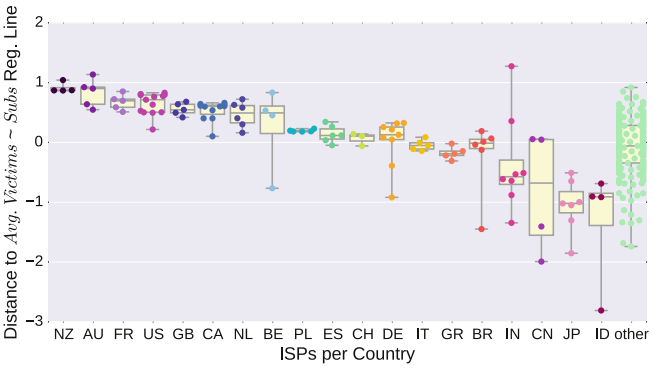
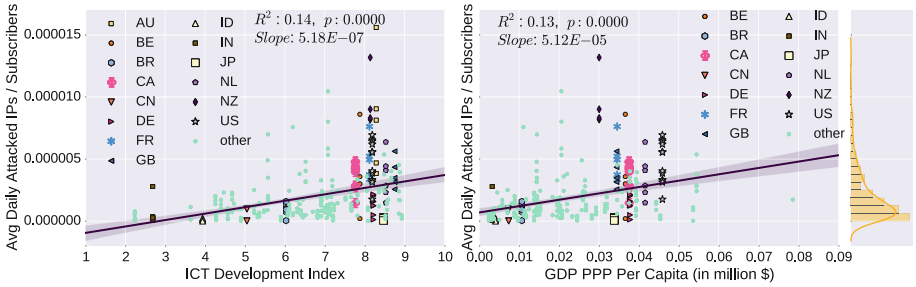


Fig. 5. Between and within country differences among ISPs

The differences between countries might be explained by institutional characteristics of the countries in which the ISPs operate. Two institutional differences that we tested for are: *(i)* the development of the ICT infrastructure of each country and *(ii)* the overall economic status of the country. In both cases we expect to observe more victims in more developed countries, as more online activity and better infrastructure might drive more motives and opportunities



(a) ICT development index (b) GDP PPP per capita

Fig. 6. Correlation of ISP victim rates with country level variables

for attacks – around online gaming, for example. The ICT development index is a well established indicator ranging from 1 to 10 with higher values for more developed countries. The index is provided by the ITU (the United Nations International Telecommunications Union) and constructed from a set of internationally agreed-upon indicators. We also looked at the gross domestic product at purchasing power parity (GDP PPP) per capita, to capture the economic status of each country [22]. From the plots in Figs. 6a and b, it is clear that both explanatory variables do correlate with ISP victim rates, but only weakly.

To consider the joint effect of the three explanatory factors that we have examined so far, i.e., the number of ISP subscribers, ICT and GDP PPP indexes, we construct several statistical models using negative binomial, generalized linear model (GLM) regression. The models predict the number of victims per ISP given a set of explanatory variables. A summary of these statistical models are presented in Table 2.

$Model_1$ only includes the attack surface size, $Model_2$ adds the ICT development index as an additional factor and finally $Model_3$ adds the GDP PPP per capita. As expected, $Model_1$ demonstrates the effect of the size of the subscriber population – i.e., the size of the 'attack surface' – in correspondence with

Table 2. Negative binomial GLM regression models with 'Log_e' link function for number of ISP victims

	Dependent variable:		
	# Victims per ISP		
	(1)	(2)	(3)
Subscribers (log_{10})	2.160***	1.996***	1.977***
	(0.079)	(0.075)	(0.074)
ICT Dev. Index (2015)		0.249***	
		(0.034)	
GDP PPP per Capita (in $1000)			0.030***
			(0.004)
Constant	−5.880***	−6.712***	−5.705***
	(0.454)	(0.468)	(0.430)
Observations	304	300	291
Log Likelihood	−2,255.880	−2,204.260	−2,128.202
θ	0.963*** (0.070)	1.097*** (0.082)	1.143*** (0.087)
Akaike Inf. Crit	4,515.761	4,414.520	4,262.404

Note: *p<0.1; **p<0.05; ***p<0.01

our earlier results (Fig. 4). The other two models demonstrate that in addition to size, the two institutional country variables considerably contribute to the variation in the number of victims per ISP, however their effects are much smaller. We interpret the results of $Model_2$ as an example. While holding everything else constant, increasing the number of subscribers by one unit (equivalent to multiplying the number of subscribers by 10 due to the log_{10} scale of the variable) multiplies the number of victims per ISP by $e^{1.996} = 7.36$. Similarly, increasing the ICT development index by one unit (while other factors are held constant) multiplies the number of victims by $e^{0.249} = 1.28$. $Model_3$ can be interpreted in a similar fashion. Note that due to the correlation of ICT development and GDP we do not include both variables in one model.

We have also examined other factors, such as 'gaming popularity' and 'piracy' which show weak correlations with victimization rates as well. Including these in separate GLM models shows a significant small effect of online gaming as captured by the average number of games owned per country on the Steam online gaming platform. This could be indicative of a possibly weak relation with online gaming and end-host victimization. However, further examination of the variable indicates strong correlations with ICT development and GDP therefore bearing little added information which the other factors did not already include in our models.

Given that the institutional factors have a weak effect, it begs the question of why, in the majority of the countries, ISP victim rates are closely clustered together. More specifically, the ISPs of only 12 of the 77 countries are dispersed by more than one order of magnitude (among them are Brazil, India, and China). Even with quite similar infrastructure and economic conditions, the differences among ISPs are larger between the countries than within them. This pattern suggests that there are specific country-level factors at work, beyond the general factors we examined.

We can only speculate why ISPs in a certain country are so clustered, but one explanation is that attackers and victims are geographically concentrated and that their interaction leads to network-effects. We know from the research on booters that many of the customers are gamers [6]. Other studies have told us that many of the victims are also related to gaming [1]. Combine this with findings from online social network analysis, inside and outside of gaming, which found that these online networks are shaped by geographical vicinity. In other words, users in online networks tend to be friends or familiar with each other in offline networks as well [23,24]. In other words, they are geographically close.

Jointly, these three factors might drive a geographically concentrated network effect: some of the victims become attackers themselves, which is easy because of the booter services. These new attackers, in turn, victimize others, and the cycle continues. This pattern fits with anecdotal evidence from news reports. In the Netherlands, for example, DDoS-ing became such a widespread phenomenon among schoolkids [13], that even those who did not play online games started to use booters, because everyone was doing it. One more technically skilled

youngster said he quit DDoS-ing, as "it became too easy" and "even my sister can do it" [25].

Overall, our findings reveal that the number of subscribers of ISPs is a very strong predictor for the number of victims per ISP (see Fig. 4). This result suggests that the chances of being victimized are surprisingly uniform across ISPs. The accessibility of DDoS-as-a-service might have caused a democratization of victims: everywhere there are now regular users deemed worthy of attack. This is a far cry from the highly publicized attacks on high profile targets like governments and enterprises. Those are attacked too, of course, but the bulk is targeted at regular netizens.

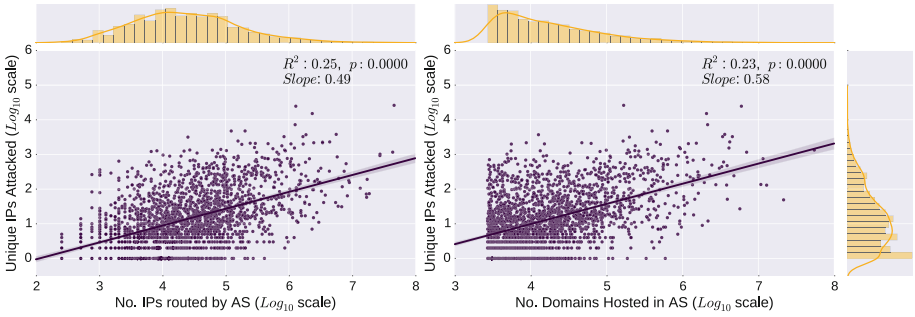
That being said, we do see significant variation in terms of victimization rates. The country-level differences are partially explained by institutional factors and partially by specific country-level effects. In the absence of direct evidence, we speculated that the remaining variation might be driven by geographically concentrated network effects.

6 Hosting Providers

In this section we take a closer look at victims located in hosting provider networks. As for ISPs, the main questions at this stage are: *How are victims distributed across different hosting ASes* and *Do some hosting providers have disproportionately more victims than others?*. Unlike broadband victims, we do not expect the dynamic nature of IP allocation to significantly effect or lead to a misrepresentation of the number of victims. Therefore we can examine the distribution of victims over networks by simply counting the number of unique victim IPs that we observe per AS.

As with broadband networks, we expect differences in customer base or network size to correlate with the number of victims. To test this, we need to estimate the size of the hosting providers. One approximation is to use the number of hosted second-level domains (2LDs) per each provider. Recall, however, that we found that only 46.6 % of the hosting victim IPs have been observed to host domains. This implies that the number of domains will not be a very reliable approximation of the attack surface size. We can use the number of routed IP addresses by each hosting provider as a second proxy for size. This metric, however, is less able to account for shared hosting (several 2LDs sharing the same IP address). As we will see below, using both proxies in combination gives the best results.

Figures 7a and b plot the number of unique victim IPs per hosting provider against the number of routed IP addresses and hosted 2LDs of the provider respectively. Both figures demonstrate a moderate effect of attack surface size on the number of victims, but size does not appear to explain a large portion of the variance as indicated by the relatively lower R^2 values. This simply means that only a small part of the variation among hosting ASes is explainable purely through the attack surface size. We can see that some hosting ASes are disproportionately attacked more (data points far above the regression line) or less



(a) Routed IPs (b) Hosted domains

Fig. 7. Correlation hosting victim counts with size estimates.

(data points far below the regression line) in relation to their size. This signals that attacks on hosting providers are also quite strongly driven by other explanatory factors. The question to consider then is *what additional factors can explain the variation that we observe after the size effect has been corrected for*? As before correcting for size effects can be achieved through dividing the number of victims per provider by the size estimate of the provider.

One possible non-size related explanatory factor that we consider is related to the popularity of the hosted content. The expectation here is that more popular content is more likely to be attacked. In our analysis we use the list of top 1 million Alexa ranking domains as a proxy for the popularity of the hosted content [26]. Given the 2LDs that we have identified per hosting provider, we use the median ranking of the subset of top 1M Alexa ranked domains as an indicator of popularity. Note that in our analysis we use reversed rankings: the most popular Alexa domain has the rank of $1,000,000$.

A second possible factor that we consider is the type of hosting service that is offered. We expect that dedicated hosting is more likely to be attacked in comparison to shared hosting and other similar cheaper services offered by hosting providers. We use the number of IP addresses that have been used by the hosting provider to host all of its 2LDs as an indicator of the type of hosting. This indicator combined with size estimates (routed IPs and hosted 2LDs) captures the spread/density of domains per available IP address. A lower density of domains per IP is an indication for more dedicated services to their customers, while higher densities are indicators of shared hosting.

Our analysis of these non size-related factors demonstrates a weak correlation with the number of victims per provider after correcting for size effects. For the sake of brevity we do not include the details and instead move on to consider the joint effect of all explanatory factors.

In a similar fashion to what we did for broadband victims, we construct several statistical models of the number of victims per hosting provider using negative binomial GLM regression. A summary of these models is presented in Table 3. They clearly demonstrate that for larger attack surfaces there are more victims.

Table 3. Negative Binomial GLM regression models with 'Log_e' link function for number of Hosting Victims

	Dependent variable:		
	# Victims per Hosting Provider		
	(1)	(2)	(3)
f_1: Routed IPs	1.198^{***}		0.507
(log_{10})	(0.040)		(0.354)
f_2: Hosted Domains		1.237^{***}	1.050^{***}
(log_{10})		(0.050)	(0.243)
f_3: IPs with Domains			−0.415
(log_{10})			(0.427)
f_4: Median Alexa Rank			0.305^{***}
(log_{10})			(0.075)
$f_1 \times f_2$			-0.338^{***}
(Interaction term)			(0.088)
$f_1 \times f_3$			0.266^{***}
(Interaction term)			(0.044)
$f_2 \times f_3$			0.198^{**}
(Interaction term)			(0.084)
Constant	-1.120^{***}	-0.988^{***}	-3.859^{***}
	(0.177)	(0.215)	(1.093)
Observations	2,203	2,203	2,203
Log Likelihood	−10,594.160	−10,703.310	−10,192.260
θ	0.421^{***} (0.011)	0.393^{***} (0.010)	0.546^{***} (0.014)
Akaike Inf. Crit	21,192.330	21,410.620	20,400.520

Note: *p<0.1; **p<0.05; ***p<0.01

$Model_3$ uses all variables to explain the variance in victimization of hosting providers. Due to the unavoidable correlations between these variables we include interaction terms which control for the covariance between them. The model demonstrates that when considered jointly, the number of hosted 2LDs and the popularity of content have a significant effect on the number of victims per hosting provider. As expected, the size-related factor has the largest effect while the popularity of content as represented by the median Alexa rank is moderately affecting the victim numbers. It also suggests that there is not enough evidence to support the hypothesis that the density of domains or type of hosting has a significant effect on victim numbers. Due to the inclusion of interaction terms, $Model_3$'s results need to be interpreted in a slightly different manner. The more complex and improved model (as indicated by the improved log likelihood) suggests that while holding all other factors constant, increasing the 'Hosted Domains' variable by one unit (equivalent to multiplying the number of hosted 2LDs by 10 due to the log_{10} scale of the variable) multiplies the number of victims by $e^{1.050-0.338+0.198} = 2.48$. Increasing the 'Median Alexa Rank' variable by one unit (equivalent to multiplying the median Alexa rank of the content by 10 due to the logarithmic scale) multiplies the number of victims by $e^{0.305} = 1.35$. Finally, note that in $Model_3$ the number of routed IPs is not a significantly contributing factor. This does not negate the size effect as observed in $Model_1$ and simply means that when considered jointly with the other factors the number of routed IPs does not add more information to the model that has not been already captured by the other included factors. Based on these results we can conclude that in addition to size factors which have the strongest effect

on the number of victims per hosting provider the popularity of content also weakly contributes to this number.

To get a better sense of the actual victims, we have taken a closer look at some of the hosting victims that are associated with domain names. Many IP addresses are associated with multiple domains, obscuring the target and potential motive of the attackers. However, a subset of around 23,855 IP addresses are associated with just a single domain name according to our passive DNS data. We took a random sample of 1 % of this set (238 domains) and checked all of them manually to assess what type of website was being attacked. Of the 238 domains, 107 no longer showed any content. Most of them could no longer be resolved, others ran into connection issues or were replaced by parking pages. Given that the victim data was collected over two years, some degree of 'link rot' is to be expected, though this decay of domains is much higher than those found in other studies (e.g. [27]), suggesting that a lot of the victims had a somewhat fleeting presence on the web, rather than being well-established businesses or organizations.

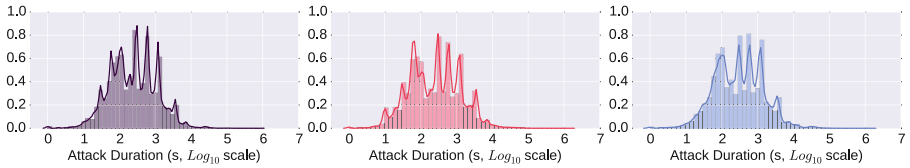
Of the 132 sites that offered content, 55 sites (42 %) were directly related to gaming. Of these, 27 were associated with a single game: Minecraft (17), followed by Counterstrike (6) and Runescape (4). The remaining 77 sites (58 %) were highly heterogeneous, including but not limited to a few large stores, an airline, two football clubs, two schools, two escort services, one porn site, several hobby forums, a casino, a nature conservancy, and Twitpic, owned by Twitter since late 2014. In short: motives for DDoS attacks are highly varied, though gaming-related feuds are the most dominant of motives. In the Minecraft community specifically, DDoS attacks seem to be part of the culture.

We can summarize our results with respect to hosting providers as follows. Hosting providers constitute the second largest group of victims in the amplification honeypot data. Some providers are attacked disproportionately more than others. This can be partially explained by the size of their attack surface. Furthermore, hosting popular content increases the number of victims. Finally, in agreement to what others have also found we see a large victimization of gaming related resources within the hosting providers.

7 Attack Duration

In previous sections we have examined the question of who gets attacked more, whether that is disproportionate and if some factors can explain the variance among victim counts. We can also approach the question of who gets attacked more from the view point of time. That is, rather than looking at victim counts we can also approach the question as *who gets attacked longer and possibly why?*

To answer these questions, we take all victim IP addresses and measure the intervals under which they were continuously attacked. These intervals are calculated regardless of which AmpPot or protocol was used to attack the victim IP. The resulting interval lengths are defined as the attack duration. Note that here, we have merged attacks that are closer than 600 s apart and consider them as one continuous attack on the victim. Given these durations, the primary



(a) Broadband ISP (b) Hosting (c) Other

Fig. 8. Distribution of attack durations for various victim types.

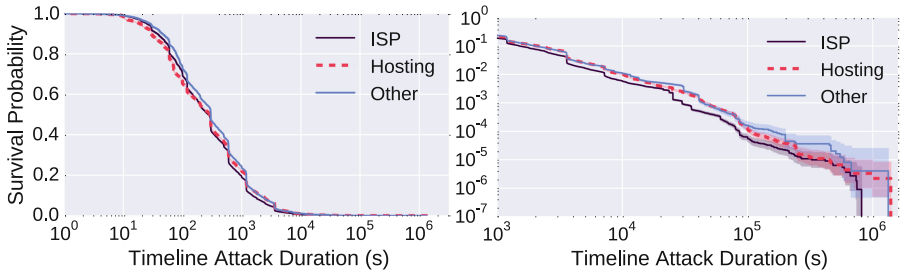
question is *whether the distribution of these durations differs per victim type.* These distributions are shown in Fig. 8.

The median attack duration for broadband ISPs, hosting and the other types of victims are 272, 285 and 300 s, respectively. One surprising observation is the frequency of relatively short attack durations. The majority of attacks are shorter than 286 s long. For attacks longer than 300 s, we observe similar distributions of attack durations for all three types of victims. Interestingly, we observe an increased number of attacks that last around 5, 10, 20, 60, or 120 min. The trend suggests that, in general, the attacks are largely originated from booter services and are most possibly driven by attackers' capabilities rather than victim types (see Fig. 8).

To further compare the differences in durations for different victim types, we use a well established statistical technique that is commonly referred to as survival analysis. The technique is used to answer questions about the proportion of a population that will survive past a certain point of time on a measurement timeline and at what rate the individuals 'survive' or 'die'. In our case, the event that we analyze is the '*end of an attack*' on a victim IP. Figure 9 demonstrates our survival analysis results. We use the Kaplan-Meier estimator to approximate the survival function [28], measuring the probability of an attack exceeding a certain duration for various victim types.

A log-rank comparison of the survival probabilities indicates a significant difference at a 0.99 confidence level between attack durations on different victim types. The log-rank chi-square statistic comparison between broadband/hosting, broadband/other and hosting/other are equal to 2,131.8, 3,493.4, and 739.3 respectively. These results indicate a significant difference among the attack durations per victim type, however in terms of magnitude, the differences seem to be quite small (see Fig. 9).

We can also compare the survival rates of each victim type using the Cox proportional hazards model. The Cox model does not depend on distributional assumptions of survival time and allows to estimate the hazard ratio defined as the relative risk based on a comparison of event rates. The hazard ratios show that relative to hosting providers, attacks end 14 % faster for broadband victims while 3 % slower for the other type of victims. While the results demonstrate that attacks are statistically shorter on broadband ISP victims the magnitudes of the differences are not large enough to have significant implications.



(a) All (b) 10% longest attacks

Fig. 9. Survival analysis of attack durations

To conclude, all victim types experience attacks ranging from short lived attacks in the order of several seconds to long attacks which last several days. In other words, there is no significant variance among the duration of attacks on victims of different types.

We have further manually analyzed victim IP addresses of the 100 longest attacks of which 98 lasted more than 24 h. They were launched against 87 unique IP addresses and 46 unique ASes. Interestingly, we do not observe any domains historically hosted on as many as 41 IP addresses (47 attacks). Of these, 6 IP addresses were directly related to gaming, including two victims against which the attacks lasted more than 16 days. Of the remaining 46 unique IP addresses, which were identified to be hosting some content, 17 were mapped to just a single domain name in passive DNS data. Of these, we have identified 6 victim IP addresses that hosted websites which provide torrent files to facilitate P2P file sharing, 4 websites related to gaming, 2 chat websites, one Internet banking website, and one TorGuard VPN website. By manual analysis of 15 IP addresses for which we observed 2 or 3 domains, we have further identified three victim IP addresses that mapped mainly to torrent, gaming, and TorGuard websites, respectively. The remaining 14 victim IP addresses mapped to more than 3 domains; 4 among them appeared to be used for shared web hosting and they mapped to 51, 346, 614, and 931 domains. To conclude, our manual analysis reveals that not only gaming but also torrent sharing-related IP addresses are among long-duration attacked victims.

8 Related Work

Much research has been devoted to analyzing the technical properties of amplification DDoS attacks: which protocols can me misused and how; how large the population of vulnerable reflectors is; how difficult or easy it is to find and misuse these reflectors; and how they could be mitigated [1,12,29,30]. We know for example that many UDP based protocols are prone to be misused (`NTP`, `DNS`, `SNMP` and `Chargen`) and we know what their amplification factors are [12]. We also know how large the populations of vulnerable devices running these

protocols are [1,5,12] and what kind of a threat they pose. Darknet and honeypot traffic reveals how perpetrators are actively scanning for such devices in the wild [1,2,12,31]. Some have even attempted attacking their own infrastructure in order to asses the potential damage of booters and surprisingly find their damage to be much smaller than the spectacular cases reported in the news [13]. Others have examined the motives behind the provision of booter services through interviews [32]. Analysis of trends also reveals how over time specific protocols rise and fall out of popularity among attackers and how remediation and intervention has affected the landscape [1,8].

Earlier work on amplification DDoS attacks have focused less on studying the victims. The most in-depth understanding comes from the special case of `NTP` attacks, which allows probing the amplifier for victim IP addresses. Czyz et al. [1] provided the most comprehensive overview. The analysis of the smaller subset of victims from leaked booter databases [6,7] also point towards gaming-related victims. We corroborate earlier findings, especially [1,8], that many of the victims are end hosts and gaming-related resources, but we also expand on this and show that the distributions have shifted. Moreover, we provide a wholly novel contribution by developing victimization rates and providing an explanatory analysis of key determinants behind victimization patterns.

Finally, part of what we know about victims is based on industry reports from DDoS mitigation providers [9–11,14]. These mostly provide information on the type of industry that is affected most by DDoS attacks and point to the gaming industry and software industry as main victims. Our results paint a rather different picture, agreeing only with those reports in that many victims are gaming-related. Industry reports seems to be vulnerable to biases related to the fact that data collection often takes place in networks of the customers of DDoS mitigation providers.

9 Discussion and Implications

This study has presented the first in-depth look at victimization patterns of DDoS amplification attacks - and thus of the booter services that drive the bulk of these attacks. We found that broadband networks harbored most of the victims (62 %), followed by hosting networks (26 %). Educational, governmental and enterprise networks make up just a small fraction of the victim population (12 %), contrary to industry reports and news items about high-profile attacks.

The population of victims is predictably distributed across broadband and hosting networks. To a large extent, the size of the user population drives the victimization rate – in broadband around 60 % of the variance in victim counts can be explained from just the number of subscribers of the provider. Further explanatory factors are ICT development and GDP per capita. We also see significant differences among countries, however, that are not explained by these institutional factors. Remarkably, within most countries, ISP victimization rates are clustered together. This implies there are specific country-level effects at play, perhaps the result of geographically concentrated network effects among attackers and victims.

In hosting provider networks, the size effect is also visible, though less pronounced. The popularity of content, as measured by Alexa rankings, had a small effect. When we looked at victims IP addresses associated with a single domain, we found that 42 % of the sites we could identify were related to gaming, most notably to Minecraft.

Attack duration did not differ significantly across the victim populations. When we examined the 100 longest attacks, 98 of which lasted more than 24 h, we found, again, mostly gaming-related content rather than high-profile targets.

What do these findings mean for the consequences of the so-called commoditization of DDoS attacks? Rather than going after high-value targets, DDoS-as-s-service has invited attackers to go after regular users. With the commoditization of attacks, victimhood has democratized. And so has criminality, in all likelihood. Assuming that the users are targeted by someone that actually knows them, rather than by a random stranger, our findings imply that the attacker population has also broadened. In short, booters have indeed drawn more attackers into the DDoS ecosystem, as the commoditization theory suggests, and this has led to a an expansion of victims among regular users, who now make up the bulk of all victims.

Overall, the fact that most victims are regular users suggests that profit is not a dominant motive anymore, assuming it ever was. The commoditization provided by booters has enabled attacks for as little as one U.S. dollar. This type of cybercrime is priced in the same range as, or even below, many entertainment products. It is now cost-effective to pursue many more motives than profit, even very frivolous ones – like harassing your schoolmates during Minecraft games or online chats. Many of the new attackers probably do not see themselves as cybercriminals. Everyone is doing it, and they are not making any money from it.

The fact that attack patterns are so proportional to the number of users might seem unsurprising, but it has far-reaching implications. Rather than a phenomenon of motivated attackers with specific objectives and targets, DDoS has become a cultural phenomenon. The closest parallel to the observed pattern seems to be wide-spread use of torrents and file lockers to download copyright-infringing materials. This suggests a new route of action for fighting the DDoS problem: rather than using criminal law to go after motivated attackers, a better approach might be what criminologists call *situational crime prevention* [3]. It shifts the focus from identifying and penalizing attackers to taking away the opportunities that trigger crime. It can draw on a much broader mix of measures, often based on civil rather than criminal law. It can range from soft measures, such as awareness campaigns for youngsters, to harder ones, like the takedown of booter accounts by providers such as PayPal [8].

What are the implications of our findings for the wider commoditization of cybercrime? Should we expect an influx of attackers and an expansion of victims in other criminal markets as well? Not per se. As Florencio and Herley have argued, cybercrime is often harder than it looks and it scales less well than one would assume at first glance [33,34]. Indeed, in many markets, we do not see the rapid expansion of crime that effective commoditization would cause. This can be explained by the fact that many of these service models do not supply

complete criminal value chains. Take fraud using banking Trojans for example. It is one thing to buy malware-as-a-service and distribute it via pay-per-install, but that doesn't mean one can successfully execute online banking fraud. There are bottlenecks elsewhere, especially in the use of money mules and other cash-out channels. Mules-as-a-service did not manage to solve this bottleneck yet.

We see the predicted effects of commoditization in DDoS attacks, because here the booter provides the value chain end-to-end. In other forms of cybercrime this seems much harder or even impossible, though some might come close, like ransomware-as-a-service using bitcoin. And indeed, we did recently see an explosion of ransomware attacks. We can only hope that for many other forms of cybercrime, bottlenecks will remain resistant to successful commoditization.

Acknowledgements. This work has been enabled through the support of NWO Pr. Nr. CYBSEC.12.003/628.001.003, SIDN, the Dutch National Cyber Security Center and Beatriu Pinos BP-A-214. We would like to thank Dr. Paul Vixie and Farsight Security for providing our pDNS data. In addition we would like to acknowledge the support of the MEXT (Program for Promoting Reform of National Universities) and PRACTICE (Proactive Response Against Cyber-attacks Through International Collaborative Exchange) programs.

References

1. Czyz, J., Kallitsis, M., Papadopoulos, C., Bailey, M.: Taming the 800 Pound Gorilla: the rise and decline of NTP DDoS attacks. In: Proceedings of ACM IMC, pp. 435–448 (2014)
2. Krämer, L., Krupp, J., Makita, D., Nishizoe, T., Koide, T., Yoshioka, K., Rossow, C.: AmpPot: monitoring and defending against amplification DDoS attacks. In: Bos, H., et al. (eds.) Raid 2015. LNCS, vol. 9404, pp. 615–636. Springer, Heidelberg (2015). doi:10.1007/978-3-319-26362-5_28
3. Thomas, K., Yuxing, D., David, H., Holt, T.J., Kruegel, C., Mccoy, D., Bursztein, E., Grier, C., Savage, S., Vigna, G.: Framing dependencies introduced by underground commoditization. In: WEIS (2015)
4. Santanna, J.J., Sperotto, A.: Characterizing and mitigating the DDoS-as-a-Service phenomenon. In: Sperotto, A., Doyen, G., Latré, S., Charalambides, M., Stiller, B. (eds.) AIMS 2014. LNCS, vol. 8508, pp. 74–78. Springer, Heidelberg (2014)
5. Kuhrer, M., Hupperich, T., Bushart, J., Rossow, C., Holz, T.: Going wild: large-scale classification of open DNS resolvers categories and subject descriptors. In: Proceedings of ACM IMC (2015)
6. Karami, M., Mccoy, D.: Understanding the emerging threat of DDoS-As-a-Service. In: Proceedings of Usenix LEET, pp. 2–5 (2013)
7. Santanna, J.J., Durban, R., Sperotto, A., Pras, A.: Inside booters: an analysis on operational databases. In: Proceedings of IFIP/IEEE IM, pp. 432–440 (2015)
8. Karami, M., Park, Y., McCoy, D.: Stress testing the booters: understanding and undermining the business of DDoS services. In: Proceedings of WWW (2016)
9. Akamai: State of the Internet / Security Q4. Technical report Akamai (2014). https://www.stateoftheinternet.com/
10. Arbor Networks: Worldwide infrastructure security report volume X. Technical report (2015). https://www.arbornetworks.com/insight-into-the-global-threat-landscape

11. Incapsula: DDoS global threat landscape report. Technical report (2015). http://lp.incapsula.com/ddos-report-2015.html
12. Rossow, C.: Amplification Hell: revisiting network protocols for DDoS abuse. In: Proceedings of NDSS, pp. 23–26 (2014)
13. Santanna, J., Van Rijswijk-deij, R., Hofstede, R., Sperotto, A.: Booters - an analysis of DDoS-as-a-Service attacks. In: Proceedings of IFIP/IEEE IM (2015)
14. Kaspersky: Statistics on botnet assisted DDoS attacks (2015). https://securelist.com/blog/research/70071/statistics-on-botnet-assisted-ddos-attacks-in-q1-2015/
15. Asghari, H., van Eeten, M.J.G., Bauer, J.M.: Economics of fighting botnets: lessons from a decade of mitigation. IEEE Secur. Priv. **13**(5), 16–23 (2015)
16. TeleGeography: Telegeography globalcomms data. http://shop.telegeography.com/products/globalcomms-database
17. CAIDA: AS classification. http://www.caida.org/data/as-classification/
18. Farsight Security: DNSDB. https://www.dnsdb.info
19. Tajalizadehkhoob, S., Korczynski, M., Noroozian, A., Ganan, C., van Eeten, M.: Apples, oranges and hosting providers: heterogeneity and security in the hosting market. In: Proceedings of IEEE/IFIP NOMS, pp. 289–297 (2016)
20. Akamai: State of the internet/security Q4. Technical report (2015). https://www.stateoftheinternet.com/downloads/pdfs/q4-2015-securityreport-ddos-stats-trends-analysis-infographic.pdf
21. Asghari, H., Ciere, M., Van Eeten, M.J.G.: Post-Mortem of a Zombie: conficker cleanup after six years. In: USENIX Security (2015)
22. PRB. Population Reference Bureau - Gross Domestic Product. http://www.prb.org/DataFinder/Topic/Rankings.aspx?ind=260
23. Ledbetter, A.M., Kuznekoff, J.H.: More than a game: friendship relational maintenance and attitudes toward Xbox LIVE communication. Commun. Res. **39**(2), 269–290 (2012)
24. Allamanis, M., Scellato, S., Mascolo, C.: Evolution of a location-based online social network. In: Proceedings of ACM IMC, p. 145. ACM Press, New York (2012)
25. Schravese, F., Born, A.: Lekker thuis providers platleggen (2015). http://www.nrc.nl/handelsblad/2015/10/17/lekker-thuis-providers-platleggen-1545974
26. Alexa: Alexa top 1M ranked sites (2015). http://s3.amazonaws.com/alexa-static/top-1m.csv.zip
27. Zittrain, J., Albert, K., Lessig, L.: Perma: scoping and addressing the problem of link and reference rot in legal citations. Legal Inform. Manage. **14**(02), 88–99 (2014)
28. Kaplan, E.L., Meier, P.: Nonparametric estimation from incomplete observations. J. Am. Statist. Assoc. **53**(282), 457–481 (1958)
29. Kuhrer, M., Hupperich, T., Rossow, C., Holz, T.: Exit from Hell? Reducing the impact of amplification DDoS attacks. In: USENIX Security, pp. 111–125 (2014)
30. Kuhrer, M., Hupperich, T., Rossow, C., Thorsten Holz, G.: Horst: Hell of a handshake: abusing TCP for reflective amplification DDoS attacks. In: Proceedings of USENIX WOOT (2014)
31. Durumeric, Z., Bailey, M., Halderman, J.A.: An internet-wide view of internet-wide scanning. In: USENIX Security, pp. 65–78 (2014)
32. Hutchings, A., Clayton, R.: Exploring the provision of online booter services. In: Deviant Behavior, pp. 1–16 (2016)
33. Florencio, D., Herley, C.: Where do all the attacks go? In: Economics of Information Security and Privacy III, pp. 13–33 (2013)
34. Florencio, D., Herley, C.: Is everything we know about password- stealing wrong? IEEE Secur. Priv. Mag. **10**(6), 63–69 (2012)