Assignment 3: Underlying factors that explain performance differences

WM0824TU Economics of Cybersecurity (2019)

Brennen Bouwmeester, Hsin Cheng, Kevin Su, and Jochem Vlug

Group 5 (DDoS), Delft University of Technology

Abstract. This paper discusses the underlying reasons for differences in security performance against DDoS attacks by hosting providers. Using the AmpPot dataset, hosting providers are compared and statistical analysis on possible explaining factors shows the reason differences exist. These factors are linked to incentives and possible countermeasures by actors in the DDoS attacks arena.

Keywords: Distributed Denial-of-Service (DDoS) · Performance metrics · Underlying factors · Externalities · Incentives for countermeasures

1 Introduction

Distributed Denial-of-Service (DDoS) attacks are considered the top operational threat in the cyber landscape (Holl, 2015; Cheung, 2017). These types of attacks overload the capacity of the target service or its surrounding infrastructure through a flood of illegitimate traffic, aiming to disrupt normal traffic and make the system unavailable for users. Disruptions resulted from DDoS attacks have caused millions in revenue losses, reputation damage, and customer attrition to many industries (Cardoso de Santanna, 2017; Chromik, Cardoso de Santanna, Sperotto, & Pras, 2015; Cheung, 2017).

Even though DDoS attacks have been around for a long time, the emerge of *Booters*, or *DDoS-as-a-service* or *DDoS-for-hire* – along with the use of amplification technique – has deteriorated the DDoS attack problem not only in the attack frequencies but also in attack power (Cardoso de Santanna, 2017; Noroozian et al., 2016). Nowadays, people are able to buy DDoS attack services with very affordable prices at the Booter websites, which are easily reachable through popular searching engine such as Google and Bing (Cardoso de Santanna, 2017). The commodification of DDoS attack has removed the technical barrier to perform attacks; together with the introduction of amplification techniques, the occurrence of reported DDoS attacks has been escalating rapidly (Chromik et al., 2015; Cardoso de Santanna, 2017).

In this assignment, the performance of different victims in protecting their availability from DDoS attacks will be compared. The variance in this performance will then be analyzed and the reason for the existence of this variance will be discussed. In other words, we will look at the underlying factors that explain the different performances by victims. To get to these factors, first possible countermeasures by three different actors in the arena were analyzed and also their incentives to perform these countermeasures or not. Adding to that, the distribution of costs and benefits are discussed for each of the countermeasures. This will be the base for finding underlying factors for the difference in performance.

Aiding this analysis. is the empirical data from the AmpPot dataset (Krämer et al., 2015), which provides not only data on the different performance by victims, but also some possible explaining factors.

2 Existing countermeasures and incentives

This section discusses the existing countermeasures of three actors. These measures will be quantified with a qualitative benefits and cost analysis. This helps to give insight into the incentives of the actors and the externalities of the measures.

These three actors have been chosen because of their importance and the diversity of their measures in the DDoS problem field. Hosting providers, also the problem owner for this study, are a crucial link in the protection of DDoS attacks. Availability of services play an important role in their business model and attacks disrupt their main services. Internet service providers form a powerful actor because they act as an intermediary for DDoS attacks. Their benefits and costs will be interesting to analyze. Governmental institutions are interesting because their measures are less technical but more focused on regulation and liability, which differs from the measures that other actors could take.

2.1 Hosting providers

Hosting providers host servers on the internet with a diverse range of services, such as email, websites, and data storage. While these are different services, they still share similar structure types so the countermeasure could be applied to everyone.
The countermeasure taken in this case is implementing Deep Packet Inspection (DPI). Brook (2018) explains DPI as a form of packet filtering, through inspection it is possible to block any malicious packets. Figure 1 shows that normal filtering only captured the header of a packet with limited information, DPI also includes information on the data of the packet (Dharmapurikar, Krishnamurthy, Sproull, & Lockwood, 2003).

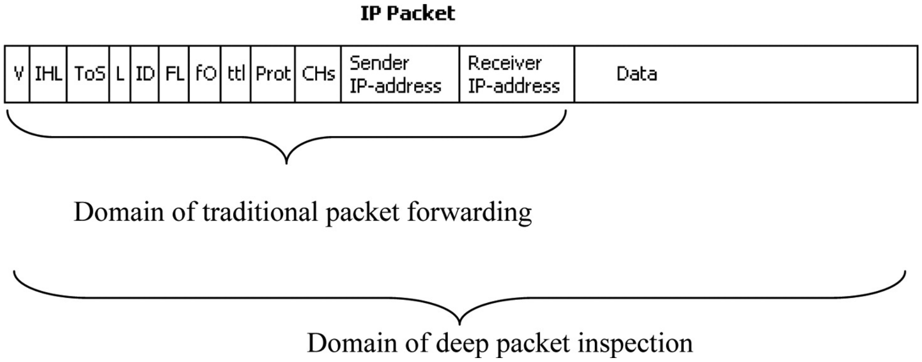

Fig. 1: DPI Diagram from (Bendrath & Mueller, 2011)

The costs and benefits of deep packet inspection is still a topic of discussion. The costs of the system would be for the hosting provider themselves, however, this service is beneficial to the clients of the hosting provider and indirectly society. It is important to discuss the perspectives of other domains as well. DPI still remains politically contested due to the concerns in privacy and strong regulation.

(Asghari, van Eeten, Bauer, & Mueller, 2013). While it improves the security of hosting providers, this comes at the cost of privacy for legitimate internet traffic.

The countermeasure is interesting for the hosting provider to perform. By looking at their existential goal, which is creating business and profit, their business model relies on availability of their servers. Disruption due to a DDoS attack would cause losses in profit, seeing reliability is an criteria for a hosting provider, and profits rely upon availability of its services. The burden is on the shoulders of the hosting providers because they cannot rely on others to protect them without certain agreements in place.

The externalities of this problem should be discussed to understand why it is that DDoS remains a problem. These externalities can be described as collateral positive effects or damages. The problem has both positive and negative externalities. In essence, the positive externality is a benefit to third parties as a consequence of an action. It should be noted that actions of other parties could have helped the hosting providers in mitigating the DDoS attacks. If a new type of DDoS attack is mitigated successfully, parties can all rely on the found technique of defending against the attack. Furthermore, as a rule of commons, when botnets target one party, another cannot be attacked. Security increases if botnets decrease in size, less and less powerful attacks take place, which is beneficial for Hosting providers, ISP's, law enforcement and society in general. However, no one wants to pay the price for tracking down and decreasing botnets in size.

Furthermore, the negatives externalities mean harm imposed on third parties as a consequence. Van Eeten and Bauer (2009) mention botnets which create a lot of negative externalities. The infected computers are responsible for a majority of the DDoS attacks, however, the users of these infected computers are often not aware that they are part of a botnet, let alone that they are damaging other networks. In essence, this makes them an accomplice to a crime without their consent or knowing.
Another negative externality of DDoS attacks is the vulnerability of hosting providers. In the USA, a profit-led country, hosting providers only invest in security if it is necessary or it proves to be profitable, this leave companies unnecessarily insecure (Gordon, Loeb, Lucyshyn, & Zhou, 2014).

2.2 Internet Service Providers

Internet Service Providers are the intermediaries who give access to the internet. They have a wide range of countermeasures to battle DDoS attacks, such as blackholing, scrubbing, traffic engineering, and local filtering (Vink, 2016). The countermeasure discussed here is RFC 2827 (ingress filterting) or more commonly known as BCP38, as it tackles the root of the problem (Ferguson, 2000). BCP38 prevents attackers to spoof their IP-adresses. It will eliminate amplification/reflection attacks, which were responsible for the highest volume attacks so far, thus stopping the attacks from starting at all. Figure 2 explains that the amplifiers will not be able to target the victim without a spoofed IP-address. For other types of attacks, it will allow law enforcement to be able to trace back the source of an attack to the attacker.

The benefits for the countermeasure seem clear: some attacks are prevented before they can start and it will be easier to track to the real source. Damas and Karrenberg (2008) focuses on other benefits as well, such as increased confidence in DDoS traffic analysis. This will help to gain better insight into the risks of the problem. However, these benefits are not necessarily for the ISPs themselves. These are mostly aimed at society as a whole. ISPs will see that the (malicious) traffic on their network

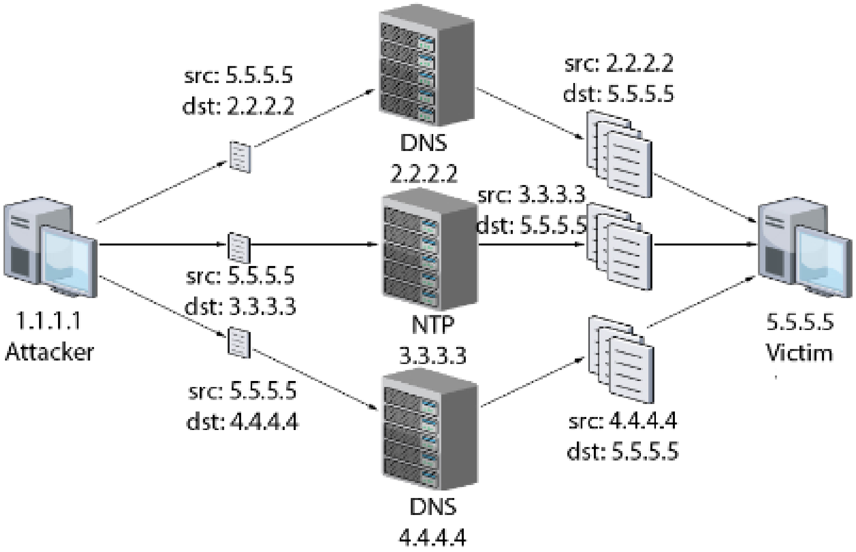

Fig. 2: Explanation of how amplifying DDoS attacks work

decreases, which will also save money. The costs for this countermeasure is for the internet providers themselves, they are the ones paying for the hardware, software maintenance, and labor to implement this. While hardware has decreased in price, maintenance and labor remains difficult (McConachie, 2014).

Discussing incentives on internet service providers remains very difficult, as it is not clear what the net effect is for different types of ISPs (Van Eeten, Bauer, Asghari, Tabatabaie, & Rand, 2010). Figure 3 shows that anti-spoofing measures differs a lot per country which means incentives for ISPs cannot be applied generally. Not all ISPs have implemented BCP38, where some attribute this to the *tragedy of the commons* phenomenon (McConachie, 2014). Implementation needs to be done on a wide scale, meaning it will only be beneficial if everyone does it. However, it seems that there is also an intrinsic reasoning why they would do it. Not having malicious traffic through their systems is beneficial to ISPs as well. As Bauer and Van Eeten (2009) mentions, there are many different security-enhancing and security-reducing incentives for ISPs; there will be some level of mitigation measures, however, it is not guaranteed that the level is the social optimum.

The externalities of the security issue for the internet service providers will become more explicit. Their role as intermediaries mean that they are the intermediary of unintended consequences in DDoS attacks. An example of this is the increase in bandwidth and technology of ISPs allowing them to handle more traffic. The externality here is that the DDoS attacks will become bigger, allowing for bigger disruptions and other actors need to implement more security measures.

2.3 Governmental institutions

Turning towards the governmental institutions, both the **EU** and the **national government** of countries can influence the existing risks, either by creating laws that keep other actors from accepting risks, or by setting standards in risk mitigation, that the other actors need to meet. This could be a law that states

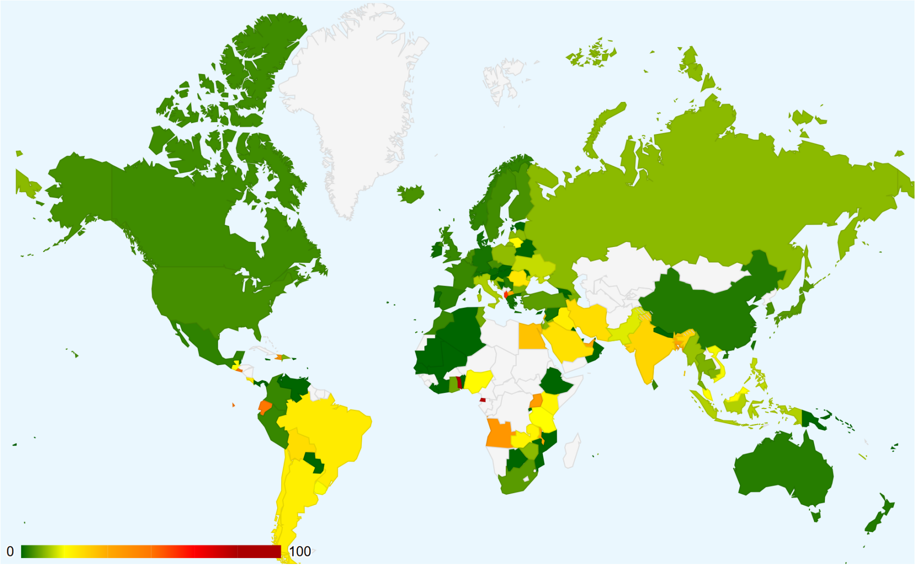

Fig. 3: Percentage of spoofing allowed in a country from CAIDA (2018)

at least a certain percentage should be spent on security, or that the probability of losing this availability should be no higher than a certain percentage per year. A third thing the governmental organizations can do is place some part of the responsibility on the desk of technology vendors. This way, risk transfer takes place from hosting providers and consumers towards these technology vendors. This could mean that the actor puts more effort into creating stable and safe connections (Assaf, 2007).

The countermeasure the government enforces would be to change the distribution of responsibility and not necessarily implement technological measures themselves. The government focuses on actors that are not necessarily affected by the attacks but play a key role in the attacks, such as intermediaries. They will have to bear the costs instead of the actual victims. It becomes problematic if this distribution is not applied internationally, meaning certain actors will be at a disadvantage in the country.

The government has the incentive to protect society from disruptions in society. A complication that may arise is that if regulation is not applied all over the world, it may slow down technological innovations and advancements. As not every company would invest in security measures so that they will be able to sell services in the country.

Externalities caused by these mitigation strategies are both positive and negative. By forcing hosting providers to invest in security measures, attack rates will fall since the security is better. This has a gained benefit since the "morale" of attackers might fall, if they no longer succeed in attacking successfully. Study shows, many attackers are teenagers drive by cultural or emotional triggers (De Cuyper & Weijters, n.d.), this type of attacker can be demotivated if DDoS attacks become harder or very expensive to perform.

Negative externalities regarding governmental regulations mostly focus on the costs of implementation. Forcing the market to invest in cybersecurity comes with big investment costs, which may or may not be necessary. According to Gordon et al. (2014), there is an upper limit to security investments to be efficient. Furthermore, transferring risks can have severe negative impacts on legality of the issue;

when an attack happens, who is then responsible and how can be accounted for the attack; how do you provide proof in a jurisdictional way?

3 Research methodology

Now that it is clear what different actors could do about the problem and what the incentives are to actually perform these actions, the differences in performance should be analysed. First, the way in which the AmpPot dataset is prepared will be explained, which is needed to perform useful statistical analyses. Then the specific metric that measures how well victims perform will be discussed, which is important to compare the providers in a useful way. Thirdly, several potential explaining factors will be discussed that might influence the difference in performance. In the next chapter, using this improved dataset, the statistical tests that will point out whether the found possible underlying factors have a significant influence on the way victims perform, will be explained, performed and concluded upon.

3.1 Difference in performance

Before finding out why some victims perform better than others, it is important to mention how this performance is indicated. Figure 4 shows the metric that is used to compare the victims to each other. Assuming that every attack in the AmpPot dataset was a successful one, except for the first attack, the victims in the dataset can be compared considering the number of successful attacks they received. The assumption that first attacks are excluded from the successful attacks comes from the finding that many victims only appear once in the dataset and it is assumed that these attacks were only trials by attackers to find out whether the victim is vulnerable or not.

To take into account that some organizations are bigger and therefore have more targets that could be attacks, the number of attacks that occurred for an organization has been divided by the number of IP domains within the organization. This creates a metric that shows how many attacks occurred per IP domain within an organization.

Important to note is that this performance indicator does not directly show how well the organizations protect themselves. An organization could also be attacked often because of their market position or their large set of enemies. Examples of factors that actually explain why certain organizations receive more attacks than others, will be discussed in the next section.

3.2 Explaining factors

Now that it is clear that significant differences exist between the performance of different victims in ensuring availability, it is important to focus on underlying reasons for these differences. In the AmpPot dataset, several possible influencing factors can already be seen. The **size of a company** might influence their likelihood of getting attacked, as putting down bigger companies maximizes the impact of an attack. Also, when a company grows to a certain size, it might be that certain security checks fall, due to heavily increasing costs.

Another factor that might influence this likelihood, is the **type of company**. It might be that educational companies are attacked more often, because students gain from having a reason to delay their projects or to be unable to take part in an exam.

A third possible explanatory factor is the **country** in which a company is **located**. It might be that companies in the United States receive more attacks, because it might be easier to attack companies there due to lower security necessities. This could be similar to the EMV system (Sullivan, 2013),

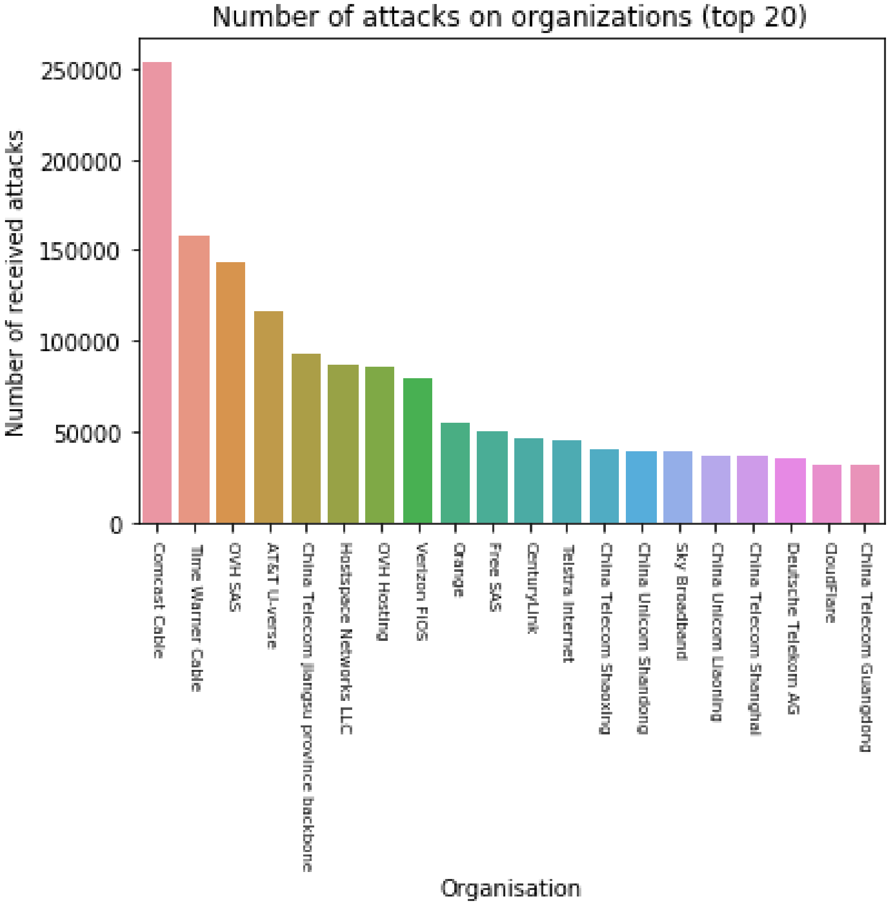

Fig. 4: Number of attacks per organization (Top 20)

where one non-participating country can be a way in. Also, cultural differences might introduce different attitudes towards security, as companies are profit-led in certain countries, such as the United States.

Another factor that might influence the number of DDoS attacks a company receives, is the number of attacks that have already occurred at this company in the past. A known vulnerable company, might attract attackers as the chance of success is relatively big. This is similar to burglaries happening twice or more at the same address. The burglars know after the first time that the house is vulnerable and attack again. Moreover, people that had their inventory stolen likely bought new replacements which can be stolen again. In DDoS attacks, this could be different as a first attack might have induced increased security. So the relation between the past number of attacks and the present number of attacks could be both negative and positive. We will refer to this variable further as **network hygiene**.

3.3 Data Preparation

The AmpPot dataset contains victim data gathered via a set of amplifier honeypots called AmpPot (Krämer et al., 2015). AmpPots mimic services having amplification attack vectors and respond to queries as if they were vulnerable, and therefore are able to incur DDoS attacks and record those events; however, their responses are filtered in order to prevent from contributing to actual attacks (Krämer et al., 2015; Noroozian et al., 2016). In total, eight AmpPots were deployed in the period of 2014 - 2015. Table 1 shows a subset of the dataset as an example.

Table 1: Sampled examples of AmpPot dataset

target_id	date	duration	service	caida_type	org	org _ipsize _seen	raw_country	paket
192.30.252.109	2015-08-25	2478	dns	Content	GitHub	87	United States	491
66.161.48.85	2015-01-13	782	ssdp	Transit/ Access	AT&T Internet Services	88377	United States	143
101.200.83.131	2015-12-17	30	ntp	Content	Aliyun Computing Co., LTD	134554	China	1582
54.250.28.38	2014-08-21	911	dns	Enterpise	Amazon.com	844486	Japan	1562
147.67.119.2	2014-12-30	85	ntp	Transit/ Access	European Commission	175	Luxembourg	480

To make sure the data is ready to provide statistical answers to the questions whether the given underlying factors are actually explaining variance or not, some preparation should take place.

First of all, as the focus is only on all possible victims, all organisations will be taken into account. Moreover, only victims that received more than 1 attack in 2014 and 2015 will be taken into account, as we assume that this first attack was only a test by attackers to find out the vulnerability of the victims. Then, NaN variables were removed from the dataset, as they could interfere with the statistical tests that had to be performed. Each attack that had a missing value in the columns country, type of organisation, name of organisation, total number of attacks, year of attack or IP domain size, was removed. This left around 8000 organisations. To find out the total number of attacks that a certain victim received, the data was grouped by organisation, and the number of rows were summed. This was also done for attacks in 2014 and attacks in 2015, to create 3 new columns of total attacks (in that year). Then, the latest occurrence, was kept in the data to keep the most accurate data about the organization. This removes some value from statistical testing as the organization might have changed country or size in the meantime.

Now that every organization occurs only once in the dataset, with the total number of attacks in 2014, 2015 and both together, the statistical testing can take place.

4 Statistical analysis

In this section, statistical tests are performed aiming to examine the aforementioned possible underlying factors. Spearman's rank correlation test is performed to test the factors on the significance of their influences on the occurrence of attacks an organisation (or victim) receives. The reason why Spearman was chosen over Pearson is because the data is not normally distributed. The normality test is based on the technique from D'AGOSTINO and Pearson (1973) which shows that the sample of number of attacks does not look Gaussian (p=0.000). The data also does not look normal with the Anderson-Darling test (Stephens, 1974). For the Spearman test it is important that there are monotonic relationships, but not necessarily linear.

As two of the possible factors are only on a nominal level, chi squared tests are performed to find out whether these variables have influence on the number of attacks.The possible factors that are considered in this study include the organisation's IP size (interval), the country where it seems to reside (nominal), its type organisation (nominal) and its "network hygiene" (interval). For Chi-Square it is important that at least one of the two variables should be either ordinal or nominal data. Furthermore, they need to be categorical independent groups.

1. *IP size* — The organisation's IP size presents the "size effect" of the analysis. The distributions of this factor versus the occurrence of attacks are shown in Fig. 5.
2. *Country where the victim resides* — For this factor, the country's ICT index and population are selected as indicators to test its influence. Next to these exploratory analyses, a chi-square test has been performed to find significance difference between countries.
3. *Type* — This factor tests whether an ISP, a hosting provider or other organisations would attract more DDoS attacks. Again, a chi-square test will shows if there exists a significant difference.
4. *Network hygiene* — This factor is measured through the occurrence of attacks the organisation received in the previous year. Spearman correlation will show if there is a significance correlation between the number of attacks in 2014 and 2015.

In the next paragraphs, it will be tested whether the factors actually explain the differences in performance significantly or not. For each of the indicated possible factors, the appropriate statistical test will be stated, applied and concluded on. For each of the tests, the number of attacks per victim will be used as dependent variable. In contrast to other assignments, this metric will not take into account that some organisations are bigger, and therefore have a bigger chance of receiving an attack towards one of their IP addresses, as this size indicator will be used as one of the explaining factors for the number of attacks.

4.1 IP size

To find out if the size of a victim matters, the IP size for each of the victim will be used as independent variable to test its correlation with the occurrence of attacks a victim receives. Figure 6 shows the relationship between the two variables. Spearman's test suggests that the variables are correlated with the correlation coefficient of $\rho = 0.402$ within 95% confidence interval. This can mean that a larger organization receives more attacks than a smaller one. This makes sense because large organizations lose the maximum value when their services are down, which makes them a likely victim. Also, the larger the IP size is of an organization, the more possible targets they have. Another theory is that the attackers might target large organizations because it will increase the reputation of the attackers. They will be able to garner more prestige and sell their services in the underground DDoS world.

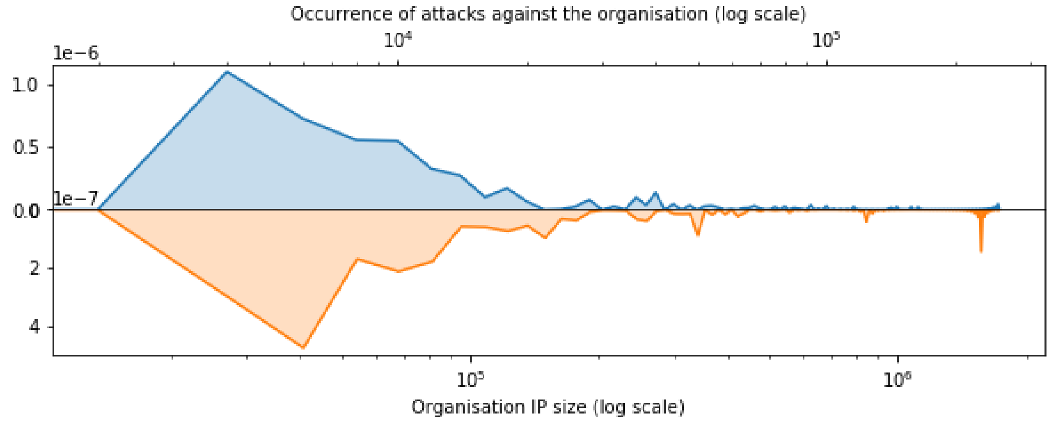

Fig. 5: Distribution (as kernel density) of organisation's IP size and the occurrence of attacks against whom, both in log scale.

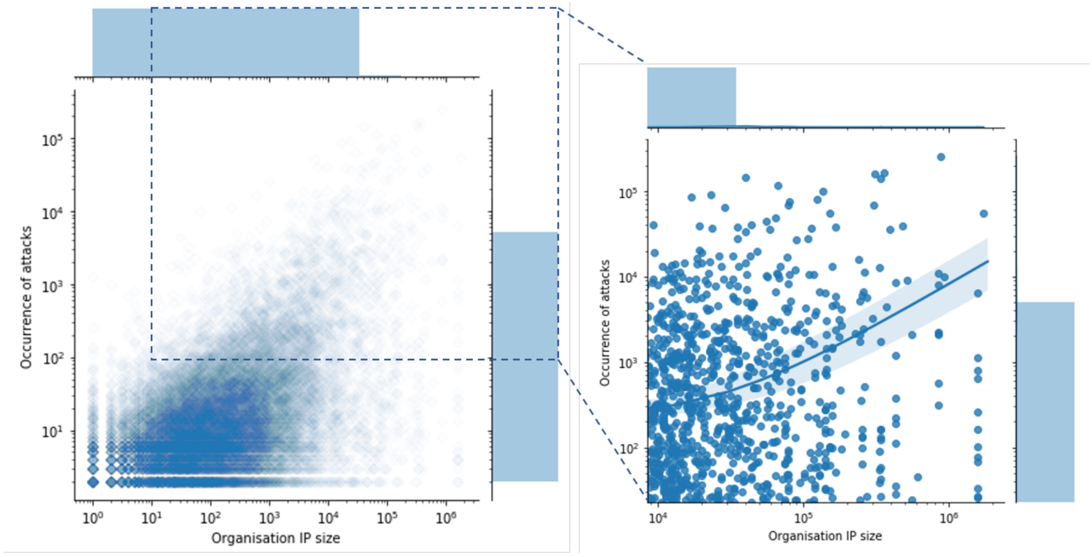

Fig. 6: Occurrence of attacks against organisation (victim) to the IP size of organisation, both in log scale.

4.2 Country of victim

The country's ICT index and population are selected as indicators to test whether the country where the victim resides has an influence on the occurrence of attacks the victim receives. Figure 7 illustrates the relationship, and Spearman's test is performed to test the correlation for the two indicators respectively. The ICT index results in a p-value of $p = 0.614$ which suggests the indicator is uncorrelated with the occurrence of attacks, while the country's population is negatively-correlated with the victim's occurrence of attacks (p = 0.00) with the correlation coefficient of $\rho = -0.020$ within 95% confidence interval. Next to this indicators of country, also the distribution of attacks over the countries themselves has been tested using a chi-square test. This test returned a p value of 0.00, which means there is a significant difference between countries. It is impossible to conclude where these differences exist based on the chi-squared test, as this only shows difference between all groups. However, the Spearman test shows correlation between the population and the number of attacks, indicating that lower populated countries receive more attacks. As no logical conclusion can be found for this relationship, we can conclude that this correlation is dependent on other charateristics of the different countries, such as the number of organizations within that country, or how strict the law is in the countries.

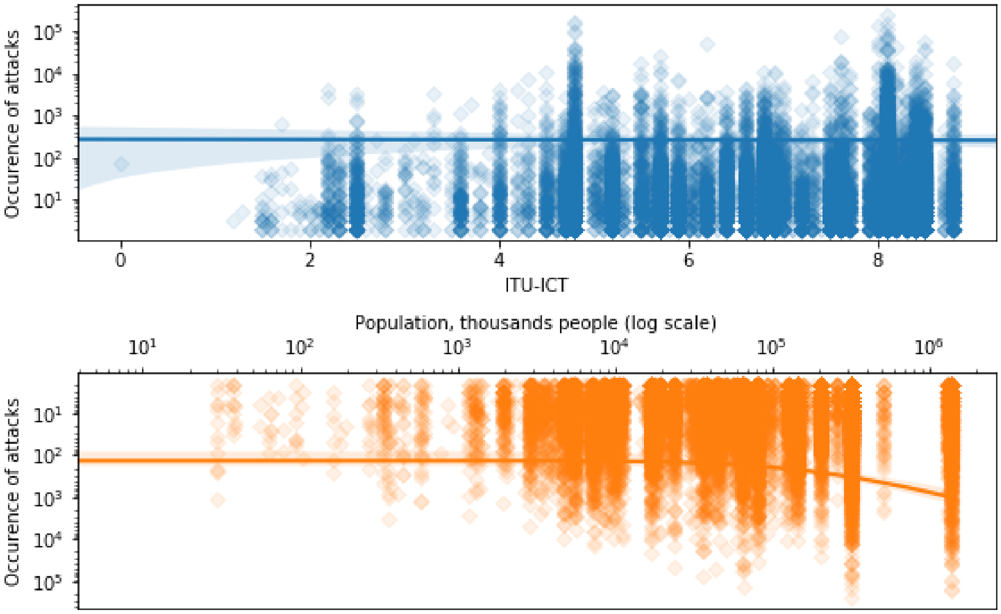

Fig. 7: Occurrence of attacks against organisation (victim) per country where they reside

4.3 Type of victim

The type of victim is labelled by two classification methods: one presented by Noroozian et al. (2016) (indicated as "AS type") and the CAIDA classified labels. Fig. 8 shows the distribution of the number of attacks each type of victim receives under different classified labels. As with the country, the type of victim is only a nominal variable. Therefore, a chi-squared test has to be performed in order to find out if the independent variable "type" has a significant influence on the number of attack a company receives. The test returned a p value of 0.00, meaning that there is a significant difference between different types of organizations. The test does not point out where the exact differences are. Although fig. 8 suggests that the most attacks take place at hosting providers, this cannot be concluded. The number of hosting providers in the dataset can distort the relationship.

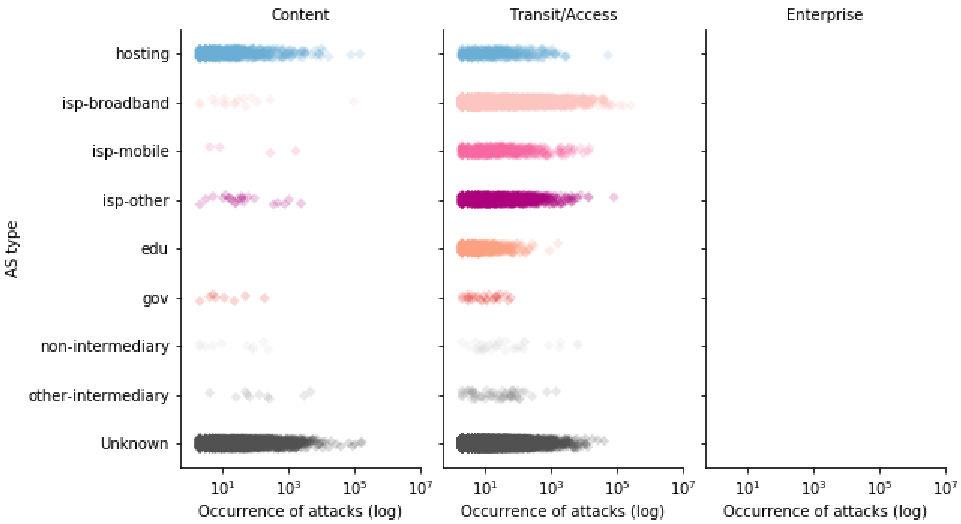

Fig. 8: Occurrence of attacks against victims by victim type: CAIDA classified labels vs AS type

4.4 Previous attacks on victims

Compared to the other underlying factors, the number of attacks that have already occurred within a certain victim is harder to connect to the number of attacks that a company receives. As the independent variable is part of the dependent variable, finding a correct statistical test is difficult. To solve this, the variable has been split up into 2 separate variables, based on the date an attack occurred. This reveals two datasets, where one is the attacks that happened before a certain date, and the other shows attacks that happened after this date. This can show whether previous attacks influence the number of additional attacks. A shortcoming of this method, is that for certain companies, the first attack happened relatively late within the scope of the dataset. This can distort the relation. Because the separated data now exists of two paired datasets, Spearman's test can be used to find out if the independent variable has a significant influence on the dependent variable. Fig. 9 shows the regression line between the number of attacks in 2014 and the number of attacks of 2015. Spearman's test suggests a correlation between the occurrence of attacks of previous year and this year with correlation coefficient of $\rho = 0.265$ and p-value $p = 0$ within 95% confidence interval. Although significant correlation exists, which suggests that organizations that received many attacks in 2014 will also receive many attacks in 2015, this does not indicate a causal relationship between the two variables. As mentioned before, the data has been distorted and some organizations only started receiving attacks in 2015, which distorts the relationship. However, that the test still suggests correlation, even though the data has been distorted, it can be stated that a relationship does exist. It also makes sense that organization that are known to be weak receive more attacks, as attackers know for sure that their attacks will be successful.

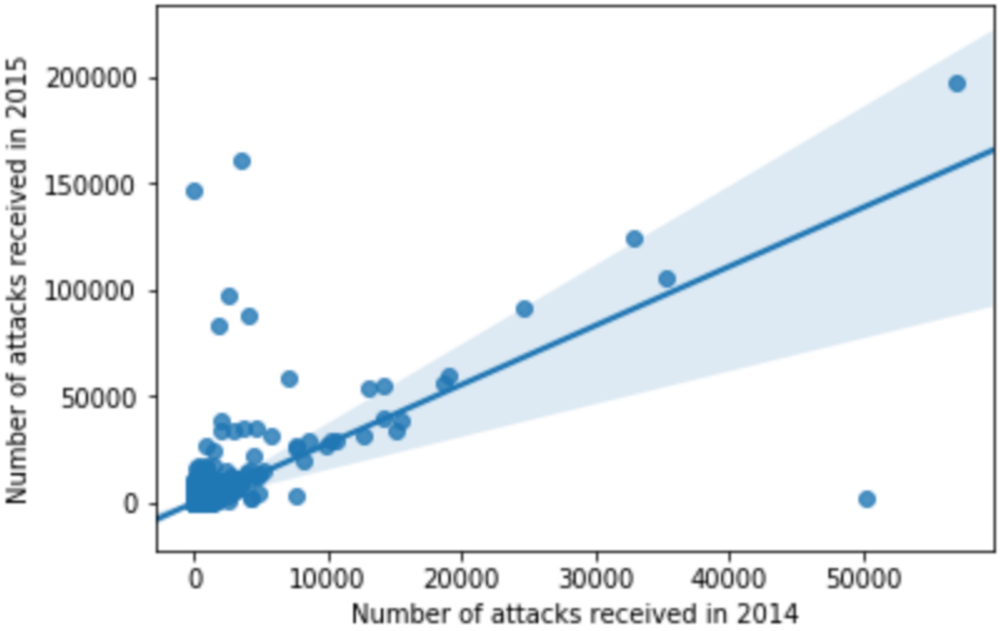

Fig. 9: Number of DDoS attacks on organisations in 2014 compared to 2015.

5 Conclusion

In this paper, the performance of different victims in protecting their availability from DDoS attacks is compared. The underlying factors that explain the different performances by victims, categorized in three actor types, are analysed.

The first actor, hosting providers, will try to mitigate possible damages by applying Deep Packet Inspection (DPI). Using this technique can be very beneficial for finding malicious behavior, but comes at the cost of privacy of its users.A problem with the countermeasures of hosting providers, is that no one wants to pay the price of tracking down attackers as it has no financial benefits for the hosting providers once they are sought after, however it does come with high costs.

The second actor, internet service providers (ISP), have a wide range of countermeasures at their disposal such as blackholing, scrubbing, traffic engineering, and local filtering. This paper has further investigated RFC 2827 (ingress filtering). This countermeasure prevents attackers from spoofing their IP-addresses. Doing so, it will eliminate the amplification/reflection attacks which are used in the highest volume DDoS attacks. Although effective at blocking a certain type of DDoS attacks, it does require additional hardware and thus extra costs for the ISP to implement.

The third actor, governmental institutions, are defined by both the EU and the national government. Governmental institutions can mitigate damages by making laws and regulations. The countermeasure discussed is the transfer of responsibility of attacks from hosting providers towards technology vendors. This way, technology vendors are incentivised to provide more robust and safe infrastructure, resulting in less attacks.

By applying the data from the AmpPot dataset to a series of statistical tests, the data can be tested on differences between actors in their performance. Doing so, an effort is made to find factors that result in good (or bad) performance.
By looking at the size, the type, the geospatial location, and the network hygiene (if networks are successfully attacked multiple years in a row) of companies, correlations may be found.
Resulting from the statistical tests on these performance indicators, the following results can be presented:
The size of companies correlates with the amount of attacks it suffers. Two reasons for this correlation can be proposed; first, the size of a company makes for more points of attack. If a company houses more data, it is more prone to attacks for it. Secondly, attacking bigger companies is more interesting for attackers since the possible damages are greater when a website goes down. Thirdly, the reputation of an attacker successfully attacking bigger companies is greatly lifted and thus this provides more incentives for attackers to aim for bigger companies.
The geospatial location of organizations shows correlation to attacks. A proposed reason for this correlation is that laws differ per country, which attackers use to their advantage. Attacks are more prone on companies located in countries with less strict regulations regarding security, these companies will have less secure server and will be attacked more often for it.
The type of organisation is also correlated to the amount of attacks. This correlation can be explained by looking at the incentives of attackers. Many attackers using DDoS have a very specific (emotional, personal) reason for attacking. Gamers have been using DDoS to attack opponents in a game, which causes harm to organisations in the game sector specifically.
The last variable, network hygiene, also shows correlation to the number of attacks. The main proposed reason for this effect is that attackers successfully attacking a company are more incentivised to return later to attack again. If an attack fails, this demotivates attackers to attack, hence the correlation with network hygiene.
The results show many correlations in the variables that were tested in the statistical analysis. These insights require further analysis as to why they exist. Further exploratory research should be conducted to find out why the correlations exist to further combat DDoS attacks.

6 Limitations

Although this paper has shown significant correlations between on the one hand the IP size, the network hygiene, the country and the type of an organization and on the other hand, the number of attacks that an organization receives, these findings have many limitations. First of all, it is important that correlations is different from causation. Correlation does not per definition mean that there exists a relationship between two variables.

Furthermore, the tests we performed were looking at the underlying factor individually. In an optimal situation, each of these variables would have been analysed as part of a whole set of underlying factors. For example, a multiple regression analysis could have shown which part of the difference in performance by the victims could be explained by which variable. This could show that the size of a company explains 40 percent of the variance, the other 3 factors explain 10 percent per factor, and 30 percent can not be explained by the data. This type of analysis would give a better insight in the underlying factors and their role, but was not performed in this report, due to time constraints and the two nominal variables in the explaining factors set. As the analyses are individual, they can be seen as relatively superficial. We did find a difference in performance per country, but were not able to conclude where these differences exist, and why they exist. The same goes for the type of organization. This is something that should definitely be looked into in further research.

This analysis has only shown four possible factors which might have influence on the occurrence of an attack. However, it is likely that there are more hidden factors which influence the attacks. For example, motivation of attackers is a factor which has not been taken into account because of the lack of data. Other hidden factors could include efficacy of law enforcement, technical trends contributing to IoT attacks (such as the growth of IoT), and types of victims could be more specific.

References

Asghari, H., van Eeten, M., Bauer, J. M., & Mueller, M. (2013, September). *Deep Packet Inspection: Effects of Regulation on Its Deployment by Internet Providers* (SSRN Scholarly Paper No. ID 2242463). Rochester, NY: Social Science Research Network. Retrieved 2019-10-11, from `https://papers.ssrn.com/abstract=2242463`

Assaf, D. (2007). Government intervention in information infrastructure protection. In *International conference on critical infrastructure protection* (pp. 29–39).

Bauer, J. M., & Van Eeten, M. J. (2009). Cybersecurity: Stakeholder incentives, externalities, and policy options. *Telecommunications Policy*, *33*(10-11), 706–719.

Bendrath, R., & Mueller, M. (2011). The end of the net as we know it? deep packet inspection and internet governance. *New Media & Society*, *13*(7), 1142–1160.

Brook, C. (2018, March). *What is Deep Packet Inspection? How It Works, Use Cases for DPI, and More* [Text]. Retrieved 2019-10-11, from `https://digitalguardian.com/blog/what-deep-packet-inspection-how-it-works-use-cases-dpi-and-more`

Cardoso de Santanna, J. (2017). *Ddos-as-a-service: Investigating booter websites* (Doctoral dissertation, University of Twente, Netherlands). (CTIT Ph.D. thesis Series No. 17-448, ISSN 1381-3617) doi: https://doi.org/10.3990/1.9789036544290

Cheung, R. (2017). *Targeting financial organisations with ddos: a multi-sided perspective* (master's thesis). Delft University of Technology, the Netherlands.

Chromik, J., Cardoso de Santanna, J., Sperotto, A., & Pras, A. (2015, 5). Booter websites characterization: Towards a list of threats. In *Proceedings of 33rd brazilian symposium on computer networks and distributed systems, sbrc 2015* (pp. 445–458). Brazilian Computer Society (SBC). (eemcs-eprint-26162)

D'AGOSTINO, R., & Pearson, E. S. (1973). Tests for departure from normality. empirical results for the distributions of b 2 and b. *Biometrika*, *60*(3), 613–622.

Damas, J., & Karrenberg, D. (2008). *Network Hygiene Pays Off - The Business Case for IP Source Address Verification.* Retrieved 2019-10-14, from `https://www.ripe.net/publications/docs/ripe-432`

De Cuyper, R., & Weijters, G. (n.d.). Cybercrime in cijfers. *Unknown*.

Dharmapurikar, S., Krishnamurthy, P., Sproull, T., & Lockwood, J. (2003). Deep packet inspection using parallel bloom filters. In *11th symposium on high performance interconnects, 2003. proceedings.* (pp. 44–51).

Ferguson, P. (2000). *BCP 38 - Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing.* Retrieved 2019-10-02, from `https://tools.ietf.org/html/bcp38`

Gordon, L. A., Loeb, M. P., Lucyshyn, W., & Zhou, L. (2014). Externalities and the magnitude of cyber security underinvestment by private sector firms: a modification of the gordon-loeb model. *Journal of Information Security*, *6*(01), 24.

Holl, P. (2015, March). Exploring ddos defense mechanisms. In *Network architectures and services.*

Krämer, L., Krupp, J., Makita, D., Nishizoe, T., Koide, T., Yoshioka, K., & Rossow, C. (2015). Amppot: Monitoring and defending against amplification ddos attacks. In *International symposium on recent advances in intrusion detection* (pp. 615–636).

McConachie, A. (2014, July). *Anti-Spoofing, BCP 38, and the Tragedy of the Commons.* Retrieved 2019-10-14, from `https://www.internetsociety.org/blog/2014/07/anti-spoofing-bcp-38-and-the-tragedy-of-the-commons/`

Noroozian, A., Korczyński, M., Hernandez Gañán, C., Makita, D., Yoshioka, K., & van Eeten, M. (2016). Who gets the boot? analyzing victimization by ddos-as-a-service. In *Proceedings of the international symposium on research in attacks, intrusions, and defenses, raid 2016.* Springer.

Stephens, M. A. (1974). Edf statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, *69*(347), 730–737.

Sullivan, R. J. (2013). The us adoption of computer-chip payment cards: implications for payment fraud. *Economic Review-Federal Reserve Bank of Kansas City*, 59.

Van Eeten, M., & Bauer, J. M. (2009). Emerging threats to internet security: Incentives, externalities and policy implications. *Journal of Contingencies and Crisis Management*, *17*(4), 221–232.

Van Eeten, M., Bauer, J. M., Asghari, H., Tabatabaie, S., & Rand, D. (2010). The role of internet service providers in botnet mitigation an empirical analysis based on spam data..

Vink, T. (2016). *How can ISPs handle DDoS attacks?* Retrieved 2019-10-02, from `https://security.stackexchange.com/questions/134767/how-can-isps-handle-ddos-attacks`