Evaluating the Impact of AbuseHUB on Botnet Mitigation
Interim Deliverable 1.0

PUBLIC VERSION

Giovane C. M. Moura, Qasim Lone,
Hadi Asghari, and Michel J.G. van Eeten
Economics of CyberSecurity Group

Faculty of Technology, Policy, and Management
Delft University of Technology
M.J.G.vanEeten@tudelft.nl
http://www.tbm.tudelft.nl/econsec

March 24, 2015

Contents

List of Figures

List of Tables

Introduction

This documents presents the Interim Deliverable (1.0) of a study, commissioned by the Ministry of Economic Affairs, to evaluate the impact of AbuseHUB. In essence, AbuseUB is a clearinghouse for acquiring and processing abuse data on infected machines [1]. It is the outcome of a public-private initiative and has a varied and evolving membership. Currently, it consists of 9 Dutch Internet Service Providers (ISPs), SIDN (the .nl registry) and Surfnet (the national research and education network operator). A key objective of AbuseHUB is to improve botnet mitigation by its members.

We set out to assess whether this objective is being reached by analyzing malware infection levels in the networks of AbuseHUB members and comparing them to those of other ISPs, within the Netherlands and in other countries. Since AbuseHUB members together comprise most of the broadband market in the Netherlands, it also makes sense to compare how country as a whole has performed compared to other countries.

This Interim complements the baseline measurement report produced in December 2013. Differently from the baseline report, and from our 2011 study into botnet mitigation in the Netherlands [2], this Interim deliverable contains new data sources we have obtained over the last year.

The document is organized as follows: in Chapter 1, we present the methodology used in this research – ISPs, datasets, and mapping between ISPs and datasets. Then, in Chapter 2, we compare bot presence in Dutch networks against several other countries. After that, in Chapter 3, we focus solely on the infections locations in the networks of ISPs, and compare the performance of Dutch ISPs against other countries. In the same chapter, we explore the impact of having national anti-botnet initiatives in the infection results.Finally, in Chapter 5, we present the summary of this report.

Chapter 1

Questions and methodology

The basic methodology employed in this report consists of collecting and analyzing Internet measurement data on infected machines. We then interpret these measurements by connecting them to other variables related to the operators of the networks containing infected machines, such as the country in which it is located and the number of subscribers. This way we can develop comparative metrics to determine the performance of Dutch ISPs, members and non-members of AbuseHUB, to each other and to ISPs in other countries with regards to botnet mitigation.

For this interim measurement, we evaluate the infection rates of Dutch ISPs over different times frames, depending on the data source. Three sources cover the period from January 2011 to December 2014. Other sources cover only parts of 2014. This interval complements the previous our report on Dutch ISPs [2] as well as the previous baseline measurement report we have produced to AbuseHUB, in which we covered data up to 2013.

One issue we could not fully resolve at this stage is the fact that the members of AbuseHUB are rather heterogeneous and that we are not yet in a position to generate reliable comparative metrics that take into account the differences in size and nature of the subscriber population when comparing them to non-members inside and outside the Netherlands. We have chosen an approximation approach, that we outline in the next section.

In the remainder of this chapter, we first present the research questions. Next, we turn to the Dutch Internet Service Providers and other network operators that are included in the analysis (Section 1.2). Next, we present in Section 1.3 the infection-related datasets we have analyzed. Last, in Section 1.4, we explain how the datasets are mapped back to the ISPs covered in Section 1.2.

1.1 Research questions

The main goal of the overall study, to be delivered in December 2015, is to assess the impact of the AbuseHUB initiative on infection rates in the networks of Dutch ISPs. To achieve this goal, four questions we articulated:

1. How do member ISPs compare to non-member ISPs?
2. How do member ISPs compare among themselves?

3. Do member ISPs have better data on the presence of bots in their networks?

4. What recommendation can be identified to improve botnet mitigation in the Netherlands?

1.2 AbuseHUB members and Dutch Internet Service Providers

In this interim report, we focus on the first two questions: How do member ISPs compare to non-member ISPs? And: How do member ISPs compare among themselves?

Given that AbuseHUB has ISPs as well as non-ISPs among their membership, there are certain limitations in comparing the members among themselves, as well as comparing members with non-members. The number of infections in a network is, to a significant extent, a function of the size of the user population in that network, as well as the type of users. To put it crudely, networks with highly different user populations cannot be meaningfully compared.

For ISPs, we have found that dividing the number of infections by the number of subscribers gives reliable comparative metrics. This works well for a subset of the AbuseHUB members. It does not work for all networks served by AbuseHUB, however. The Autonomous System (AS) of SIDN is basically a type of corporate network and the machines in the network are owned by the organization itself. Surfnet is different from all the others also. It receives abuse data for 16 highly heterogeneous ASes. A similar complication holds for some ISPs, who are also receiving abuse data for other ASes than the one with their broadband customers. All of this means that there is no meaningful way to generate comparative metrics for all 35 ASes covered by AbuseHUB.

We deal with the heterogeneity within AbuseHUB in three ways. First, we generate comparative metrics for those ASes of AbuseHUB that provided broadband access, predominantly to consumers. For those ASes, the number of subscribers that reside in the network is a good basis to take the size of the network into account. The data on the number of subscribers we got from the TeleGeography's GlobalComms database. We disaggregated them further using data supplied to us by the members of AbuseHub. In short, we compare oranges to oranges by including a subset (albeit a large subset) of AbuseHUB Ases. This is continuation of the approach in our earlier reports.

To include the non-broadband access ASes of AbuseHUB in the study, we have generated additional metrics and units of analysis. These can deal with the heterogeneity, but at the expense of being less comparable across networks. There are basically two workarounds. First, we will look at the absolute number of infections in all ASes of AbuseHUB (raw count or indexed) and we will look at the Netherlands as a whole. To see whether AbuseHUB has led to a faster improvement in mitigation compared to ISPs in other countries, we will analyze the data at the country level. Since AbuseHUB covers a large part of the Dutch broadband user population, we will take the Netherlands as a whole as a proxy for AbuseHUB and look at what has happened to infection rates for

ISP	ASes	$\sum$ IPv4[1]
KPN (incl. Telfort)	286, 1134, 1136, 20143, , 8737, 49562	13,278,748
Tele2	15670, 34430, 13127, 20507	1,861,847[2]
Telfort	5615	573440
UPC	6830	4,087,268[3]
RoutIT	28685	235,776
SIDN	1140,48283	3,584
SOLCON	12414	145,408
SURFnet	1101, 1102, 1103, 1104, 1124, 1125, 1126, 1128, 1132, 1133, 1139, 1145, 1161, 1837, 1888, 25182	11,345,152
XS4ALL	3265	1,123,328
Ziggo	9143	3,705,088
ZeelandNet	15542	1,751,104
Total AbuseHUB IPs:		**38,110,743**
Total Dutch IPs [3]:		**49,094,701**
Ratio AbuseHUB/NL		**77.6%**

Table 1.1: Evaluated Internet Service Providers

the country as a whole. At that level, we can generate relative metrics by dividing by the number of Internet users in each country. The second workaround is to used indexed metrics when possible, to see how much the Netherlands has improved over time compared to other countries. Those results are discussed in the next chapter.

Table 1.1 presents the ISPs evaluated in this report. This list was obtained directed with the AbuseHUB. The count of ISP size was performed using BGP data from Dec, 2013. As can be seen, the AbuseHUB covers 77.6% of the Dutch IP address space.

1.3 Evaluated botnet infection datasets

As covered in [2], there is currently no authoritative data source to identify the overall population of infected machines around the world. Commercial security providers typically use proprietary data and shield their measurement methods from public scrutiny. This makes it all but impossible to correctly interpret the figures they report and to assess their validity.

The publicly accessible research in this area relies on two types of data

[1]Obtained from BGP tables on March 14th, 2014, combining all the ASes of each ISP. We use Maxmind's [4] geolocation databses to filter out IPs employed in The Netherlands.

[2]Only IPs geo-located to the Netherlands. Tele2 had 1,878,016 allocated.

[3]Only IPs geo-located to the Netherlands. AS 6830, in fact, had 19,239,936 globally allocated.

sources:

- Data collected external to botnets. This data identifies infected machines by their telltale behavior, such as sending spam or participating in distributed denial of service attacks;
- Data collected internal to botnets. Here, infected machines are identified by intercepting communications within the botnet itself, for example by infiltrating the command and control infrastructure through which the infected machines get their instructions.

Each type of source has its own strengths and weaknesses. The first type typically uses techniques such as honey pots, intrusion detection systems and spam traps. It has the advantage that it is not limited to machines in a single botnet, but can identify machines across a wide range of botnets that all participate in the same behavior, such as the distribution of spam. The drawback is that there are potentially issues with false positives. The second type typically intercepts botnet communications by techniques such as redirecting traffic or infiltrating IRC channel communication. The advantage of this approach is accuracy: bots connecting to the command and control server are really infected with the specific type of malware that underlies that specific botnet. The downside is that measurement only captures infected machines within a single botnet. Given the fact that the number of botnets is estimated to be in the hundreds [5], such data is probably not representative of the overall population of infected machines.

Neither type of data sources sees all infected machines, they only see certain subsets, depending on the specific data source. In general, one could summarize the difference between the first and the second source as a trade-off between representativity versus accuracy. The first type captures a more representative slice of the problem, but will also include false positives. The second type accurately identifies infected machines, but only for a specific botnet, which implies that it cannot paint a representative picture.

Taking these criteria into account, we have obtained the following data sources, in which we group into two categories: global sources and Dutch sources (only NL IP addresses). From these sources, only spam is categorized as "external" to the botnet; all the other ones are internal obtained either via sinkholes or sandboxes.

1.3.1 Global data sources

Spam trap dataset (Spam)

Spam data are obtained from a spamtrap we have access to – the same source as used in the baseline report. It might not be fully representative of overall spamming trends, and also there is no guarantee that the listed spam sources are indeed originating from botnets, though so far that is still the main platform for distribution. The more important limitation is that the spam has become a less important part of the botnet economy, as witnessed in the substantial drop in overall spam level. The reports of security firms seem to confirm these overall trends. Symantec reported a significant decrease in the volume of spam messages, "from highs of 6 trillion messages sent per month to just below 1

trillion" [6] until 2012 (See Figure 4.8). Cisco, TrendMicro and Kaspersky show that the spam volume since that period has been fluctuating, but staying at more or less the same level (see [7] and [8]). All of this means that the source is becoming less representative of overall infection levels.

Shadowserver Sinkhole Conficker data (Conficker)

Established in 2004, the Shadowserver Foundation comprises volunteer security professionals that "gathers intelligence on the darker side of the Internet". They have created the Conficker working group, which provides reports and data on "the widespread infection and propagation of Conficker bots" [9].

Several members of the working group run sinkholes that continuously log the IP addresses of Conficker bots. The sinkholes work in this fashion: computers infected with Conficker frequently attempt to connect to command and control servers to receive new payloads (i.e., instructions). In order to protect the botnet from being shut down, Conficker attempts to connect to different C&C domains every day. The working group has succeeded in registering some of these domain names and logging all connections made to them. Since these domains do not host any content, all these connections are initiated by bots. Therefore, we can reliably identify the IP addresses of the Conficker bots.

The Conficker dataset is unique in several ways. First of all, unlike the other two datasets, it is not a small sample of a much larger population, but rather captures the universe of its kin. This is because of the way the bot works – most of them will eventually contact one of the sinkholes. Second, this dataset is basically free from false positives, as, apart from bots, no other machine contacts the sinkholes. These features make the dataset more reliable than the spam or DShield datasets. The difference, however, is that the dataset is only indicative of the patterns applicable to one specific botnet, namely Conficker. Although Conficker has managed to replicate very successfully, with around several million active bots at any given moment, it has not been used for any large-scale malicious purposes – or at least no such uses have been detected yet. This means ISPs and other market players may have less powerful incentives to mitigate these infections, different from spam bots, for example. These differences make the Conficker dataset complementary to the two other sets.

Overall, the Conficker dataset adds a fresh, robust and complimentary perspective to our other two datasets and brings more insight into the population of infected machines worldwide.

Zeus Gameover Botnet (Peer and Proxy)

Zeus botnet started making headlines in 2007, as a credential stealing botnet. The first version of Zeus was based on centralized command and control (C&C) servers. The botnet was studied by various security researchers and multiple versions were also tracked [10, 11, 12, 13].

In recent years Zeus has transformed, into more robust and fault tolerant peer-to-peer (P2P) botnet, known as P2P Zues or Gameover. The botnet supports several features including RC4 encryption, multiple peers to communicate stolen information, anti-poising and auto blacklist. It also can be divided into *sub-botnet*, based on BotIDs , where each sub-botnet can be used to carry-out diverse task controlled by different botmasters.

The botnet is divided into three sub-layers, which provide following functionality.

- **Zeus P2P Layer (Peer):** This is the bottom most layer and contains information of infected machines. Bots in P2P layer exchange peer list with each other in order to maintain updated information about compromised machines.
- **Zeus Proxy Layer (Proxy) :** A subset of bots from P2P layer are assigned the status of proxy bots. This is done manually by the botmaster by sending proxy announcement message. Proxy bots are used by Peer-to-peer layer bots to fetch new commands and drop stolen information.
- **Domain Generation Algorithm Layer:** DGA layer provides fall backup mechanism, if a bot cannot reach any of its peers, or the bot cannot fetch updates for a week. Zeus algorithm generates 1000 unique domain names per week. Bots which lose connection with all connected peers search trough these domains until they connect to live domain.

  More details about architecture and functioning of the botnet can be found in literature [14, 15].

This dataset is sub-divided into three feeds, GameOver Peer, GameOver Proxy and GameOver DGA. The botnet is spread in around 212 countries with on average 95K unique IP addresses per day. Hence it is gives us insight of botnet infection level at global level, and compare various countries and ISPs.

ZeroAccess

ZeroAccess is a Trojan horse, which uses a rootkit to hide itself on Microsoft Windows Operating Systems. The botnet is used to download more malware and open backdoor for botmaster to carry out various attacks including click fraud and bitcoin mining.

The botnet is propagated and updated through various channels including compromised website redirecting traffic and dropping rootkit at potential host or updating the already compromised host through P2P network.

ZeroAccess also provide global view with bots in around 220 countries with an average of about 12K unique IP addresses per day.

Morto

Morto is a worm that exploits the Remote Desktop Protocol (RDP) on Windows machines to compromise its victims. It uses a dictionary attack for passwords to connect as Windows Administrator over RDP with vulnerable machines in the network. After successfully finding a vulnerable machine, it executes a dropper and installs the payload.

We have a time series data of Morto for past 4 years with an average of 5k daily unique IP addresses distributed globally. This is relatively small, but it complements our other data sources by providing a longitudinal perspective.

1.3.2 Netherlands-only data sources

In addition to the global feeds, we have obtained access from Shadowsever Foundation to botnet data pertaining only to the Netherlands.

Shadowserver's bot feed

Shadowserver collects list of infected machines by monitoring IRC Command and Controls, IP connections to HTTP botnets, or IP's of Spam relay [16]. This Report contains comprehensive list of IP addresses in The Netherlands of compromised machines which are infected with different malware or botnets.

This datasource enable us to compare ISPs within The Netherlands and also AbuseHUB members with non-members.

Shadowserver's Microsoft Sinkhole

Microsoft shares with Shadowserver Foundation data from botnet sinkholes[4]. We have also obtained this data for IP addresses located in the Netherlands.

1.4 Mapping offending IP addresses to Dutch ISPs

For each unique IP address that was logged in one of our data sources, we looked up the Autonomous System Number (ASN) and the country where it was located. The ASN is relevant, because it allows us to identify what entity connects the IP address to the wider Internet – and whether that entity is an ISP or not.

However, there are some ISPs in Table 1.1 that operate in various countries across Europe. We employ IP-geolocation databases [17] from Maxmind [4] to single out IP addresses used in The Netherlands from the other European countries when classifying the attacking IP addresses from each ISPs.

As both ASN and geoIP information change over time, we used historical records to establish the origin for the specific moment in time when an IP address was logged in one of our data sources (e.g., the moment when a spam message was received or network attack was detected). This effort resulted in time series for all the variables in the datasets, both at an ASN level and at a country level. The different variables are useful to balance some of the shortcomings of each – a point to which we will return in a moment.

We then set out to identify which of the ASNs from which botnet IP data belonged to ISPs. To the best of our knowledge, there is no existing database that maps ASNs onto ISPs. This is not surprising. Estimates of the number of ISPs vary from around 4,000 – based on the number of ASNs that provide transit services – to as many as 100,000 companies that self-identify as ISPs – many of whom are virtual ISPs or resellers of other ISPs' capacity.

So we adopted a variety of strategies to connect ASNs to ISPs. First, we used historical market data on ISPs – wireline, wireless and broadband – from TeleGeography's GlobalComms database 2013 [18] . We extracted the data on all ISPs in the database listed as operating in a set of 40 countries, namely

[4] `https://www.shadowserver.org/wiki/pmwiki.php/Services/Microsoft-Sinkhole`

all 34 members of the Organization for Economic Co-operation and Development (OECD), plus one "accession candidate" and five so-called "enhanced-engagement" countries.

The process of mapping ASNs to ISPs was done manually. First, using the GeoIP data, we could identify which ASNs were located in each of the 40 countries. ASNs with one percent of their IP addresses mapped to one of the 40 countries were included in our analysis. For each of these countries, we listed all ASNs that were above a threshold of 0.5 percent of total spam volume for that country.

We used historical WHOIS records to lookup the name of the entity that administers each ASN in a country. We then consulted a variety of sources – such as industry reports, market analyses and news media – to see which, if any, of the ISPs in the country it matches. In many cases, the mapping was straightforward. In other cases, additional information was needed – for example, in case of ASNs named after an ISP that had since been acquired by another ISP. In those cases, we mapped the ASN to its current parent company.

It is important to notice that ISP change their AS size over time (mergers, selling/buying blocks). To cope with this, we use historical BGP data and produce our metrics matching the timestamp of the botnet data with the BGP tables of the respective period.

1.5 Compensating for Known Limitations in Internet Measurements

Our approach allows us to robustly estimate the relative degree in which ISP networks harbor infected machines. It has certain limitations, however, that need to be compensated for. The effects of three technical issues need to be taken into account when interpreting the data: the use of Network Address Translation (NAT), the use of dynamic IP addresses with short lease times. The key issue is to understand how these technical practices affect the number of machines that are represented by a single unique IP address.

NAT means sharing a single IP address among a number of machines. Home broadband routers often use NAT, as do certain other networks. This potentially underrepresents the number of infected machines, as multiple machines show up as a single address. Dynamic IP addresses with short lease times imply that a single machine will be assigned multiple IP addresses over time. This means a single infected machine can show up under multiple IP addresses. As such, it over-represents the number of infected machines. Both of these practices counteract each other, to some extent. This limits the bias each of them introduces in the data, but this does not happen in a consistent way across different networks.

This is a classic problem in the field of Internet measurement: how many machines are represented by a single IP address? Ideally, one IP address would indicate one machine. But reality is more complicated. Over an extended time period, a single address sometimes indicates less than one machine, sometimes more than one. This varies across ISPs and countries. Earlier research by Stone-Gross *et al.* [19] has demonstrated that in different countries, there are different ratios of unique IP addresses to infected machines – referred to as "DHCP

churn rates" .

In this report, to control for the bias caused by churn rates, we use shorter time scales when counting the number of unique IP addresses in a network, for all the datasets.

On shorter time scales, the potential impact of churn is very limited. Earlier research found that churn starts to affect the accuracy of IP addresses as a proxy for machines on timescales longer than 24 hours. 12 We therefore worked with a time period of 24 hours. All our comparative analyses are based on the daily average number of IP addresses from an ISP network. This compensates for churn, but has a downside: in these estimates, the number of infected machines may be now grossly undercounted, depending on the prefix, AS and/or ISP evaluated.

While the number of bots measured in a 24 hour period is the most reliable for comparisons across networks, it cannot indicate the actual infection rate of a network in absolute terms. For absolute estimates – in other words, of the actual number of infected machines – we use larger time periods, depending on the situation: months, quarters or even the whole 18-month measurement period. For sources we have checked both the daily average number of unique IP addresses, and the total number of unique IP addresses for that particular metric. That way, we can compensate for the various measurement issues. Patterns that hold across these different measurements can be said to be robust and valid. These measurement issues are revisited in more detail in each section of the findings chapter.

The result of this approach is time series data on the number and the location of infected machines across countries and ISPs. We have paid special attention to whether these machines are located in the networks of the main ISPs in the wider OECD.

1.5.1 Actively Measuring DHCP churn

Currently, our team at TU Delft is working on producing an active-measurement based approach that are able to capture, for each ISP/AS and network block, their respective DHCP churn rates.

We present in [20] the first attempt towards estimating ISP and Internet-wide DHCP churn rates, in order to better understand the relation between IP addresses and hosts, as well as allow us to correct data relying on IP addresses as a surrogate metric. We proposed an scalable active measurement methodology and then validate it using ground truth data from a medium-sized ISP. Next, we built a statistical model to estimate DHCP churn rates and validate against the ground truth data of the same ISP, estimating correctly 72.3% of DHCP churn rates. Finally, we apply our measurement methodology to four major ISPs, triangulate the results to another Internet census, and discuss the next steps to more precisely estimate DHCP churn rates.

The next steps in this research is to measure large scale ISPs and their dynamics over time. We hope to be able to produce DHCP-churn normalized metrics for the final version of this report.

Chapter 2

The Netherlands compared against other countries

In this chapter, we use our global data sources (Section 1.3) to rank the Netherlands against all other countries, and, more specifically, against a selected set of countries which we considered relevant reference points (Germany, Great Britain, France, Finland, Italy, Spain, United States and Japan).

The aim of this chapter is two answer the following questions:

- How does the Netherlands rank against other countries?
- Is the Netherlands improving over time, compared to other countries?

The idea is that if the ranking of the Netherlands has improved, this suggests a positive impact of AbuseHUB (see Section 1.1 for details on why we use country-level comparisons to study the impact).

To answer the questions, we compute the daily number of unique IP addresses from each global data feeds we have obtained, and aggregate it into quarters or weeks, depending on the time span of the data.

In Section 2.1, we analyze the average performance of the Netherlands, compared to other countries, also taking into account the number of Internet users. In Section 2.2, we present an analysis on the evolution over time of performance of the Netherlands against other countries, taking into account their respective time series.

2.1 Country ranking

Table 2.1 shows the average number of daily unique IP addresses for each global feed we have analyzed, for both top 10 countries with the highest number and for the countries of interest we have mentioned before. Analyzing this table, we can see that the Netherlands only has a modest share of the overall problem. Its high rank number (column #), indicating there are many countries with more infected machines.

The ranking of the Netherlands varies from 36th to 132nd, for the sources we have analyzed. This numbers, however, does not account for the different number of Internet users in each countries, which create unfair compari-

son conditions (e.g., the US has a population $\sim$ 19 times of the Netherlands). To compensate for this, we have produced a ranking in which the number of unique IP addresses seen in the infection data is normalized by the Internet user population of the country, obtained from the World Bank [21]. These results can be seen in Table 2.2.

In terms of it relatively infection rate, so taking the number of Internet users into account, we can see that the Netherlands does quite well. It consistently ranks above 140, meaning there are 140 countries with higher infection rates per user. When compared to a subset of interesting countries, we can conclude that the Netherlands performs above average. It is among the least infected countries in that group. Only for spam, does the Netherlands perform mediocre.

Our global feeds allow to state that the Netherlands, on average, ranks quite well in the world, also in comparison to countries that are relevant points of reference and that care about botnet mitigation. We can also see, however, that Finland still outperforms the Netherlands. It has adopted effective mitigation practices before anyone else and has been a consistent top performer.

2.2 Country performance over time

So the Netherlands ranks quite well in the world and among countries that care about botnet mitigation. The question is to what extent we can credit AbuseHUB with the result. In other words, to what extent did AbuseHUB reinforce the practices of ISPs and make them more effective?

Such questions of causality are difficult to answer under the best of circumstances. However, we might gain some clues as to the impact by looking at the development over time. Since AbuseHUB was not fully operational in the early part of our measurements, we might gauge its impact from how infection rates developed since 2011.

Was the Netherlands already ranked well in 2011 or did it improve? Table 2.3 shows the evolution of the ranking of the countries of interest for the global feeds, broken into years.

First and foremost, we can observe that most of the reference countries have improved over time, with a few exceptions. Some have improved more, others less than the Netherlands.

We have also looked at the speed of clean-up across the reference countries. Figures 2.1–2.5 shows the time series of daily unique IP addresses for the Netherlands and chosen countries. These figures show that (i) the Netherlands infection rates are relatively low in comparison with the compared countries, as also shown in Tables 2.1 and 2.2. From these graphs, it is hard to judge whether there has been an acceleration in clean-up in the Netherlands.

To get a better sense of the relative speed of clean up, we have generated indexed time series. Figures 2.7 – 2.9 show the infection rates of the reference countries all index at 1 at start of the measurement period – i.e., we have divided all daily averages by the first daily average of the first measurement. We only performed this for the data sources that span more than one year (Morto, spam, and Conficker). In this way, all countries start with a value equal to 1 and their variation shows the percentage of infections that have increased or reduced.

Top 10 Countries

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
	IPs	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**
1	GB	4536.67	UK	3652.43	CN	307422.54	CN	351.76	US	1830.03	IN	27733.18
2	JP	4251.1	IT	1317.69	BR	187905.77	IR	137.92	ES	894.54	RU	13590.24
3	IT	3754.36	JP	1106.8	RU	128092.49	BR	78.4	TR	830.66	VN	8535.64
4	US	2542.71	BY	997.17	IN	118912.98	TR	66.41	IN	762.3	BR	8199.17
5	UA	2016.02	GB	760.24	VN	111371.68	JP	44.39	IT	721.99	US	8006.29
6	IN	1690.31	KZ	697.95	KR	69882.05	TW	40.4	JP	623.31	PK	7320.34
7	UK	1665.18	UA	668.14	IT	69629.65	US	38.62	BR	443.99	ID	5697.34
8	FR	1210.62	IR	626.79	AR	62840.21	RU	34.61	MY	439.65	BY	5105.02
9	ID	948.74	ID	620.12	TW	62652.1	DE	34.61	TH	414.16	UA	4940.40
10	KR	842.48	VN	504.19	ID	62375.01	TH	32.57	VN	390.93	SA	4493.54

Countries of Interest

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
CC	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**
NL	**62**	52.98	**132**	10.43	**85**	1299.95	**36**	9.27	**52**	50.17	**40**	683.47
DE	19	398.24	31	152.53	22	31574	9	34.61	14	326.91	21	2489.28
GB	1	4536.67	5	760.24	29	18076.93	19	17.33	15	310.82	26	2042.28
FR	8	1210.62	22	226.65	28	18761.83	35	9.29	11	365.54	32	1256.59
FI	173	3.15	183	2.37	135	118.03	166	1.79	122	6.24	FI	58.35
IT	3	3754.36	2	1317.69	7	69629.65	15	22.47	5	721.99	22	2451.62
ES	28	270.24	54	69.13	15	50135.93	23	13.58	2	894.54	17	2997.32
US	4	2542.71	23	207.49	12	57109.84	7	38.62	1	1830.03	5	8006.29
JP	2	4251.1	3	1106.8	20	32216.55	5	44.39	6	623.31	18	2875.94

Table 2.1: Average Daily Unique IP addresses ranking

Top 10 Countries

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
	IPs	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**
1	SS	26700.00	SS	13100.00	SS	12900.00	PW	2	SS	500000.00	SS	30600.0
2	VA	4166.67	SH	625.00	RO	5402.16	SM	1.87	AS	14473.68	VA	4375.0
3	MP	241.55	MS	351.25	BG	4651.16	MC	2.58	CK	8666.67	NR	3571.42
4	TC	239.84	GE	219.28	VA	4000.00	GI	1.71	PW	6707.32	TK	2500.0
5	AI	239.83	SM	196.47	TW	3352.54	VU	1.95	VG	5471.96	NU	1818.18
6	MH	233.10	BY	191.59	MK	3199.93	AW	5.01	GP	4525.12	WF	1757.66
7	SM	213.53	MP	190.86	HU	2946.08	GD	1.94	MC	4372.81	AS	1694.07
8	GE	195.21	TO	170.45	AL	2874.16	IM	1.77	BG	4362.52	SH	1250.0
9	MC	172.38	MC	157.57	MP	2747.81	KY	1.75	ES	4058.59	BY	980.85
10	KI	166.44	TC	152.44	VN	2715.58	VI	1.99	GE	3561.21	MS	800.84

Countries of Interest

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
CC	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**
NL	**185**	3.34	**206**	0.65	**192**	81.97	**141**	0.58	**127**	395.44	**150**	43.09
DE	165	5.7	190	2.18	112	452.48	148	0.49	78	796.42	163	35.67
GB	24	79.22	115	13.27	140	315.66	171	0.3	72	955	164	35.66
FR	77	23.18	168	4.34	135	359.27	176	0.17	57	1238.9	187	24.06
FI	207	0.6	208	0.49	220	24.48	161	0.37	188	47.91	211	12.10
IT	15	104.11	57	36.54	21	1931.03	139	0.62	12	3483.9	120	67.99
ES	144	7.5	194	1.93	38	1404.13	159	0.38	9	4058.59	93	8394
US	140	9.1	202	0.74	161	206.02	180	0.13	63	1168.5	178	28.88
JP	56	38.77	126	10.09	147	293.87	156	0.4	68	1006.37	183	26.23

Table 2.2: IP addresses/Million Internet Users Ranking

CC	Conficker				Morto				Spam			
	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014
NL	176	183	196	195	68	75	94	76	108	179	156	144
DE	91	105	124	144	31	56	62	80	158	144	115	148
GB	126	135	145	149	64	76	97	115	146	152	133	136
FR	113	128	134	142	115	121	128	133	163	193	162	173
FI	209	214	217	219	141	151	149	143	199	205	206	213
IT	13	22	22	27	103	155	111	111	103	124	60	100
ES	31	34	23	37	57	59	70	70	129	53	39	114
US	150	154	161	162	79	93	93	99	178	164	136	92
JP	138	134	139	152	132	53	157	152	184	157	165	156

Table 2.3: Countries Yearly Ranking (normalized by each countries' Internet Users numbers)

As can be seen, the curves for the Netherlands indicate slightly faster clean up for spam and Conficker, though nothing dramatically different from the other reference countries. We should add that in many of these countries, there have also been anti-botnet efforts. In that sense, we should not expect the Netherlands to be dramatically different.

In sum, the evidence is inconclusive. We will be able bring more data to this issue for the final deliverable by analyzing the feeds obtained throughout 2015.

2.3 Main findings

In this chapter, we have presented the evidence that the Netherlands is doing relatively well in terms of botnet mitigation, compared to the rest of the world. Also among a set of reference countries, the Netherlands performs above average. It is unclear from the available evidence whether AbuseHUB has accelerated the process of mitigation by ISPs. It might have, but most reference countries have followed suit. We hope that an additional year of measurements will provide more clarity.

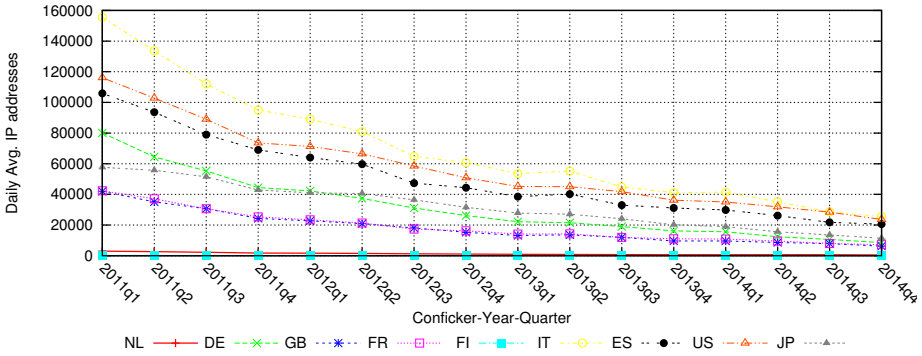

Figure 2.1: Conficker Countries - Daily Average

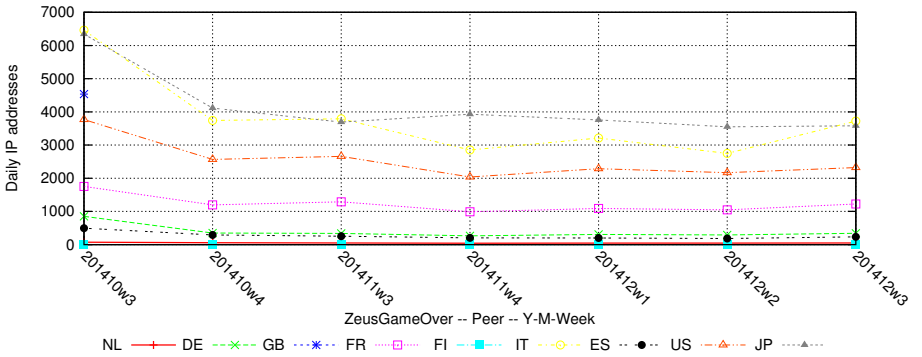

Figure 2.2: GameOver Peer Countries - Daily Average

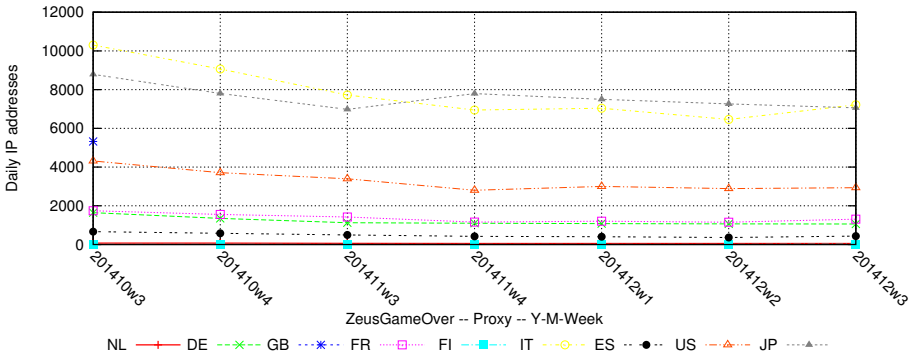

Figure 2.3: GameOver Proxy Countries - Daily Average

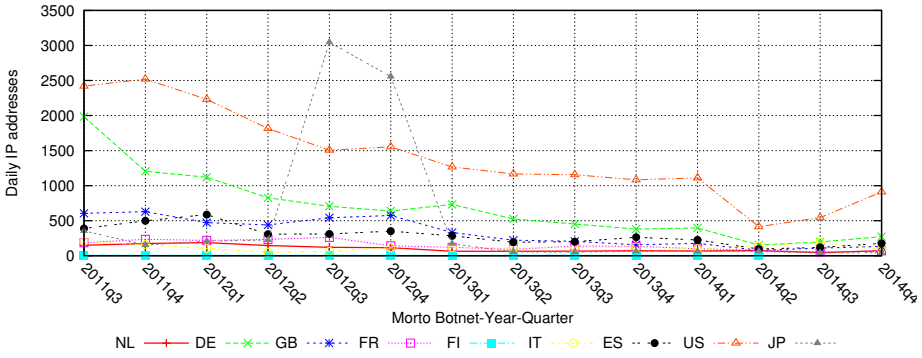

Figure 2.4: Morto Countries - Daily Average

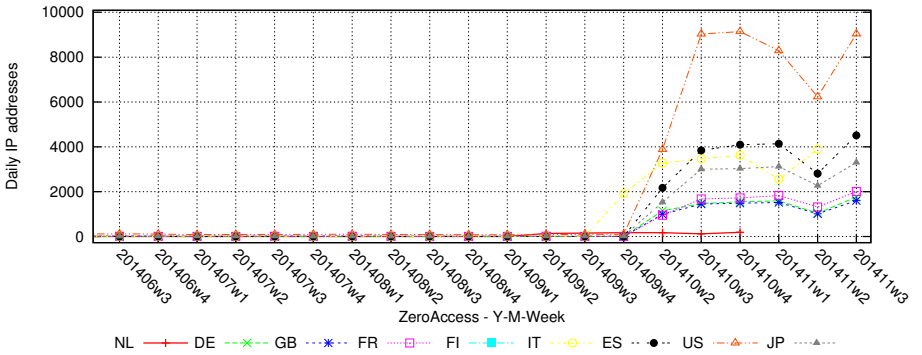

Figure 2.5: ZeroAccess Countries - Daily Average

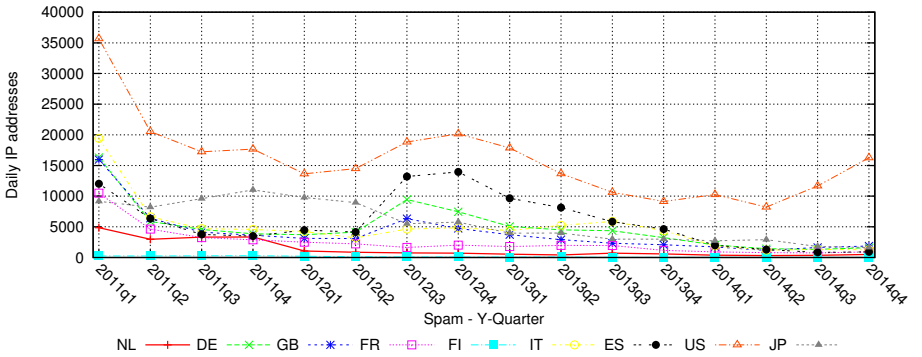

Figure 2.6: Spam Countries - Daily Average

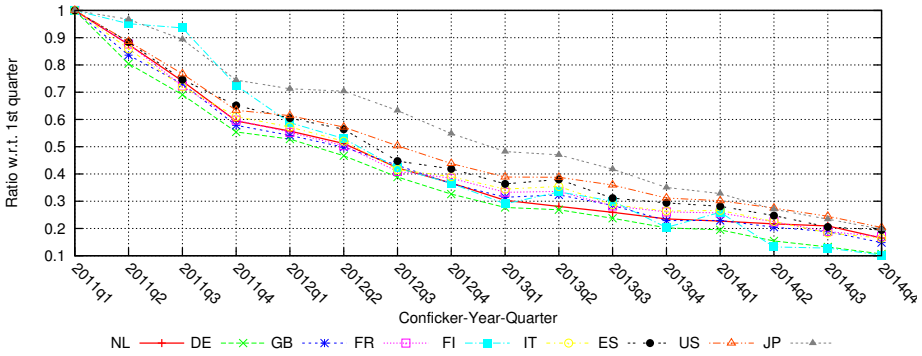

Figure 2.7: Conficker Countries - Indexed w.r.t. first quarter

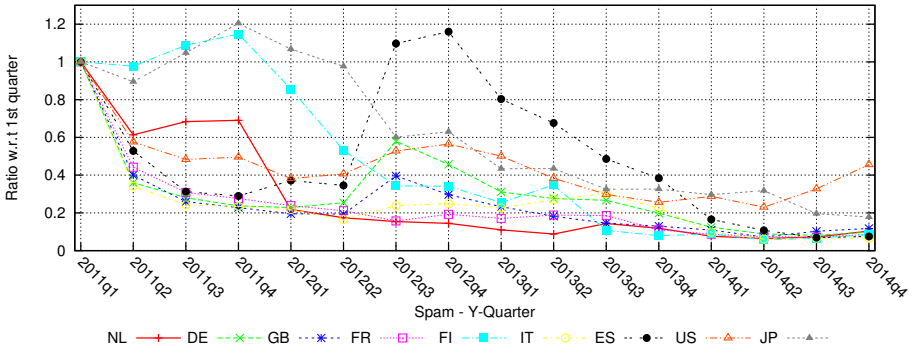

Figure 2.8: Spam Countries - Indexed w.r.t. first quarter

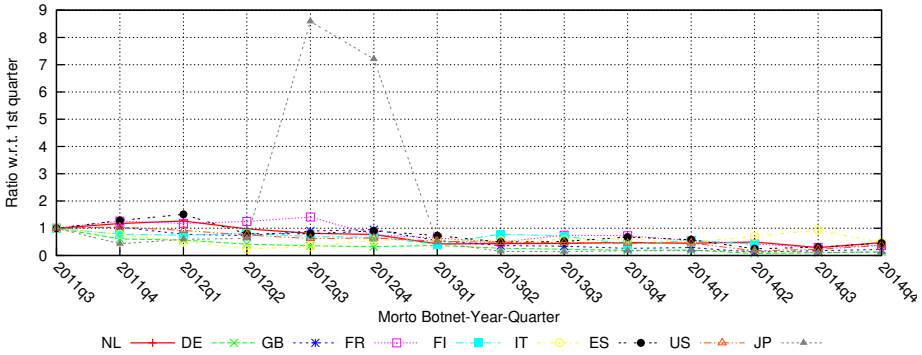

Figure 2.9: Morto Countries - Indexed w.r.t. first quarter

Chapter 3

Dutch ISPs and ABIs efficacy

In Chapter 2, we presented an analysis of the performance of botnet infections in the Netherlands in comparison to other countries. The approached means that for each country we include all infected machines with an IP addresses that geolocates to that country, irrespective of what network it is.

In this chapter, however, we only analyze a subset of the infected IP addresses: only those that belong to ISPs. We exclude national research networks (NRENS), hosting providers, government networks, and all non-ISPs addresses, following the methodology described in Section 1.4.

We investigate two research questions:

- How do Dutch ISPs rank in comparison to ISPs in other countries. It's important to emphasize that our ISP mappings cover 60 countries, so not the whole world.
- How does the Netherlands compare against countries with an anti-botnet initiative (ABI) similar to AbuseHUB and against countries without an ABI?

We cover the first research question in Section 3.1 and the second in Section 3.2. Finally, in Section 3.3, we present the main findings of this chapter.

3.1 Countries's ISP rankings

Table 3.1 shows the average number of daily unique IP addresses for each global abuse feed we have analyzed, for both top 10 countries with the highest number and for the countries of interest we have mentioned before. This table can be seen as a subset of Table 2.1, since it only lists IP addresses associated to Autonomous Systems belonging to ISPs in these countries.

Analyzing this table, we can see that the Netherlands only has a modest share of the overall problem. Its high rank number indicates there are many countries with more infected machines (see column #; rank 1 means the highest observed infection rate, rank 60 the lowest). The ranking of the Netherlands varies from 32nd to 60th, for the sources we have evaluated.

These numbers, while informative about the absolute size of the problem, cannot be compared across countries. Larger countries have more Internet subscribers and that is an important driver of the number of infections. To compensate for these differences, we count the number of infections for the ISPs of each country and normalize it by the total number of Internet subscribers of those ISPs. The subscriber data was obtained from the Telegeography GlobalComms database [18]. These results can be seen in Table 3.2. When compared to a subset of reference countries, we can conclude that the Dutch ISPs performs above average. It is among the least infected countries in that group.

Our global feeds allow to state that the Dutch ISPs, on average, rank quite well in the world, also in comparison to countries that are relevant points of reference and that care about botnet mitigation.

3.2 Anti-botnet initiative countries

For the second research question presented in the introduction of this report, we explored how countries with a national Anti-Botnet Initiative (ABI) performed in comparison to each other and to countries without such an initiative.

We classify each country into one of the three following categories:

- Mature ABI: countries that have ABIs initiatives for longer periods of time.
- New ABI: countries that have only recently started a national ABI.
- No ABI: a selected group of countries without an ABI, but otherwise more or less comparable to the other groups.

Table 3.3 shows which countries we have grouped into the three categories accordingly to the presence and status of their national ABIs.

Figures 3.1–3.6 show the scatter plots of countries ISPs' and their respective number of subscribers (summed for all mapped ISPs for each country) and daily unique IP addresses (daily average), for the global data sources we have evaluated. For example, for The Netherlands (NL), we have summed the average number of IPS for the mapped ISPs and divided by their sum of subscribers.

Analyzing these figures, we can observe that Dutch ISPs (marked with 'x') perform well in comparison to other ISPs. Also, it is clear that countries with mature ABIs, such as the Netherlands, tend to have lower infection rates per subscriber than members of the other groups, for most of the sources. However, these figures also clearly show that there is a lot of variance in each group. Some ISPs in countries without an ABI have lower infection rates that some ISPs in countries with a mature ABI, for example. In other words, the presence of an ABI does not dictate ISP performance in botnet mitigation. Some ISPs will still perform poorly, while other ISPs do well in the absence of a national initiative.

Tables 3.4 through 3.9 show the average ranking of each group among the 60 countries – in other words, we averaged the rank numbers of the countries in each group to get a sense of the overall position of the group.

Top 10 Countries ISPs

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
	CC	**#**	**CC**	**#**	**CC**	**#**	**CC**	**#**	**CC**	**#**	**CC**	**#**
1	GB	3632.9	IT	6459.3	CN	613049.1	CN	1379.0	ES	1505.5	PK	7057.49
2	IT	3069.3	JP	4920.4	BR	251287.2	BR	461.7	US	1421.8	VN	6924.5
3	JP	2807.9	GB	4253.3	VN	123622.2	US	380.7	TR	1214.8	BR	6443.40
4	US	1730.1	BY	2673.7	KR	88290.6	TR	215.1	IT	908.8	RU	5456.9
5	FR	1024.1	US	2100.7	RU	87887.7	DE	186.2	BR	804.1	BY	4483.2
6	KR	743.5	VN	1747.6	IT	80119.1	RU	144.4	VN	701.4	ID	4440.5
7	CN	647.6	ID	1712.4	AR	73718.6	JP	134.3	TH	633.8	SA	3912.0
8	TR	619.2	UA	1622.9	ID	60667.74	TH	119.2	JP	606.4	US	3233.8
9	VN	569.1	TR	14548.5	TR	58784.3	IT	117.2	DE	593.6	CN	3162.0
10	ID	529.4	FR 1161.1		ES	55310.0	VN	94.0	AR	519.0	AR	2697.2

Countries of Interest

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
CC	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**
NL	**50**	**30.2**	**52**	**32.9**	**51**	**1100.3**	**29**	**20.3**	**32**	**84.27**	**36**	**323.24**
DE	13	337.1	13	1078.3	12	40896.6	5	186.2	9	593.6	20	1678.1
GB	1	3632.9	3	4253.3	25	19647.6	12	65.3	13	338.7	22	1488.8
FR	5	1024.1	10	1161.1	23	21877.4	25	33.63	11	437.5	29	735.5
FI	60	6.4	60	10.3	58	155.5	54	4.2	58	8.39	52	45.5
IT	2	3069.3	1	6459.3	6	80119.1	9	117.2	4	908.82	18	1931.5
ES	18	236.1	24	437.4	10	55310.0	11	76.5	1	1505.5	11	2612.3
US	4	1730.1	5	2100.7	15	39205.0	3	380.7	2	1421.8	8	3233.8
JP	3	2807.9	2	4920.4	22	22450.4	7	134.3	8	606.4	25	1241.7

Table 3.1: Average Daily Unique IP addresses ranking for ISPs only

Top 10 Countries ISPs - normalized by million subscribers

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
	CC	**#**	**CC**	**#**	**CC**	**#**	**CC**	**#**	**CC**	**#**	**CC**	**#**
1	BY	247.6	BY	1403.8	ID	37120	CY	46.8	BG	176.491	PK	4062.0
2	IT	228.78	UA	618.5	VN	36162.2	ZA	44.14	TR	143.396	BY	3165.2
3	CY	168.5	IT	481.4	PK	35600.5	IS	36.8	SI	140.587	MA	2650.3
4	GB	168.1	ID	390.8	EG	23099.2	MT	33.7	HR	131.458	KZ	2056.0
5	UA	158.3	PE	362.1	RO	21546.8	TH	28.8	VN	130.653	SA	1823.0
6	EE	141.7	VN	325.5	BR	20259.2	BR	28.4	GR	127.955	ID	1794.6
7	PE	134.3	CY	299.5	RU	20188.6	TR	27.9	ES	124.042	PE	1701.6
8	ZA	129.8	ZA	299.0	AR	19007.6	EG	26.3	LT	123.144	VN	1627.98
9	MT	122.63	MY	293.7	KZ	18913.8	LU	23.7	TH	120.839	RS	1228.8
10	ID	120.8	KZ	292.7	CL	14832.4	IL	23.7	CY	108.352	RO	884.2

Countries of Interest

#	GameOver Peer		GameOver Proxy		Conficker		Morto		ZeroAccess		Spam	
CC	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**	**#**	**IPs**
NL	**58**	**4.5**	**60**	**4.9**	**58**	**182.8**	**59**	**3.25**	**54**	**12.5**	**51**	**51.91**
DE	51	14.3	44	45.8	43	1707.21	39	8.07	44	21.7	48	64.61
GB	4	168.1	14	196.8	46	1067.6	58	3.35	51	15.68	44	73.95
FR	33	41.7	43	47.2	47	1050.2	60	1.52	47	17.81	56	33.5
FI	3	3.8	58	6.2	60	105.9	50	4.53	59	5.01	57	29.00
IT	2	228.7	3	481.4	28	6376.7	33	8.90	20	67.7	33	147.0
ES	46	19.4	46	36.0	31	5430.3	45	7.05	7	124.0	20	241.1
US	43	22.6	48	27.49	54	532.6	49	4.53	46	18.6	53	41.1
JP	20	76.4	23	133.96	52	724.2	50	4.22	48	16.5	55	37.6

Table 3.2: Daily average of unique number of IP addresses seen in data source, normalized by 10^6 (million) subscribers in 60 countries

Group	Country Codes
Mature ABI	AU, DE, IE, JP, KR , FI
Recent ABI	BE, ES, HR, RO, IT, FR, PT, US
No ABI	UK, NO, CZ, NZ, HK, LU

Table 3.3: ABI countries group

Group	2009	2010	2011	2012	2013	2014
Non ABI	46	48	47	46	46	45
New ABI	34	35	34	31	31	30
Matured ABI	44	46	46	45	45	46
NL	59	58	58	58	59	59

Table 3.4: Comparison rankings for Conficker

We can see that countries with a mature ABI rank better than those with a recent or no ABI. This holds across all sources except for the GameOver Zeus feeds. Furthermore, we can also see that the Netherlands ranks substantially better than the other groups, even than the other countries with a mature ABI.

As for improvements over time, the findings are less clear. The dominant pattern is that the groups are quite stable. In some sources, the mature ABI countries improve, in others they are stable or getting slightly worse. For the Netherlands, the ranking remains stable, but that is to some extent to be expected, since it started already among the highest ranks. Only in terms of spam, did we witness a significant improvement over the period when Dutch ISPs increased their efforts and launched AbuseHUB.

3.3 Main Findings

The main finding presented in this chapter is that, in general, Dutch ISPs perform substantially well compared to ISPs from 60 other countries. Moreover, we can also see clearly in our results that the presence of mature Anti-botnet initiatives (ABIs) correlates with lower infection rates per subscribers. However, the lack of an ABI or having a relatively recent ABI does not imply necessarily better of worse performance at the level of the individual ISP. There is a lot of variance at that level that is sometimes larger than the variance among countries. With or without a mature ABI, some ISPs do well, others do worse. In the end, ABIs seem to nudge company policies in the right direction, but they do not dictate it.

Group	2014
Non ABI	39
New ABI	37
Matured ABI	34
NL	58

Table 3.5: Comparison rankings for GameOverPeer

Group	2014
Non ABI	43
New ABI	38
Matured ABI	37
NL	60

Table 3.6: Comparison rankings for GameOver Proxy

Group	2011	2012	2013	2014
Non ABI	33	31	32	23
New ABI	36	38	37	36
Matured ABI	39	38	44	43
NL	57	56	57	53

Table 3.7: Comparison rankings for Morto

Group	2011	2012	2013	2014
Non ABI	40	39	35	40
New ABI	37	36	33	37
Matured ABI	45	44	44	42
NL	44	56	58	55

Table 3.8: Comparison rankings for Spam

Group	2014
Non ABI	40
New ABI	28
Matured ABI	44
NL	54

Table 3.9: Comparison rankings for ZeroAccess

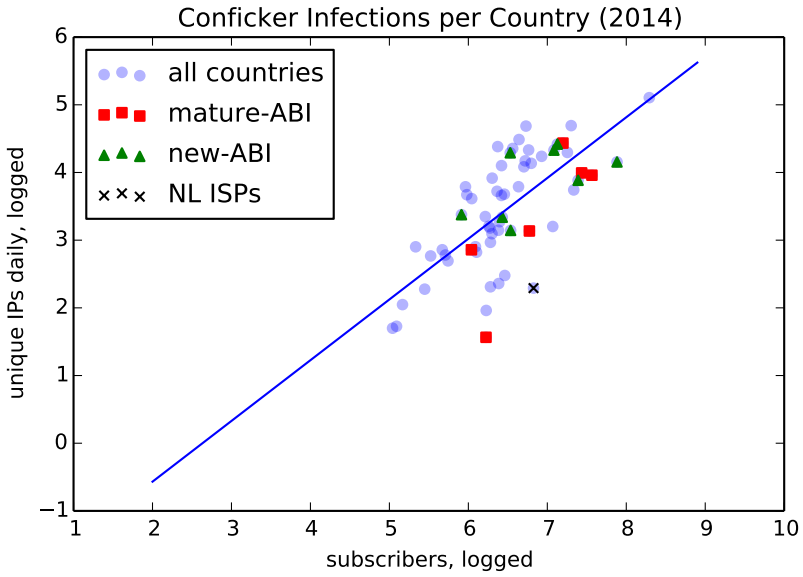

Figure 3.1: Conficker Countries-ABIs scatter plot

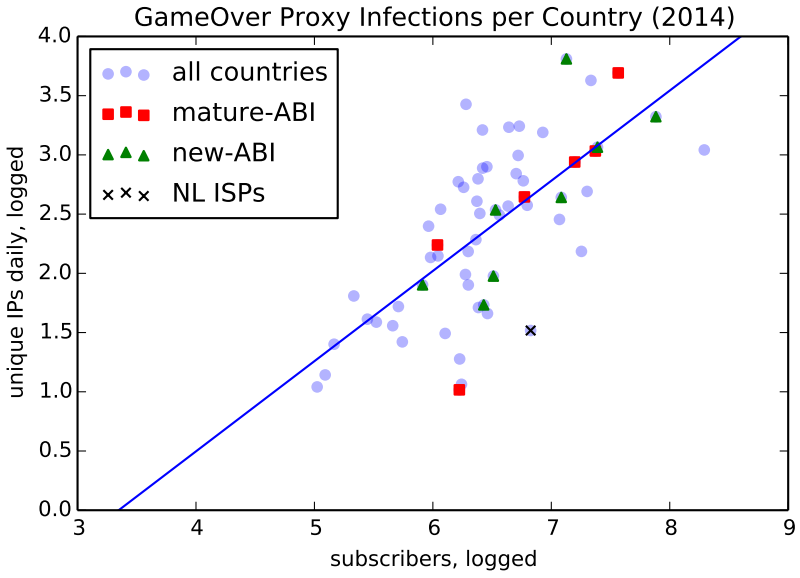

Figure 3.2: Game Over Proxy Countries-ABIs scatter plot

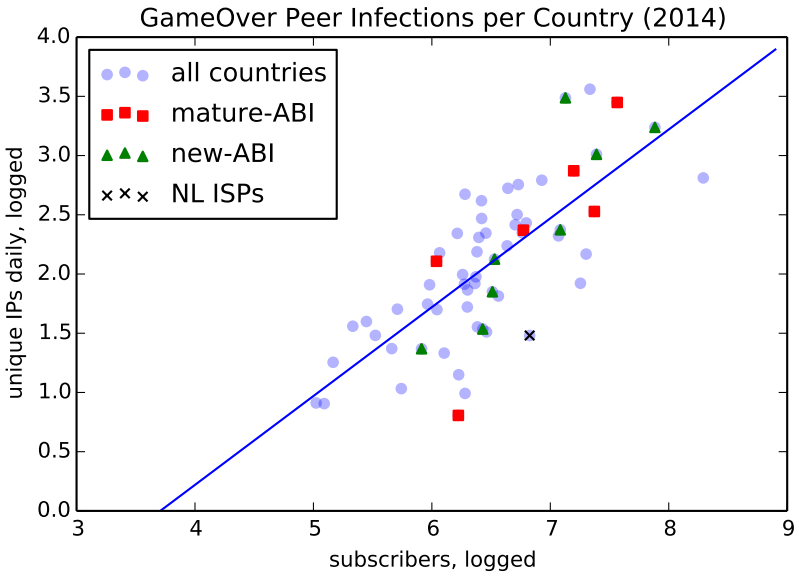

Figure 3.3: GameOver Peer Countries-ABIs scatter plot

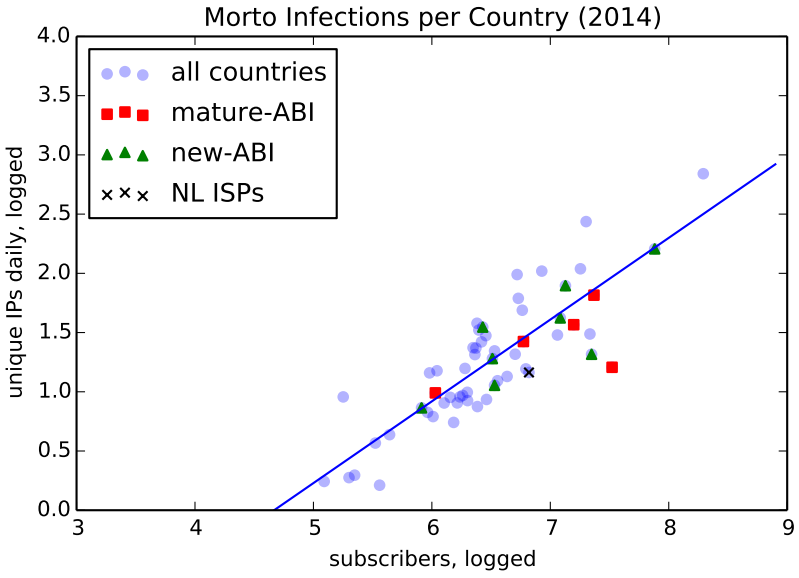

Figure 3.4: Morto Countries-ABIs scatter plot

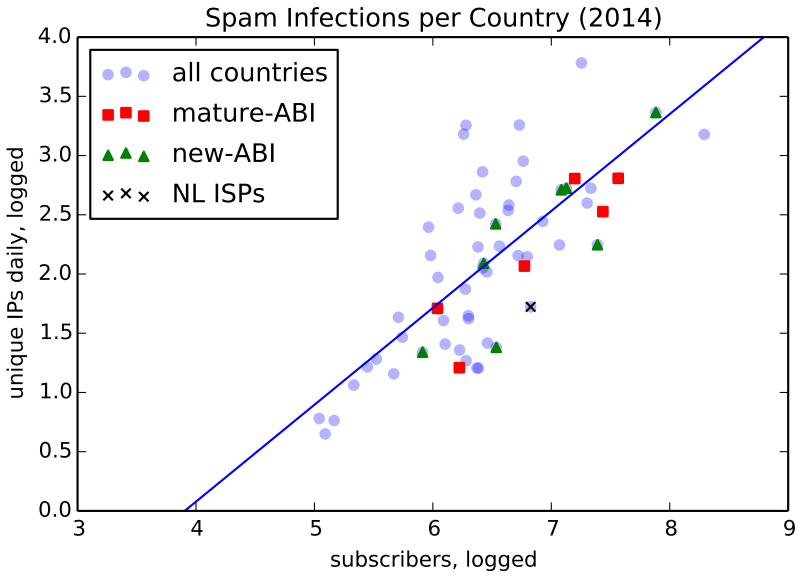

Figure 3.5: Spam ABIs Countrie

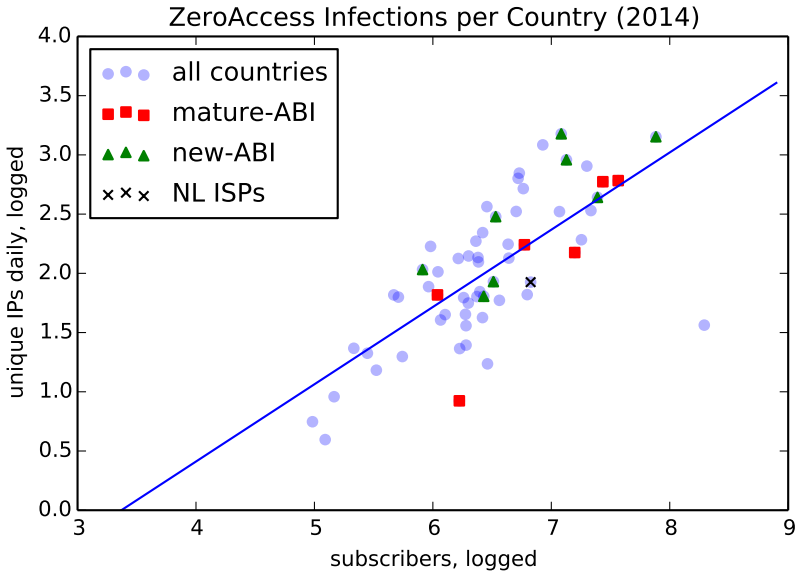

Figure 3.6: ZeroAccess Countries-ABIs scatter plot

Chapter 4

AbuseHUB members compared against non-members

In this chapter, we take a deeper look into the infections within the Netherlands. In our 2011 study, we found that the bulk of all infections were located in the networks of the main broadband providers – around 80 percent, to be precise. In the baseline report of December 2013, we found that this distribution had changed. It was unclear how these finding should be interpreted, since it was based on few data sources.

The broader set of data sources we now have available, allows us to revisit this issue. In addition to the global sources we have analyzed in the previous chapter, we also include two NL-only data feeds we have obtained, as described in Section 1.3.

We explore the distribution of infections in the Netherlands across AbuseHUB members versus non-members. The list of ASes belonging to the first group is provided in Table 1.1. The non-members comprise all other ASes that geo-locate to the Netherlands, which are too many to list here in full.

It is important to emphasize that we *expect the majority of infections to be located in the networks of AbuseHUB members*, for two reasons. First, AbuseHUB members cover 77.6% of the Dutch IP address space.Secondly, the majority of infected computers are typically home users' machines, which are concentrated in the networks of ISPs and thus in AbuseHUB networks, since its members cover the vast majority of the Dutch retail market. If most of infections are indeed located in the networks of AbuseHUB members, that does no imply that the members are worse than non-members: it simply implies that they cover most of vulnerable population.

We will look at the most relevant ASes per source. The idea is to determine:

- How infections are distributed over AbuseHUB members versus non-members;
- How this distribution changes over time.

The rest of this chapter is organized as follows: Section 4.1 discusses the daily averages of infection for members and non-members. In Section 4.2, we

	GameOver Peer	GameOver Proxy	Conficker	Morto
Member	32.69	35.33	826.16	8.59
Non-member	19.03	23.56	427.75	10.50
	ZeroAccess	**Shadowserver Bot**	**Shadowserver MS**	**Spam**
Member	31.14	2786.55	276.93	344.97
Non-member	28.16	1770.10	179.46	339.50

Table 4.1: AbuseHUB members × non-members – average of daily IP addresses

present an analysis on the distribution across of members and non-members over time. Then, in Section 4.3, we explore which non-members ASes contribute the most infections in each data source. Finally, Section 4.4 presents a summary of the main findings.

4.1 Infections in member versus non-member networks

Table 4.1 shows the average daily number of unique IP addresses per group, covering the whole period for which each data source is available. As expected, AbuseHUB members are still responsible for the most of infected IP addresses – which exception of Morto botnet. We analyze in more detail in which member networks these bots are located in Chapter **??**.

The difference among members and non-members are not negligible. In average, across all sources, we observed an average of 6,456 daily IP addresses: members are responsible for 61% of those, while 39% is in non-members ASes. In light of the fact that AbuseHUB members are responsible for 77.6% of the pool of addresses, this distribution is remarkable.

Furthermore, the fraction located within AbuseHUB members is going down over time. The pattern confirms the shift we observed in the observed in the baseline report: compared to 2010, a growing proportion of the infections reside in the networks of non-members. This means that AbuseHUB members are improving faster than non-members.

4.2 Distribution over time

Figures 4.1 to 4.7 show the temporal behavior of AbuseHUB Members × non-members. Notice that the time period during which we track the distribution is different for each data source, depending on its availability.

The dominant pattern across the different sources is that AbuseHUB members improve over time, while non-members are relatively static. This fits with the finding that non-members make up a larger portion of the problem than earlier studies observed.

The shifting distribution implies that non-member ASes become increasingly important in botnet mitigation in the Netherlands. To some extent, this finding is surprising. Since most infections occur in end user devices, we

would expect that broadband providers would harbor most infection. AbuseHUB members cover most of the broadband market in the Netherlands, but its portion of the botnet problem is decreasing. This begs the question: what non-member networks contribute the most infections? We explore this question in the next section.

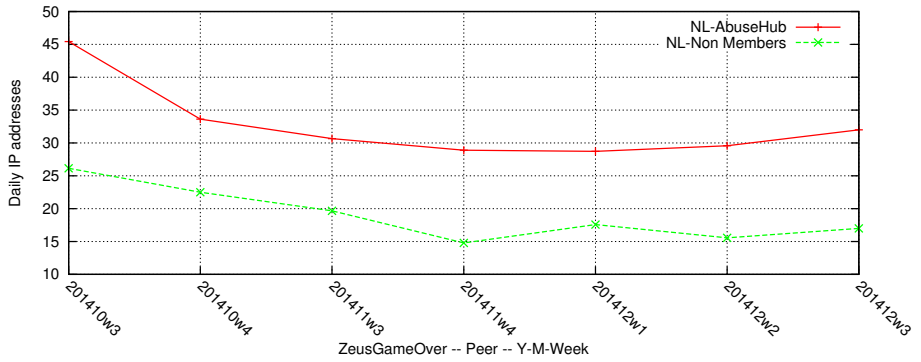

Figure 4.1: Zeus Peer Countries

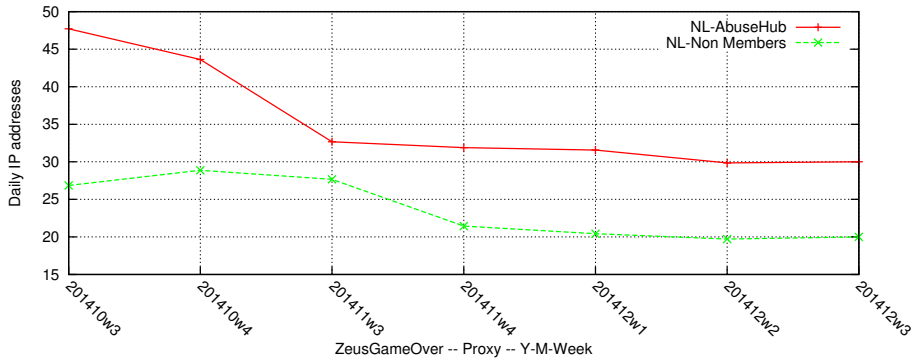

Figure 4.2: Zeus Proxy Countries

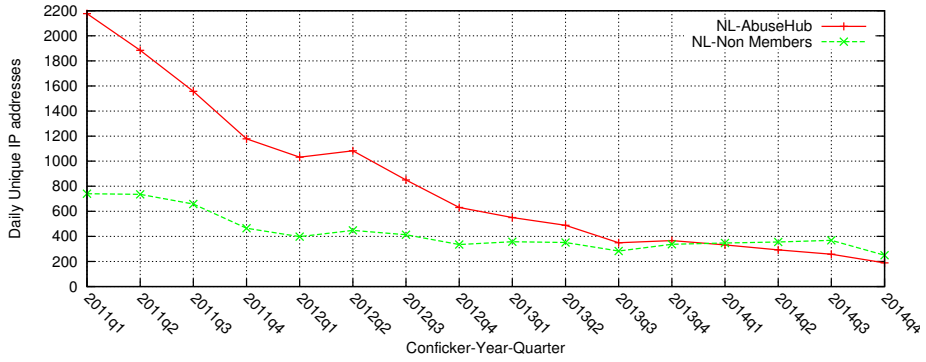

Figure 4.3: Conficker Countries

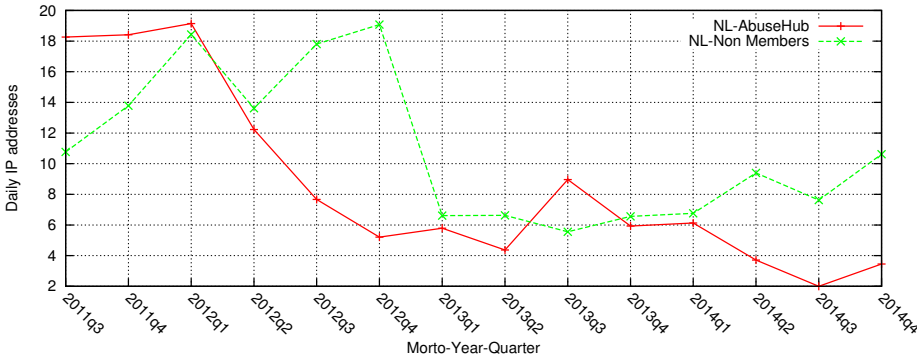

Figure 4.4: Morto members Countries

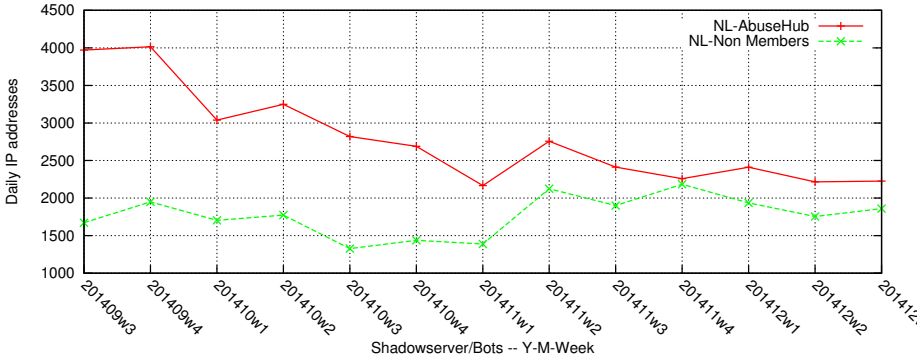

Figure 4.5: Shadowserver Bots

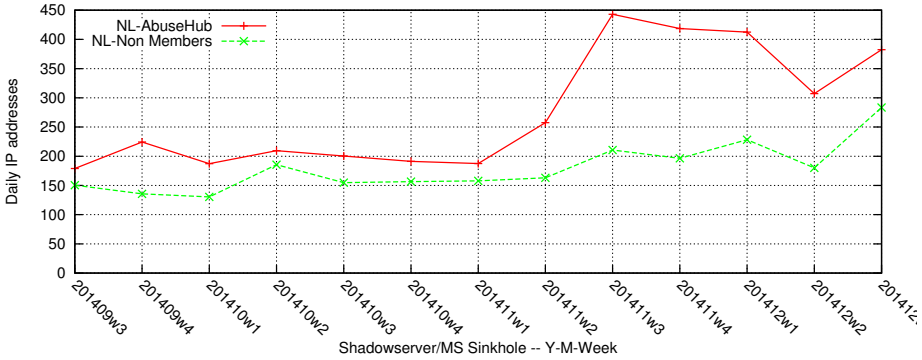

Figure 4.6: Shadowserver MS Bots

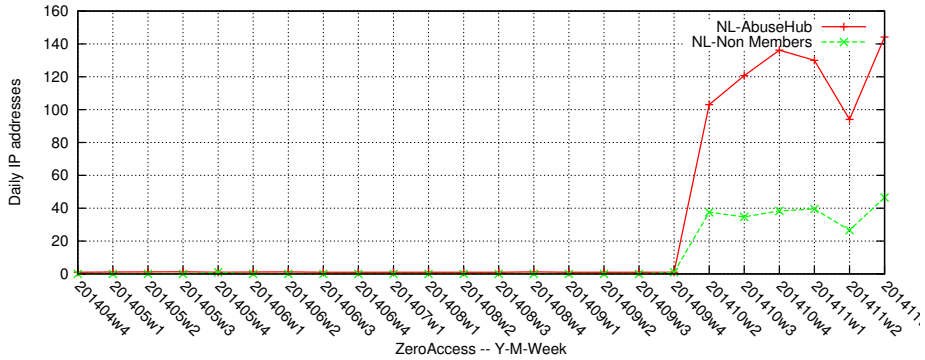

Figure 4.7: ZeroAccess

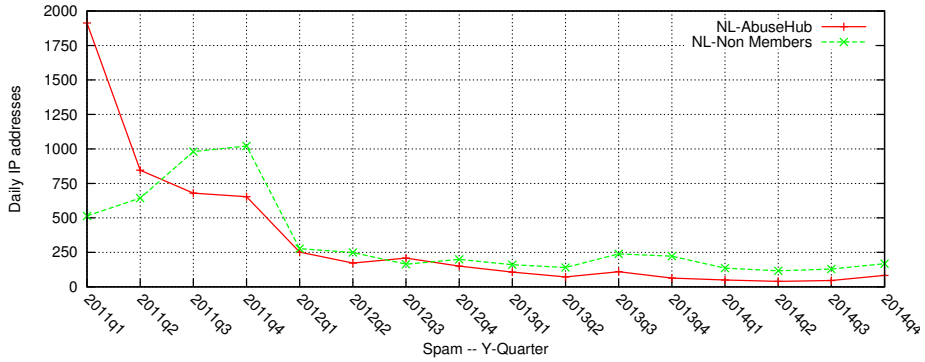

Figure 4.8: Spam

4.3 Most infected non-member networks

In the previous section we have shown that the number of infected IP addresses in AbuseHUB non-members ISPs corresponds to 39% of the total observed in the Netherlands. In this section, we present, for each data source, which ASes contribute the most to that infected population, based on the observed daily average number of infections.

Tables 4.2 – 4.6 show the results for each data source. Not all AS names may be familiar to the reader. Across all tables, two main groups of networks dominate: smaller broadband providers that are not members of AbuseHUB (Online, CAIW) and hosting providers.

The latter is a bit puzzling. For some sources, most notably spam, we know that hosting infrastructure is being used. However, we do know why hosting equipment would show up in botnet sinkholes, since the malware on which the botnet is based only infects Windows end user machines (home or business). We can speculatively suggest two explanations. First, hosting providers are not know as access providers, but they may nevertheless offer access services, bundled with their hosting services. Second, some of these infections may reside in hosted Virtual Machines running Windows. Further analysis is needed to clarify this situation.

4.4 Main findings

Across the different data sources, we can observe that the majority of the infected machines reside in AbuseHUB member networks. This was expected, as the members cover the bulk of all the Dutch IP addresses and of the retail market, in which bots tend to be concentrated. When compared over time, however, we see that the proportion of the infected population in AbuseHUB members is going down. Infections in member networks are diminishing faster than in non-members.

A substantial portion of the botnet problem (on average, 37 percent) resides in non-member networks, most notably smaller broadband providers and hosting providers. Some data sources suggest this portion is growing. All of this suggests that the impact of AbuseHUB can increase by recruiting members among the remaining broadband providers and larger hosting providers – or, to put it differently, that these parties can benefit from joining AbuseHUB in dealing with their botnet problem.

AS	Daily IPs	AS Name
8075	5	Microsoft Corporation,US
29396	2.69048	UNET Unet Network, The Netherlands,NL
34311	1.8	LG-AS LG CNS Europe B.V,NL
30925	1.70732	SPEEDXS-AS CBizz B.V.,NL
12469	1.58333	INFONET-NETHERLAND KPN B.V.,NL
60781	1.5	LEASEWEB-NL LeaseWeb B.V.,NL
5390	1.39394	EURONET Online Breedband B.V.,NL
8426	1.32353	CLARANET-AS ClaraNET LTD,GB
5400	1.30435	BT British Telecommunications plc,GB

Table 4.2: GameOver Peer – Top 10 Non-members

AS	Daily IPs	AS Name
49544	3.14286	INTERACTIVE3D i3d B.V.,NL
30925	1.80952	SPEEDXS-AS CBizz B.V.,NL
29396	1.7619	UNET Unet Network, The Netherlands,NL
43350	1.62069	NFORCE NFOrce Entertainment BV,NL
34311	1.51613	LG-AS LG CNS Europe B.V,NL
60781	1.5	LEASEWEB-NL LeaseWeb B.V.,NL
32475	1.5	SINGLEHOP-INC - SingleHop,US
39309	1.3871	EDUTEL-AS Edutel B.V.,NL
51430	1.33333	ALTUSHOST-NET AltusHost B.V.,NL
8426	1.32353	CLARANET-AS ClaraNET LTD,GB

Table 4.3: GameOver Proxy – Top 10 Non-members

AS	Daily IPs	AS Name
5390	198.807	EURONET Online Breedband B.V.,NL
16265	145.114	LEASEWEB-NETWORK LeaseWeb B.V.,NL
26496	77.2273	AS-26496-GO-DADDY-COM-LLC - GoDaddy.com, LLC,US
15435	72.0795	KABELFOON CAIW Diensten B.V.,NL
200130	66.0795	DIGITALOCEAN-ASN-1 Digital Ocean, Inc.,EU
21155	65.4943	ASN-PROSERVE ProServe B.V.,NL
20857	56.8046	TRANSIP-AS TransIP B.V.,NL
43350	50.7614	NFORCE NFOrce Entertainment BV,NL
34233	42.7857	SUPERIOR-AS Superior B.V.,NL
36351	42.1818	SOFTLAYER - SoftLayer Technologies Inc.,US
12871	41.2273	NL-CONCEPTS KPN B.V.,NL

Table 4.4: ShadowServer Botnet Top 10 Non-members

AS	Daily IPs	AS Name
5390	35.2278	EURONET Online Breedband B.V.,NL
43350	35	NFORCE NFOrce Entertainment BV,NL
57043	16.2051	HOSTKEY-AS HOSTKEY B.V.,NL
16265	9.2987	LEASEWEB-NETWORK LeaseWeb B.V.,NL
15435	7.49367	KABELFOON CAIW Diensten B.V.,NL
60781	5.66667	LEASEWEB-NL LeaseWeb B.V.,NL
12871	5.21519	NL-CONCEPTS KPN B.V.,NL
29396	4.05128	UNET Unet Network, The Netherlands,NL
0	3.98815	a placeholder for non-routed networks -Reserved AS-,ZZ
28878	3.97436	SIGNET-AS Signet B.V.,NL

Table 4.5: ShadowServer Microsoft — Top 10 Non-members

AS	Daily IPs	AS Name
32748	8.11111	STEADFAST - Steadfast Networks,US
49981	4.34981	WORLDSTREAM WorldStream,NL
16265	3.77097	LEASEWEB-NETWORK LeaseWeb B.V.,NL
57043	3.04959	HOSTKEY-AS HOSTKEY B.V.,NL
16276	2.93478	OVH OVH SAS,FR
25074	2.90909	INETBONE-AS PlusServer AG,DE
35017	2.57303	SWIFTWAY-AS Swiftway Sp. z o.o.,GB
29073	2.39316	ECATEL-AS AS29073, Ecatel Network,NL
42267	2.28395	SHIRYO-AS Shiryo Networks B.V.,NL
50673	2.24786	SERVERIUS-AS Serverius Holding B.V.,NL

Table 4.6: Morto – Top 10 Non-members

AS	Daily IPs	AS Name
16265	52.0118	LEASEWEB-NETWORK LeaseWeb B.V.,NL
21155	14.5569	ASN-PROSERVE ProServe B.V.,NL
24875	11.2612	NL-ISPSERVICES Avira B.V.,NL
16131	9.87811	GRAFIX-IS Voiceworks BV,NL
25525	9.49055	REASONNET-AS Reasonnet IP Networks B.V.,NL
12573	9.22644	WIDEXS WideXS B.V.,NL
15703	9.10425	TRUESERVER-AS TrueServer BV AS number,NL
25459	8.38634	NEDZONE-AS NedZone Internet BV,NL
15879	8.24778	ASN-IS IS Group B.V.,NL
20857	7.91717	TRANSIP-AS TransIP B.V.,NL

Table 4.7: Spam — Top 10 Non-members

Chapter 5

Interim conclusions

This interim deliverable provides a preliminary assessment of the impact of the AbuseHUB initiative in reducing the presence of infected computers in the networks of Dutch ISPs. Several key findings can be taken from this report:

- The Netherlands as a whole performs above average in botnet mitigation, not just compared to the rest of the world, but also compared to a set of reference countries (Germany, Great Britain, France, Finland, Italy, Spain, United States and Japan);
- Dutch ISPs (a subset of all networks in the Netherlands) perform above average in comparison with ISPs of other countries.
- More noteworthy is the fact that the Dutch ISPs perform above average compared to ISPs in countries with a similar mature national Anti-Botnet Initiative (ABI);
- The presence of a mature ABI in a country correlates with lower infection rates per subscriber. There is a lot of variance among ISPs, however. Many ISPs in countries without ABIs are also performing well, while some ISPs in countries with an ABI perform below average. In the end, ABIs seem to nudge company policies in the right direction, but they do not dictate it.
- The evidence is, so far, inconclusive whether the operational launch of AbuseHUB has accelerated mitigation. In the worldwide ranking of all infected sources in each country, the Netherlands has improved between 2011 and 2014, but so have the rankings of the reference countries. When we compare only the ISP networks in 60 countries, a more precise comparison, we find that the Netherlands is already at the top of the ranking in the sources that go back to 2011. That makes it very difficult to see any effect of AbuseHUB, since the rank can not really get substantially better. We do see an improvement in dealing with infected machines sending spam, but the other rankings are stable over time. Our preliminary interpretation of this finding is that the process of improved mitigation was already underway and is now supported by AbuseHUB. AbuseHUB is not the start of that process.

- Within the Netherlands, AbuseHUB members still harbor the majority of botnet infections. This is expected as they cover 77.6% of the entire Dutch IPv4 address space, and the bulk of the most vulnerable population: retail users. That being said, the distribution has been shifting towards non-members. The latter have contributed, on average, 39% of the infections. Members appear to have improved faster than non-members;
- The most infected non-members are smaller ISPs that have not (yet) joined AbuseHUB and hosting providers. These operators may benefit from connecting with AbuseHUB and vice versa.

Bibliography

[1] AbuseHUB, "Abuse Information Exchange," AbuseHUB, Jan. 2014. [Online]. Available: http://www.abuseinformationexchange.nl

[2] M. van Eeten, H. Asghari, J. Bauer, and S. Tabatabie, "ISPs and Botnet Mitigation: A Fact-Finding Study on the Dutch Market," Duthc Ministry of Economic Affairs, The Haghe, The Netherlands, Tech. Rep., 2011. [Online]. Available: http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2011/01/13/internet-service-providers-and-botnet-mitigation.html

[3] Maxmind, "Allocation of IP Addresses by Country," Mar 2015. [Online]. Available: https://www.maxmind.com/en/allocation-of-ip-addresses-by-country

[4] ——, "Maxmind," 2012. [Online]. Available: http://www.maxmind.com/

[5] L. Zhuang, J. Dunagan, D. R. Simon, H. J. Wang, and J. D. Tygar, "Characterizing Botnets from Email Spam Records," in *Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*. Berkeley, CA, USA: USENIX Association, 2008.

[6] B. Krebs, "Spam Volumes: Past & Present, Global & Local," Jan 2013. [Online]. Available: http://krebsonsecurity.com/2013/01/spam-volumes-past-present-global-local/

[7] Cisco Systems, "Spam overview - SenderBase," 2014. [Online]. Available: http://www.senderbase.org/static/spam/#tab=1

[8] TrendMicro USA, "Global Spam Map," 2015. [Online]. Available: http://www.trendmicro.com/us/security-intelligence/current-threat-activity/global-spam-map/

[9] ShadowServer, "Conficker Working Group," December 2013. [Online]. Available: http://www.confickerworkinggroup.org/

[10] "Zeus Tracker." [Online]. Available: https://zeustracker.abuse.ch/

[11] H. Binsalleeh, T. Ormerod, A. Boukhtouta, P. Sinha, A. Youssef, M. Debbabi, and L. Wang, "On the analysis of the zeus botnet crimeware toolkit," in *Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on*. IEEE, 2010, pp. 31–38.

[12] M. Riccardi, D. Oro, J. Luna, M. Cremonini, and M. Vilanova, "A framework for financial botnet analysis," in *eCrime Researchers Summit (eCrime), 2010*. IEEE, 2010, pp. 1–7.

[13] N. Falliere and E. Chien, "Zeus: King of the bots," *Symantec Security Respons e (http://bit. ly/3VyFV1)*, 2009.

[14] "Gameover Zeus." [Online]. Available: http://blog.shadowserver.org/2014/06/08/gameover-zeus-cryptolocker/

[15] D. Andriesse and H. Bos, "An analysis of the zeus peer-to-peer protocol," 2013.

[16] "Botnet-Drone." [Online]. Available: https://www.shadowserver.org/wiki/pmwiki.php/Services/Botnet-Drone

[17] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye, "IP Geolocation Databases: Unreliable?" *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 2, pp. 53–56, Apr. 2011.

[18] TeleGeography, "GlobalComms Database Service," Mar. 2014. [Online]. Available: http://www.telegeography.com/research-services/globalcomms-database-service/

[19] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your botnet is my botnet: analysis of a botnet takeover," in *Proceedings of the 16th ACM conference on Computer and communications security*. ACM, 2009, pp. 635–647.

[20] G. C. M. Moura, C. Gañán, Q. Lone, P. Poursaied, H. Asghari, and M. van Eeten, "How Dynamic is the ISPs Address Space? Towards Internet-Wide DHCP Churn Estimation," in *Networking Conference, 2015 IFIP (to appear)*, 2015. [Online]. Available: http://repository.tudelft.nl/view/ir/uuid:a41254a9-812a-4044-8b43-14035499dfba/

[21] The World Bank, "Internet users (per 100 people)," 2015. [Online]. Available: http://data.worldbank.org/indicator/IT.NET.USER.P2