Bad Neighborhoods on the Internet

Giovane C. M. Moura, Ramin Sadre, and Aiko Pras

ABSTRACT

Analogous to the real world, sources of malicious activities on the Internet tend to be concentrated in certain networks instead of being evenly distributed. In this article we formally define and frame such areas as Internet Bad Neighborhoods. By extending the reputation of malicious IP addresses to their neighbors, the bad neighborhood approach ultimately enables attack prediction from unforeseen addresses. We investigate spam and phishing bad neighborhoods, and show how their underlying business models, counter-intuitively, influences the location of the neighborhoods (both geographically and in the IP addressing space). We also show how bad neighborhoods are highly concentrated at a few Internet Service Providers and discuss how our findings can be employed to improve current network and spam filters and incentivize botnet mitigation initiatives.

INTRODUCTION

Spam, phishing, and distributed denial-of-service (DDoS) attacks are just three examples of attacks that have a significant impact on both the Internet and the real world. It is estimated that the losses incurred by spam reach more than US$ 20 billion yearly in the United States alone [1], while DDoS attacks can take down important services, such as the infamous 2007 Estonian attacks, in which citizens could not access their online banks, government, and media websites.

Behind these attacks we usually find a large number of hosts spread all over the world. Some of these attacks are carried out by botnets, which are a large set of distributed compromised machines (called bots or zombies), usually "hijacked" computers located at homes, schools, and businesses, controlled by a botmaster [2].

Even though malicious hosts are spread all over the world, there is evidence that they are, in fact, concentrated in certain networks. Collins *et al.* [3], for example, employed the term "spatial uncleanliness" for clusters of compromised hosts on the Internet, while Chen *et al.* [4] have shown that the distribution of malicious IP addresses is non-uniform. Such concentration resembles the crime distribution in the real world: crime occurs in many places, but tends to be concentrated in certain areas, which are sometimes referred to as "bad neighborhoods."

This resemblance between the real world and the Internet regarding the malicious activities distribution motivated Wanrooij *et al*. to introduce the term Internet Bad Neighborhood [5]. In the real world it is statistically more likely that a crime will be perpetrated in a bad neighborhood in comparison with other locations. The same principle holds for the Internet: the probability of a host behaving badly increases if its neighboring hosts also behave badly. The rationale behind this idea is that different subnetworks have different security policies, and poorly managed subnetworks are more likely to be more often compromised than better managed ones, ultimately increasing the concentration of malicious IP addresses.

In [5] bad neighborhoods were defined in terms of spamming hosts within the same /24 subnetwork, thus hosts that share the same IP address prefix. The contribution of this article is to formalize the bad neighborhood concept and also extend it to hosts within the same Internet Service Provider (ISP) or the same geographic location. Also, we cover not only spam, but also another type of attack (phishing), and show that counter intuitively, the location of bad neighborhoods varies significantly with the underlying business model employed by the attackers.

The main motivation to carry out this research is that knowledge about the concentration of malicious hosts is valuable in attack prediction [6]. Traditionally, blacklisting has been the approach of choice to predict attacks, in which sources involved in previous attacks are filtered in future connections. The bad neighborhood concept furthers traditional blacklisting and improves attack prediction by extending the reputation of malicious IP addresses to their immediate neighbors — that is, by blacklisting their neighboring IP addresses, which are, in turn, more likely to carry out attacks due to the typical concentration [5, 7].

This article summarizes some of the main findings presented in [8], with the aim to make that research comprehensible to a broad audience. We refer the interested reader to [8] for details regarding data collection, measurement approach, statistical analysis, and algorithms. Note that due to space constrains, this article is limited to spam and phishing; additional forms of bad behavior can also be found in the aforementioned document.

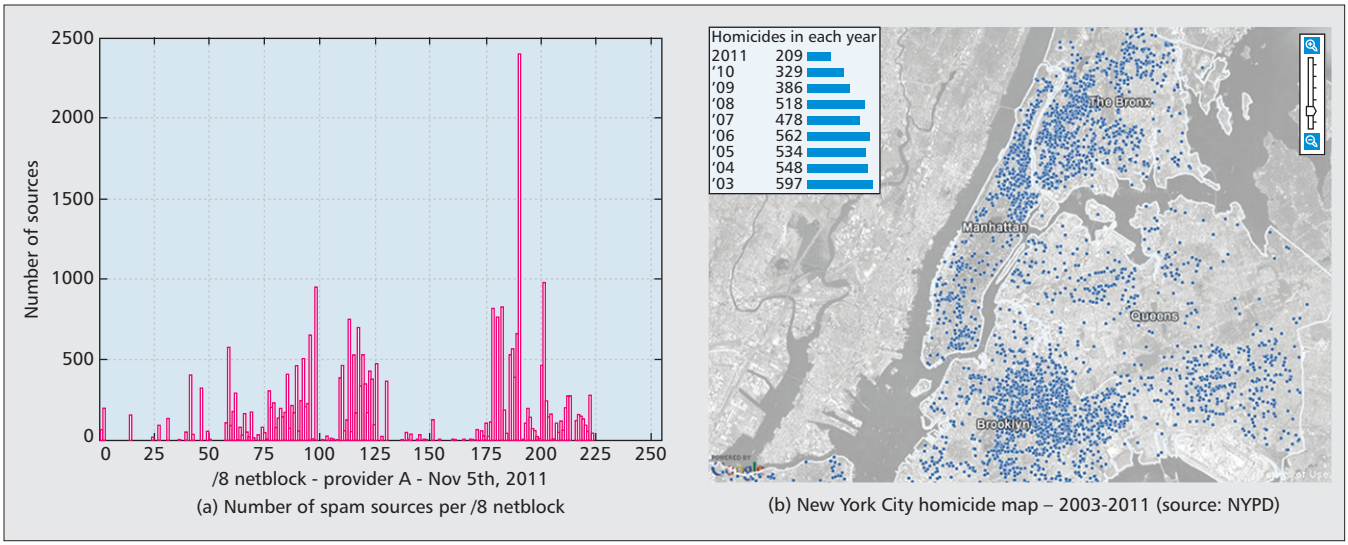
This article is divided as follows. In the following section we formalize the bad neighborhood concept (BadHoods hereafter). Then we cover spam and phishing datasets used in this

Giovane C. M. Moura is with Delft University of Technology.

Ramin Sadre is with Aalborg University.

Aiko Pras is with the University of Twente

0163-6804/14/$25.00 © 2014 IEEE



(a) Number of spam sources per /8 netblock

(b) New York City homicide map – 2003-2011 (source: NYPD)

Figure 1. Distribution of malicious activities on the Internet and in the real world.

article and show which ISPs concentrate most of the malicious hosts. By addressing ISP-based BadHoods we can determine whether certain ISPs "are more tolerant" regarding malicious behavior than others. Then we focus on the geographical distribution of BadHoods by analyzing if certain countries "host" more malicious systems than others. We then present the potential applications of this work. We present conclusions in the final section.

The Internet Bad Neighborhood Concept

Previous research already showed that malicious hosts are not evenly distributed over the IP addressing space [3, 4]. This can also been seen in Fig. 1a, in which we show the distribution of spamming hosts, per /8 prefix,[1] that targeted Provider A — a major hosting provider in The Netherlands — on November 5th, 2011.

This concentration resembles crime distribution in the real world. To illustrate this analogy, consider Fig. 1b, which shows the homicide locations in New York City for the period between 2003 and 2011. As can be seen, some neighborhoods concentrate more homicides than others. Therefore, to reduce these rates more rapidly, the New York Police Department should improve security in higher rate neighborhoods first.

With the purpose of filtering spam in mind, Wanrooij [5] extended the usage of traditional blacklists by spam filters by aggregating individual spamming hosts into BadHood lists (/24), and considering the reputation of bad neighborhoods to filter spam (the likelihood of a message being spam increases if hosts in the same subnetwork of the sender had previously sent spam), and also by scanning messages for URLs of phishing sites. Even though a BadHood-based spam filter was presented previously [5], there was no formal definition, nor were other issues investigated (e.g. changes over time, variation according to application and measurement points).

Therefore we define an Internet Bad Neighborhood as a set of IP addresses clustered according to an **aggregation criterion** in which **a number of IP addresses** perform **a certain malicious activity** over a specified **period of time**.

In this definition, *aggregation criterion* stands for the basic building blocks/criteria employed to cluster malicious IP addresses into BadHoods, which define the size of the neighborhood. Different criteria can be employed for this purpose. The most direct one is network prefixes (e.g, /24, /18, and so on), which can be employed in firewalls and Intrusion Detection and Prevention Systems (IDPS). With this in mind, we have proposed and evaluated two algorithms to aggregate /32 addresses into various network prefixes in [8]. (We present below the results of aggregating IP addresses into two other criteria: geographical location and ISPs.)

The *number of IP addresses*, in turn, is the number of observed malicious IP addresses in the analyzed datasets. It is important to emphasize that this number will likely differ from the total number of IP addresses in the neighborhood, since some addresses within the bad neighborhood could actually be "good IP addresses." For example, an IP-based /24 BadHood, such as 10.10.10.0/24, has a fixed size of 256 IP addresses. However, in general only a fraction of these carry out malicious activities, and some of these addresses may not even be in use. The same principle applies for bad neighborhoods in the real world: there are good innocent citizens living in these places too.

A *certain malicious activity* refers to the abused application by the BadHood (e.g. spam, DDoS, phishing). Therefore, a single host might belong to multiple Bad Neighborhoods. Finally, *period of time* is the time frame used to define a bad neighborhood (e.g. day, weeks). This is an important variable since bad neighborhoods are expected to change over time, due to hosts constantly cleaning-up/compromising and the dynamics of the Dynamic Host Configuration Protocol DHCP [9, 10]. Therefore, as traditional

[1] *We use the Classless Inter-Domain Routing (CIDR) notation for network prefixes/blocks, see RFC 4632.*

As any network security approach, ours also have its limitations too. We build entirely upon the results of third-party detection systems, so the accuracy of the BadHood blacklists is linked to the quality of the sources and, as consequence, is it is prone to false positives (legitimate IPs being flagged as malicious).

blacklists, BadHood blacklists should be constantly updated in order to capture the variation in the source of attacks.

It is important to emphasize that by definition, BadHoods are specific with regard to the measurement point. However, we have also shown in [8] that depending on the application, there is a significant overlap for different measurement points. For example, blacklists provided by blacklist providers (such as the Composite Block List (CBL)[2] or Spamhaus' blacklists) function as supersets of blacklists observed by other measurement.

Singling out Internet Bad Neighborhoods

In the real world crime statistics are of importance when deciding if a neighborhood should be considered "bad" or not. These statistics are generated by companies, police departments, and governments, by keeping track of malicious activities perpetrated in neighborhoods, based on the reports and charges pressed by the victims.

We propose an analogous approach to find Internet BadHoods. The idea is to compile statistics per neighborhood based on the security incidents observed by targets (analogous to victims), which are devices connected to the Internet. Targets should be monitored employing an *intrusion detection* system. As output, the sources of the attack are identified based on the source IP address. After that, a blacklist containing the IP addresses of the sources is generated (a /32 raw blacklist) and used as an input to the aggregation process, in which IP sources (/32) are aggregated into BadHoods, according to an aggregation criterion (e.g. IP prefix such as /24, or geographical information). In the end, a final *BadHood blacklist* is generated.

As with any network security approach, ours also has its limitations. We build entirely upon the results of third-party detection systems, so the accuracy of the BadHood blacklists is linked to the quality of the sources and, as a consequence, it is prone to false positives (legitimate IPs being flagged as malicious). Moreover, the aggregation process into BadHoods may incur more false positives, since legitimate hosts may wind up being blacklisted too. However, as shown in [7], effective attack detection (true positives) and wrongful detection (false positives) depend not solely on the BadHood blacklists, but how they are used in the algorithms — for example, classifying as spam e-mail messages from BadHoods that have at least 10 or 20 malicious IP addresses.

Moreover, our approach does not address fine-grained attack attribution [11], that is, determining the real source of the attack (offending IP address). Attribution is a time-consuming task since malicious users may hide behind a series of computers or use spoofed IP addresses, not to mention the effects incurred by usage of DHCP and NAT by ISPs [9, 10]. We instead blacklist the last IP address in the attack chain, regardless of its owner's intention (we have addressed the ethical issues in [8]). We chose this since we assume the point of view of a network administrator that aims at predicting attacks in real-time.

ISPs and Bad Neighborhoods

For the purpose of mail filtering, we typically aggregate malicious IP addresses into larger network prefixes (e.g. /24) [5, 7]. In this section, however, we aggregate the addresses into ASs, and investigate the relationship between BadHoods and ISPs. We present the evaluated datasets and discuss the findings below.

Evaluated Datasets

As described above, the first step to evaluate BadHoods is to obtain logs of attacks from real-world production networks. Since we want to evaluate if the results hold for different applications, we have obtained blacklists for spam and phishing. For spam, we employed the Composite Block List (CBL), which is a spam blacklist generated by blacklisting every IP address that spams one of CBL's own spam trap infrastructures. CBL is one of the most used for spam filtering and has been previously employed in several research works. For phishing, in turn, we employed data from Phishtank, which is an open community web site in which anyone can "submit, verify, and track phishing websites."[3] It provides a blacklist of URLs of forged websites. Since we need IP addresses instead of URLs to proceed with our analysis, we have resolved all URLs to IP addresses using Google Public DNS servers (`8.8.8.8` and `8.8.4.4`).

To keep the time variable from our BadHood definition as the control variable, we chose the same time frame for both applications: from July 19 to 25, 2012. We then generated a final blacklist for each of the data sets, containing all /32 unique IP addresses observed in the monitoring period. In the end we obtained 9,320,197 unique /32 IP addresses of spam sources, and 3,016 unique /32 IP sources of phishing sites.

ISP-Based Bad Neighborhoods

In this work we employ the autonomous system number (ASN) associated with each IP address in question to determine the ISP it belongs to. As discussed in [8], not every AS is an ISP; other types of organizations may own an ASN. However, as covered in the same material, for most cases the worst organizations are actually the ISPs themselves. To map IP addresses to ASNs, we employ the MaxMind GeoLite ASN database,[4] which is based on BGP routing tables. After that we ranked the ASs by the number of malicious IP addresses and, for spam, also according to the ratio of spamming IP addresses (# of malicious IPs/# of announced IP addresses by the AS). Tables 1 and 2 show the Top 10 ASN (ISPs) for spam and phishing, respectively. Analyzing these tables, we can make the following observations for these datasets.

BadHoods are Highly Concentrated at the ISP Level: At the moment of our analysis there were 42,201 active ASs (announced on global BGP tables), and 35 percent of those were found sending spam. When considering the top 20 ASs, we found that they concentrate almost 50 percent of all spamming IP addresses observed in our data sets, even though they announce less

[2] *http://cbl.abuseat.org/*

[3] *http://www.phishtank.com/*

[4] *http://dev.maxmind.com/geoip/legacy/geolite*

Ranked by Number of Blacklisted IPs						
#	Blacklisted IPs	ASN	AS Name	Ratio (percent)	Announced IPs	Country
1	687,107	AS9829	BSNL (Bharat Sanchar Nigam)	15.4	4,439,552	IN
2	523,679	AS45595	Pakistan Telecom Company	19.11	2,739,968	PK
3	485,944	AS25019	SaudiNet	27.65	1,757,440	SA
4	396,885	AS45899	VNPT Corp	16.88	2,351,104	VN
5	258,996	AS4134	Chinanet	0.23	110,884,096	CN
6	199,679	AS6713	Itissalat Al-MAGHRIB	7.5	2,660,864	MA
7	174,056	AS24560	Bharti Airtel, Telemedia	11.02	1,578,752	IN
8	171,575	AS17803	BSES TeleCom Limited	15.62	1,097,984	IN
9	170,318	AS6147	Telefonica del Peru S.A.A.	12.32	1,381,376	PE
10	156,308	AS7738	Telecomunicacoes da Bahia S.A.	3.63	4,300,800	BR
Ranked by Ratio						
#	Ratio (%)	ASN	AS Name	Blacklisted IPs	Announced IPs	Country
1	62.55	AS37340	SpectraNet Limited	3,523	5,632	NG
2	55.56	AS50604	SC Media SUD SRL	1,138	2,048	RO
3	43.77	AS31208	OJSC MegaFon Network	1,793	4,096	RU
4	40.81	AS131222	Udyog Vihar	78,992	193,536	IN
5	39.2	AS57704	SpeedClick for ITC	803	2,048	PS
6	37.03	AS56995	NetStream Technology	1,517	4,096	PS
7	35.97	AS36912	Orange Cameroun SA	2,947	8,192	CM
8	35.93	AS43766	MTC KSA Mobile	552	1,536	SA
9	35.35	AS58251	Dade Pardazi Novin Yaran Tosei	181	512	IR
10	34.17	AS50948	Behkoush Rayaneh Afzar Co.	700	2,048	IR

Table 1. Top 10 Spam ASes (ranked by blacklisted IPs and ratio).

than five percent of the total number of IPv4 addresses. Moreover, the #1 AS (BSNL) is responsible for more than than seven percent of all the spamming IP addresses observed in our datasets.

Since ASs can have different "sizes" (that is, the number of IP addresses that they announce), we also ranked ASs according to the ratio of spamming IP addresses, using data from BGP reports from Hurricane Electric.[5] Table 1 shows the results. We found that AS37340 had 62.55 percent of its addresses involved in spam. Nine of the top ten ASs, however, are small ones (announced IPs < 10,000). Such ISPs, having such alarming rates of infected IP addresses, typically "neglect/turn a blind eye" to malicious activities in their networks [12] and are truly "spam havens," from which spammers can operate almost freely.

Other studies and industry reports also showed concentrations of malicious hosts in ISPs. For example, van Eeten *et al.* [13] found similar concentrations in ISPs for volume of spam messages. Other reports, such as SpamRankings,[6] show similar figures. Even though the ISP's position may vary according to the measurement point and measurement time frame, the same concentration pattern remains consistent across different studies.

BadHood Locations Vary with the Underlying Business Model/Exploited Application: Comparing Tables 1 and 2, we can see that ASs differ for spam and phishing. The explanation for this variation lies with the specifics of the application

[5] *http://dev.maxmind.com/geoip/legacy/geolite*

[6] *http://www.spamrankings.net/*

The technical realization of the phishing business model requires dependable web servers in which forged websites can be hosted. Therefore, they are typically located in hosting and cloud providers. This also explains the difference between the number of phishing IP addresses and spamming IP addresses per AS.

#	Blacklisted IPs	ASN	AS	Country	Service
1	140	AS36351	SoftLayer Technologies Inc.	US	Cloud provider
2	92	AS32475	SingleHop	US	Cloud provider
3	92	AS16276	OVH Systems	FR	Hosting provider
4	87	AS46606	Bluehost Inc.	US	Hosting provider
5	77	AS21844	ThePlanet.com Internet Services, Inc.	US	Merged with Softlayer
6	50	AS24940	Hetzner Online AG RZ	DE	Hosting provider
7	48	AS47583	Aurimas Rapalis	LT	Hosting provider
8	46	AS32613	iWeb Technologies Inc.	CA	Hosting provider
9	44	AS26496	GoDaddy.com, LLC	US	Hosting provider
10	40	AS7162	Universo Online S.A.	BR	Hosting and content provider

Table 2. Top 10 phishing ASs (ordered according to the absolute number of sources).

being exploited for the attack and its underlying business model and its realization. To generate revenue, spammers need to overcome spam filters and persuade end users to click on their links and acquire products/services. Given current mail filters' high spam detection rates (and that only a tiny fraction of messages that make it to the user ultimately lead to a purchase), spammers heavily rely upon sending a massive amount of messages [14] in order to maximize the probability of a purchase. However, to minimize the effects of blacklisting, spammers prey on vast amounts of free untainted IP addresses to carry out their spam campaigns, usually taking the form of botnets [12].

In contrast, the technical realization of the phishing business model requires dependable web servers in which forged websites can be hosted. Therefore they are typically located in hosting and cloud providers. This also explains the difference between the number of phishing IP addresses and spamming IP addresses per AS. These differences in the bad neighborhoods does not only occur for spam and phishing; in fact, we showed in [8] that BadHoods are application-specific, that is, they vary according to the exploited application.

BadHoods Geographical Location

In this section we aggregate the datasets previously described using their country of origin as aggregation criteria. Several previous reports have shown the geographical distribution of spamming hosts. Our contribution, however, is to investigate whether the same countries account for most of the BadHoods for different applications.

To proceed with it, we needed to perform IP geolocation, which consists of determining the Internet users' geographical location based on their IP address. There are currently two main paradigms to perform this task. Active IP techniques are typically based on network delay measurements, but they do lack scalability and present a high measurement overhead. The database-driven approach, on the other hand, consists of "a database engine" (e.g. SQL/MySQL) containing records for a range of IP addresses. Poese *et al.* [15] found that databases perform very well when geolocating IP addresses to country-level (96 percent to 98 percent success rate, depending on the database). In this article we employed the Maxmind database, since it is one of the most precise commercially available databases [15].

Figures 2a and 2b show the geographical distribution and concentration of malicious IP addresses after being aggregated into country-based BadHoods (colors represent the log-scale absolute number of malicious host per country). Analyzing these figures, we can make the following observations.

Spamming Hosts are Distributed All Over the World: In total, the top 20 countries were responsible for 76.31 percent of all the spamming IP Addresses. Moreover, the countries having more spamming hosts are located in Asia, followed by South America.

Different from Spam, Phishing Sites are Not Found All Over The World: In fact, less than 40 percent of the countries in the world were found having phishing hosts (92 out of 250). Phishing hosts are mostly concentrated in advanced economy nations, with the US leading. The reason is their business model, as covered above: phishing relies upon dependable hosts (offline time means lost business opportunity), and currently most of the datacenters/cloud/hosting providers are located in the United States and other advanced economy nations.[7]

The BRIC Countries (Brazil, Russia, India,

[7] *A map of data centers per country can be found at http://www.datacentermap.com.*

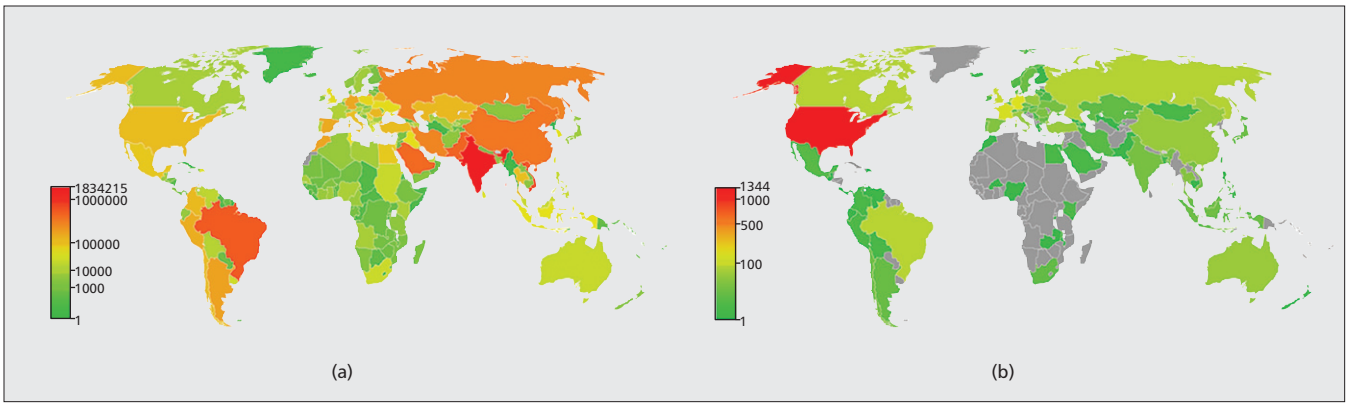


Figure 2. Bad neighborhoods geographical location and concentration.

and China) are Among the Countries with Most Spamming Hosts: These countries are currently experiencing significant economic growth, and in comparison with countries with advanced economies, still have a significant part of their population without Internet access. (The Internet penetration ratios are: BR – 40.6 percent; RU – 43.0 percent; IN – 7.5 percent; CN – 34.3 percent; world – 35 percent).[8] The Internet penetration should increase between nine percent and 15 percent per year until 2015 in the BRIC countries.[9] Combining a growing economy with a large demand for Internet access, we can expect the number of malicious hosts in these countries to increase as more users obtain Internet access, if measures are not taken to improve the security in the networks in these countries. For example, if India would have the same Internet penetration rate as a comparable large country (e.g. the United States at 79 percent) while keeping the same ratio of malicious hosts, it alone would have almost 20 million spammings hosts, which is more than twice the current number we have observed in the entire world in our datasets.

The absolute number of spam sources per country is also correlated to the countries' population and Internet penetration rates. Therefore, we have also shown in [8] that when the population is used as a normalizing rate, we see that the BRIC countries do not rank in the Top 20. In fact, we see five countries that deploy Internet censorship measures (#1-Saudi Arabia, #2-Belarus, #6-Kazakhstan, #8-Vietnam, #16-Tunisia)[10] among the top 20 countries. We believe that while trying to circumvent censorship, users in these countries might wind up getting their computers infected by accessing open proxies, malicious websites, or installing malicious tools.

Existing and Future Applications

In this section we discuss three main applications of the Internet Bad Neighborhood concept: attack prediction, lightweight spam filtering, and botnet mitigation incentives.

Attack Prediction

Blacklisting is a technique widely employed to defend against malicious traffic on the Internet. Its roots can be traced back to as early as 1997, in which lists of IP addresses involved in spam and other objectionable behavior were shared as Border Gateway Protocol (BGP) feeds. Later, in 2010, the IETF standardized a Domain Name Server (DNS)-based approach to share blacklists and whitelists [16].

Blacklists, such as the ones evaluated in this article, can be seen as an attempt to predict attacks based on historical past [6]. The Bad Neighborhood concept furthers traditional blacklisting and improves attack prediction by *extending the reputation of malicious IP addresses to their immediate neighbors* — that is, by blacklisting their neighboring IP addresses. Neighbors, in this case, may refer to hosts within the same prefix (e.g., /24, /23), or even coarser aggregation criteria, such as ASN. This approach has proved to be effective in predicting spam messages [5, 7].

Based on that, we envision attack prediction models as a promising area to apply the bad neighborhood idea. BadHood-based models should take into account not only the observed concentration of malicious hosts, but also the other findings demonstrated in [8] that this article builds upon. To mention a few such models, as also shown here, should be tailored to the type of attack (or exploited application), and the historical past should be considered to cope with increasingly stealthy attacks. We have seen that 40–95 percent of BadHoods are likely to strike more than once within a one week period, and that 85 percent of these do carry out a second attack within the first five days from the first attack [8]. Such prediction models are related to other works, such as by Soldo *et al.* [6], where the authors have employed a recommendation system to predict attacks.

Lightweight Spam Filtering

Traditional machine learning techniques employed to filter spam, such as Bayesian networks, despite working relatively well, have some drawbacks, such as being CPU-intensive. The authors in [5] found that mail severs spend 64 percent of CPU time on spam filters. To cope with that, the authors employed a bad neighborhood-based mail filter that also scanned for phishing URLs in the message bodies that led, for their environment, to a 20-fold throughput gain over SpamAssassin, a renowned spam filter.

Their study, however, was carried out on a small scale (<10k messages). After that we eval-

8 http://www.itu.int/ITU-D/ict/statistics/material/excel/2010/

9 http://www.bcg.com/documents/file58645.pdf

10 http://opennet.net/

Our algorithm uses as parameter the number of malicious addresses per BadHood to tell if a message is spam, and differently from [5], we did not look into the messages for malicious URLs. We were able detect more than 90 percent of the spamming sources, employing solely IP lookups.

uated the effectiveness of various /24 blacklists in simple spam filtering in [7]. In total, we have evaluated more than one million messages. Our algorithm uses as a parameter the number of malicious addresses per BadHood to tell if a message is spam, and different from [5], we did not look into the messages for malicious URLs. We were able to detect more than 90 percent of the spamming sources, employing IP lookups only.

However, one major concern is false positives: legitimate messages must not be flagged as spam. We have shown in the same study that, depending on the data set, an aggressive BadHood-based algorithm can lead to a large number of false positives, which is unacceptable for e-mail. By tuning our algorithm, we were able to significantly reduce the false positive rate, at the expense of reducing spam detection.

Based on that, one main application of the bad neighborhood concept is to develop and evaluate multi-layer mail filtering algorithms. At the first layer, a BadHood-based algorithm is employed to filter out e-mail from the most dangerous neighborhoods, that, at the same time, keeps the false positives rate low. The second layer would comprise analysis of URLs and/or contents within the messages (we learned after [8] that companies such as Google, IBM, and Symantec employ similar approaches in their mail filter products, that is, subnetwork reputation as their first line of defense against spam).

Botnet Mitigation Incentives

One of the implications of the concentration of malicious hosts in bad neighborhoods is that ISPs themselves form a centralized control point to tackle such attacks. Even though they can perform such activities, it does not mean that they should, or even that they are legally allowed. Therefore, the findings here presented should be used to provide incentives [13] for ISPs to tackle malicious traffic originated in their own networks.

Moreover, statistics compiled at the country level should also be employed to encourage initiatives organized by countries to improve security in their own networks, via legislation (similar to the United States' CAN SPAM act and the European Union's Directive on Privacy and Electronic Communications (2002/58)), and other public-private initiatives, for example, by coordinating efforts at the national level through national Computer Emergency Readiness Teams (CERTs) and/or national cybersecurity centers.

Conclusions

Malicious IP addresses are not evenly distributed on the Internet [3, 4]. In this article we propose (and formalize) to frame such concentrations as Internet Bad Neighborhoods, analogous to real world bad neighborhoods. The Bad Neighborhood concept furthers traditional blacklisting for network defense and improves attack prediction by extending the reputation of malicious IP addresses to their immediate neighbors.

We have shown that Internet BadHoods are highly concentrated at the ISP level, and not only at subnetworks. The top 20 ASs, which are somehow comparable to ISPs, concentrated almost 50 percent of all spamming IP addresses observed in our data sets, from a total of more than 40 thousand active during our measurements. In the worst case a single ISP concentrates almost eight percent of all the spamming addresses observed for the entire world in our datasets. We also found that some ISPs have an alarming ratio of more than 60 percent of their announced IPs involved in spam — typically small ISPs which operate like spam havens. We also showed that the position in the ranks varies according to the analyzed measurement data, but the concentration pattern holds for different studies/reports.

Another finding is that the location of the BadHoods varies according to the underlying business model. While spam is distributed all over the world (but concentrated in Southern Asia), phishing Bad Neighborhoods, on the other hand, are mostly concentrated in the United States and other developed nations. *Phishing* relies upon dependable hosts (offline time means lost business opportunity), and currently most data centers/cloud/hosting providers are located in the United States and other advanced economy nations. Spammers, on the other hand, base their business model on sending vast amounts of spam messages from untainted IP addresses to minimize traditional blacklisting efficiency, usually employing botnets.

We have seen how BadHoods are also clearly visible at the country level. Out of 229 countries found having spamming hosts, a single one (India) was found concentrating almost 20 percent of worldwide spamming IP addresses, followed by Vietnam and Brazil. We also discussed the potential alarming implications and showed how these results can be employed to incentivize public-private initiatives to mitigate botnets in the networks of ISPs.

Finally, we addressed the potential applications for BadHood-related research: attack prediction, lightweight spam filter, and fostering botnet mitigation incentives. This article covered IPv4 BadHoods, since IPv6 attacks are not yet so common. With the increasing adoption of IPv6, however, we can expect more attacks from IPv6 hosts. As we have covered in [8], an aggregation-style approach like the Internet Bad Neighborhoods approach is a necessity when dealing with IPv6 attacks. As covered in RFC 6177, ISPs will likely assign /48 prefixes to home users, leading to 2^{48} possible IPv6 addresses per costumer. Therefore, BadHoods should be aggregated at least using /48 prefixes for IPv6-based BadHoods. However, further investigation will be necessary to confirm if our findings hold for IPv6 addresses.

References

[1] J. M. Rao and D. H. Reiley, "The Economics of Spam," *The Journal of Economic Perspectives*, vol. 26, no. 3, 2012, pp. 87–110.

[2] E. Cooke, F. Jahanian, and D. McPherson, "The Zombie Roundup: Understanding, Detecting, and Disrupting Botnets," *Proc. Steps to Reducing Unwanted Traffic on the Internet Wksp.*, Berkeley, CA, USA: USENIX Association, 2005, pp. 6–6.

[3] M. P. Collins *et al.*, "Using Uncleanliness to Predict Future Botnet Addresses," *Proc. 7th ACM SIGCOMM Conf. Internet Measurements*, ser. IMC '07. New York, NY, USA: ACM, 2007, pp. 93–104.

[4] Z. Chen, C. Ji, and P. Barford, "Spatial-Temporal Characteristics of Internet Malicious Sources," *Proc. INFOCOM 2008, IEEE 27th Conf. Computer Commun.*, Apr. 2008, pp. 2306–14.
[5] W. van Wanrooij and A. Pras, "Filtering Spam from Bad Neighborhoods," *Int'l. J. Network Management*, vol. 20, no. 6, Nov. 201, pp. 433–44.
[6] F. Soldo, A. Le, and A. Markopoulou, "Blacklisting Recommendation System: Using Spatio-Temporal Patterns to Predict Future Attacks," *IEEE JSAC*, vol. 29, no. 7, Aug. 2011, pp. 1423–37.
[7] G. C. M. Moura *et al.*, "Evaluating Third-Party Bad Neighborhood Blacklists for Spam Detection," *Proc. IFIP/IEEE Int'l. Symp. Integrated Network Management (IM 2013)*, Ghent, Belgium, May 2013.
[8] G. C. M. Moura, "Internet Bad Neighborhoods," Ph.D. dissertation, University of Twente, Enschede, The Netherlands, Mar. 2013, available: http://dx.doi.org/10.3990/1.9789036534604.
[9] M. Fabian and M. Terzis, "My Botnet is Bigger Than Yours (Maybe, Better Than Yours): Why Size Estimates Remain Challenging," *Proc. 1st USENIX Wksp. Hot Topics in Understanding Botnets*, Cambridge, USA, 2007.
[10] B. Stone-Gross *et al.*, "Your Botnet is My Botnet: Analysis of a Botnet Takeover," *Proc. 16th ACM Conf. Computer and Communications Security*, ACM, 2009, pp. 635–47.
[11] D. A. Wheeler and G. N. Larsen, "Techniques for Cyber Attack Attribution," Institute for Defense Analyses, Alexandria, VA, USA, Tech. Rep., 2003, available: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA468859.
[12] A. Ramachandran and N. Feamster, "Understanding the Network-Level Behavior of Spammers," *Proc. 2006 Conf. Applications, Technologies, Architectures, and Protocols for Computer Communications*, ser. SIGCOMM '06. New York, NY, USA: ACM, 2006, pp. 291–302.
[13] M. van Eeten *et al.*, "The Role of Internet Service Providers in Botnet Mitigation: An Empirical Analysis Based on Spam Data," *Proc. WEIS 2010: 9th Wksp. Economics of Information Security*, 2010.
[14] D. McCoy *et al.*, "PharmaLeaks: Understanding the Business of Online Pharmaceutical Affiliate Programs," *Proc. 21st USENIX Security Symposium*, Bellevue, Washington, USA: USENIX Association, Aug. 2012.
[15] I. Poese *et al.*, "IP Geolocation Databases: Unreliable?" *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 2, Apr. 2011, pp. 53–56.
[16] J. Levine, "DNS Blacklists and Whitelists," RFC 5782 (Informational), Internet Engineering Task Force, Feb. 2010.

Biographies

Giovane C. M. Moura (g.c.moreiramoura@tudelft.nl) is a Postdoctoral Researcher with the Economics of Cybersecurity group at Delft University of Technology (TU Delft), The Netherlands. He received a Ph.D. degree from the University of Twente (UT) for his thesis entitled "Internet Bad Neighborhoods," and a master's degree from the Federal University of Rio Grande do Sul (UFRGS), Brazil. His research interests include Internet measurements, traffic monitoring and analysis, and design of network intrusion detection systems.

Ramin Sadre (rsadre@cs.aau.dk) is an Assistant Professor with the Distributed and Embedded Systems group at Aalborg University, Denmark. He received a Ph.D. degree from the University of Twente for his thesis titled "Decomposition Based Analysis of Queuing Networks." His research interests include traffic modeling, the design and analytical performance evaluation of communication systems, and the design of network intrusion detection systems.

Aiko Pras (A.Pras@utwente.nl) is a Professor with the Design and Analysis of Communication Systems Group (DACS) at the University of Twente, The Netherlands. He received a Ph.D. degree for his thesis titled "Network Management Architectures." His research interests include network management technologies, network monitoring, measurements and security. He is chairing the IFIP Technical Committee 6 on "Communications Systems," and is Project Leader of the European Network of Excellence on "Management of the Future Internet" (FLAMINGO). He is a steering committee member of several conferences, including IM/NOMS and CNSM, and a series/associate editor for the *International Journal of Network Management* . See also: http://wwwhome.cs.utwente.nl/~pras/bio.html.

Malicious IP addresses are not evenly distributed on the Internet. We propose to frame such concentrations as Internet Bad Neighborhoods — analogous to real world bad neighborhoods. The Bad Neighborhood concept furthers traditional blacklisting for network defense and improves attack prediction.