Can ISPs Help Mitigate IoT Malware? A Longitudinal Study of Broadband ISP Security Efforts

Arman Noroozian*, Elsa Turcios Rodriguez*, Elmer Lastdrager†,
Takahiro Kasama‡, Michel van Eeten*, and Carlos H. Gañán*
*TU-Delft (Netherlands), †SIDN Labs (Netherlands), ‡NICT (Japan)

***Abstract*—For the mitigation of compromised Internet of Things (IoT) devices we rely on Internet Service Providers (ISPs) and their users. Given that devices are in the hands of their subscribers, what can ISPs realistically do? This study examines the effects of ISP countermeasures on infections caused by variants of the notorious Mirai family of IoT malware, still among the dominant families. We collect and analyze more than 4 years of longitudinal darknet data tracking Mirai-like infections in conjunction with threat intelligence data on various other IoT and non-IoT botnets across the globe from January 2016 to May 2020. We measure the effect of two ISP countermeasures on Mirai variant infection numbers: (i) reducing the attack surface (i.e., closing ports that are used by the malware for propagation) and (ii) ISPs increasing their general network hygiene and malware removal efforts (as observed by proxy of the remediation of infections of other families of IoT and non-IoT malware and reductions in the number of DDoS amplifiers in their networks). We map our infection data to 342 broadband providers that have the bulk of the broadband market share in their respective 83 countries. We find that the number of infections correlates strongly with the number of ISP subscribers (R^2=0.55). Yet, infection numbers can still vary by three orders of magnitude even for ISPs with comparable subscriber numbers. We observe that many ISPs, together with their subscribers, have reduced their attack surface for IoT compromise by blocking traffic to commonly-exploited infection vectors such as Telnet and FTP. We statistically estimate the impact of these reductions on infection levels and, counter-intuitively, find no significant impact. In contrast, we do find a significant impact for improving general network hygiene and best malware mitigation practices. ISPs that were more successful in reducing DDoS amplifiers and non-Mirai malware infections in their networks also end up with significantly lower Mirai infection rates. In other words, rather than investing in IoT-specific countermeasures like reducing the attack surface, our findings suggest that ISPs might be better off investing in general security efforts to improve network hygiene and clean up abuse.**

***Index Terms*—Mirai, Internet of Things, IoT, Malware, ISP, Countermeasure, Remediation**

1. Introduction

Poorly-secured Internet-of-Things (IoT) devices have been with us for more than five years now. They have become the staple of threat assessments and even parody[1]. During this period, millions of devices were compromised by the Mirai malware family [1]. Mirai not only caused the first peak of infections, but it has persisted as a dominant malware family until today. One recent industry report named Mirai the "king of IoT malware" [2]. In 2019, Kaspersky reported that Mirai is still the leading malware family and responsible for 21% of the IoT infected devices [3]. What keeps Mirai a relevant threat is that it exploits default credentials, a problem that has still not been fixed by many manufacturers. The Open Web Application Security Project (OWASP) describes this as the top threat for IoT [4]. Additionally, the release of Mirai's source code has allowed attackers to add exploit code on top of its credential-based attacks and create newer variants which go beyond launching Denial of Service (DoS) attacks. According to industry reports, compromised IoT devices have been abused for purposes ranging from DoS attacks to the installation of tor nodes, packet sniffers, and trojans, all the way to performing crypto-jacking, DNS hijacking, and credential collection [5].

As long as manufacturers keep releasing new insecure devices into our markets, the brunt of remediating infected IoT lies with both the end users who own the devices and the Internet Service Providers (ISPs) where more than 80% of the devices are located [6]. While recent work has studied the practices and perceptions of end users when it comes to IoT security [7–9], little attention has been paid to the role of ISPs. One exception is a study that found ISPs to be able to use quarantining of infected devices as a way to enforce remediation by the customer [6]. We know virtually nothing however, of what other ISP practices might be effective.

In this paper, we explore two additional security strategies. First, can ISPs stem the spread of IoT infections by reducing the attack surface for the malware? In other words, does it help to close network ports that are used for propagation? ISPs can administer default router configurations that block ports of commonly exploited services such as Telnet and FTP [10]. This is similar to past approaches in mitigating spam, where port 25 would be blocked to prevent the distribution of spam from consumer connections. Second, are general ISP security measures for network hygiene and abuse remediation also effective against Mirai infections? In other words, are ISP practices from the area of fighting Windows-based malware [11, 12] applicable to IoT as well? And does better network hygiene correlate with lower infection rates for Mirai?

We analyze the impact of these countermeasures by examining their effect on Mirai infection numbers. Mirai is not only a dominant IoT malware family that has been

1. https://twitter.com/internetofshit

2021 IEEE European Symposium on Security and Privacy (EuroS&P) | 978-1-6654-1491-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/EUROSP51992.2021.00031

DOI 10.1109/EuroSP51992.2021.00031

around for five years, but many of its variants also have the benefit of being trackable, thus providing us with a longitudinal view of its evolution. Furthermore, its infections are relatively straightforward to remediate in most cases (i.e., changing the default passwords and power cycling the device). These features—being dominant, being observable, providing a longitudinal view, and allowing a remediation path that ISPs can recommend to their users—makes Mirai infections eminently suitable for observing what ISPs can achieve in combating IoT malware. In other words, if ISPs are not able to combat the tide of Mirai, it would be unlikely that they will achieve better results against the much more sophisticated emerging IoT malware variants, as these are harder to detect and harder to clean up.

The question of whether these ISP security efforts are helpful can only be answered empirically by analyzing long-term longitudinal data. To this end we track Mirai infections across 342 ISPs operating in 83 countries over the period of January 2016 to May 2020 using darknet data. This period covers Mirai's inception point, its peak and extends all the way to current day Mirai variants. To complement the infection data, we collect longitudinal data on the number of customer connections where common propagation ports for Mirai were open – more specifically TCP/23+2323 (Telnet), TCP/21 (FTP) and TCP/7545 (TR069). Finally, using threat intelligence data from Spamhaus [13], we develop proxies for network hygiene and malware mitigation in ISP networks from DDoS amplifier data as well as infection data for non-Mirai (IoT and non-IoT) botnets (e.g., Qsnatch, Satori and a variety of Windows malware families).

We start by analyzing the infection rates for Mirai, followed by analyzing trends in the attack surface of ISPs and other types of infections in their networks. We then bring all our datasets together and statistically model the current Mirai-like infection levels and estimate the impact of three key explanatory factors: (i) the socioeconomic environment in which the ISP operates, (ii) the ISP's attack surface, as measured by open ports used by Mirai to propagate, and (iii) their security efforts to improve hygiene and combat other malware families. In short, our main contributions are as follows:

- We present a long-term empirical analysis of infection rates for Mirai malware across 342 global broadband ISPs and find that the number of subscribers explains 55% of the variance in the number of infections, but even ISPs with similar subscriber populations still vary by several orders of magnitude in the number of Mirai infections,
- For these ISPs, we also observe trends in their attack surface for Mirai and find that more than 75% have a reduced attack surface when examining Telnet and FTP ports over time, and that ISPs with worse initial hygiene have more Mirai-like malware infections on average at the end of the study.
- We construct an analytical model of factors that drive Mirai infection rates and then statistically model these factors, where we are able to explain 84% of the variance in the number of Mirai infections of ISPs.
- Remarkably, we find that reductions in attack surface have no significant impact on the Mirai infection rates. In contrast, security efforts to combat other types of botnets appear to also reduce Mirai malware infection levels. ISPs in the top 25% of non-IoT malware reduction (e.g., Windows malware) have 29% lower Mirai infection rates than ISPs in the bottom 25%. Similarly, ISPs in the top 25% of other IoT malware reduction have 48% lower Mirai infection rates than those in the bottom 25%.

Our results suggests that, ISPs can indeed play a significant role in combating IoT malware, but that they might be better off focusing their efforts on their general network hygiene and abuse mitigation measures (e.g., [11, 12]) rather than on IoT-specific countermeasures such as reducing the attack surface. This is consistent with ISPs including IoT malware in their existing quarantining and abuse handling processes. Prior work has found that notification and quarantining does, in fact, lead to remediation [14].

2. Related Work

Our study builds upon a wealth of academic literature and security industry reports that study various IoT malware strands, develop techniques to understand and track their behavior, and devise various mitigation and remediation strategies against their spread.

Understanding Mirai has received considerable attention within the literature [1, 15–17] given the record breaking DDoS attacks that were launched using the original Mirai botnet [1], and the fact that the release of its source code in 2016 has played a pivotal role in accelerating the trend of IoT device exploitation [18]. A great deal of the current IoT malware families are derivatives or at least employ similar techniques as those used by Mirai, for instance the brute forcing of a set of default administrative credentials to gain access to poorly secured IoT devices (cf. [19]). This attention is also partly driven by Mirai's relatively more aggressive and easily detectable scanning in comparison with other IoT malware which operate more stealthily. Mirai's source code revealed that its scanner component aggressively sends out random TCP SYN probes with a sequence number equal to the destination IP address being scanned to detect other devices which it may be able to infect. This scanning pattern is detectable in darknet traffic data. At the same time this 'fingerprint' appears to be the same for Mirai's closely related malware variants as well which have reused its original source code [6, 17]. This means that Mirai infections and those of its derived Mirai-like variants are among the more easily detectable and traceable IoT malware instances. A relative ease of tracking, has also lead to an abundance of commercial and non-commercial abuse feeds that track various Mirai components and make their data available to relevant and interested parties [13, 20, 21] including the subjects or our study: broadband ISPs. Furthermore, unlike other IoT malware, several studies have shown how Mirai-like infections may be effectively remediated by ISPs through notification and quarantining [6, 14]. As such, our study utilizes Mirai-like malware as a yard stick by which to understand and study the IoT malware landscape and their security problems from an ISP perspective, with the aim of empirically understanding the role they may play in mitigating the problem.

Well before the emergence of Mirai, in 2010, Cui and Stolfo demonstrated through scanning large portions of the IP space that a significant number of embedded devices (over 500k) have weak default factory-set administrative credentials, while also being publicly reachable on the Internet [22]. Thereby, predicting the rise of IoT Malware well before the emergence of infamous strands like Mirai.

Apart from in-depth studies of Mirai itself, a substantial part of the literature of course examines other IoT malware [23], for instance, Bashlite [24], Satori [25], Fbot, ADB.miner [26], Hajime [27], and VPNFilter [28], with each malware strand appearing to exhibit increasing levels of sophistication with the passing of time, yet also retaining or sharing certain features.

Necessary for, and often employed in, such studies are tools developed in literature that devise honeypot systems, or measurement techniques for detecting IoT malware via darknet traffic [29, 30] that help trace malware infections. For instance IoTPot [31] and Candy Jar [32] which may be used to emulate or use physical IoT devices to deliberately attract malware and study its behavior. Alternative high-interaction yet more scalable systems with physically attached IoT devices [33], or systems that support a large number of emulated device firmware [34] to capture a broader set of malware samples have also been developed without being restricted to certain firmware or device architectures.

Our work directly, as well as indirectly, relies on the various techniques for fingerprinting and tracking of infected devices developed in all such earlier work, first to collect our own data and second to understand, contextualize and analyze third party data that we utilize in our own study.

Another subject receiving considerable attention within the literature is that of devising techniques to enumerate and identify marketed IoT devices or those that get easily compromised, i.e., identifying their type, and manufacturer [10, 35–37], firmware [38], in addition to an overall evaluation of device security for more and less popular IoT devices observed in the wild [39, 40]. While some identification techniques may rely on the DNS behavior of IoT devices [41], others rely on their data flows [42, 43]. Such studies lay the foundation for alternative paths to solving the IoT security problem than the one which this study concerns it self with. Many of the currently discussed alternative solutions require the availability of such information to identify and hold certain entities accountable which these techniques will produce for instance. Some alternative paths propose the standardization and production of security labels for IoT products [44–46] as solutions for instance, others tackle the problem through the supply chain of IoT devices and propose to hold vendor and retailers liable [47], while legal scholars consider whether manufacturer liability may be an option [48]. Meanwhile, policy and regulation is emerging around the globe that requires minimum security standards in the manufacturing and marketing of IoT devices, for instance guarantees to provide software updates or patch vulnerabilities [49–51] and even proposal to make weak passwords illegal [52].

3. Methodology and Data

Analytical Model

For our study, we collect a variety of datasets that speak to potential driving factors of Mirai-like infections in broadband ISP networks. These factors are listed in the analytical model depicted in Figure 1. Among the factors, security efforts to reduce attack surface size, as well as security efforts to combat botnets in general, are our main factors of interest since we aim to understand whether such security practices may help in combating the spread of IoT malware such as Mirai and its variants.

Our analytical model stipulates that Mirai-like malware infections in broadband networks are partly driven by the following factors. First are institutional factors, for instance, characteristics of the country in which an ISP operates. Examples of such characteristics are a country's ICT infrastructure development, overall wealth, regulations, and other socioeconomic factors that directly and indirectly influence compromise levels at a macro scale view of the phenomenon. Next are characteristics of the ISP itself which influences how exposed an ISP is. In other words how large the attack surface of an ISP is. Network hygiene factors also play a role here. Next, are the security practices of the ISP and its customers, which may include controls and countermeasures to limit and reduce exposure and/or efforts to remediate security incidents. In other words, the overall joint security efforts and practices of the ISP and its customers. Note that our analytical model closely relates to a model of compromise underpinning a large body of empirical security research [53, 54].

As we cannot observe these factors directly, we approximate their effects on infection levels by proxy of several empirically observable and quantifiable variables. Examples of such proxy indicators are illustrated in Figure 1. For instance, ICT Development Index may be seen as a proxy for the institutional environment factor. The number of ISP subscribers may be seen as an approximation for ISP exposure or the size of its attack surface, and the numbers of various IoT and non-IoT malware infections in ISP networks as proxies for network hygiene. With respect to Mirai variants specifically, exposed ports, especially ones through which Mirai variants have been known to propagate, may also proxy attack surface. A reduction in the number of such exposed ports may indicate security effort intended to narrow the attack surface and combat Mirai-like malware. Similarly, reductions in the number of other malware infections could also proxy security effort, but efforts that are not specifically focused on combating Mirai(-like) infections but rather indicative of measures to improve general network hygiene and combat abuse through anti-abuse best practices such as notification and quarantining. Hence our use of the term 'generic' security effort which we will use throughout the paper to refer to such practices.

Notwithstanding the role of consumers in securing their IoT devices, our analysis considers the effects of their actions to be intertwined with that of the ISP. Since we have no data that directly speaks to the security efforts of individual device owners, we have no straight forward way to disentangle the effects of their actions from the security efforts of ISPs. Thus we interpret our data as a proxy

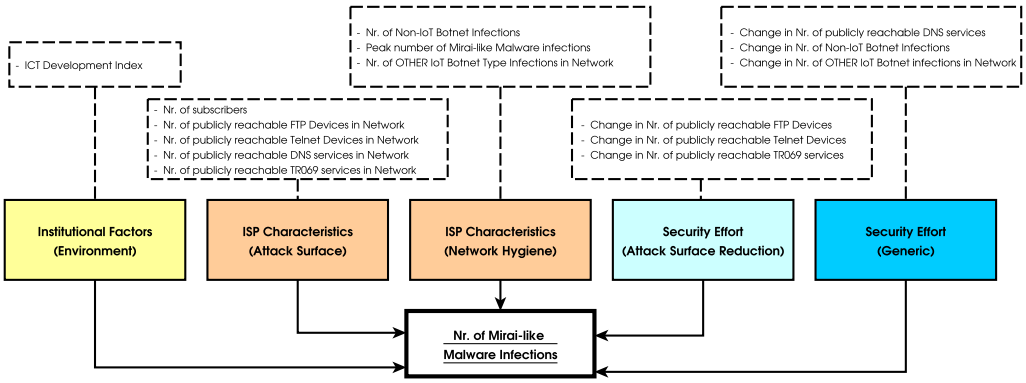


Figure 1: Analytical Model of Mirai-Like Infections in ISP Networks

indicator for the joint efforts of individual consumers and ISPs when our data speaks to factors of security effort. IoT malware is *de facto* remediated through the joint effort of the ISPs and their customers anyway. Consumers, who are unlikely to notice that their IoT devices have been compromised (since IoT malware typically do not impede the functioning of infected devices) are assisted by their ISPs who are in a better position to detect infections. ISPs are thus able to assist consumers by notifying them of the infection, and may even provide them with instructions on how to cleanup their device. In the end, it is upon the consumers to act upon the notification. ISPs rely on their cooperation to clean their network. Here however, an ISP may exert further influence over their customers by for instance (dis)allowing certain types of traffic to pass through their network, or by quarantining infected customers. ISPs may also impose certain default and more secure configurations on equipment which they provide to their customers, such as modems and routers. Nevertheless, it is through the joint effort of ISPs and their customers that Mirai-like malware may be remediated, and it is therefore reasonable to interpret our empirical data as approximating the joint security efforts of both individual consumers and their ISPs when their actions cannot be disentangled. Thus, in this paper ISP security effort refers to the joint security efforts of ISPs and their customers.

Later, in section 6, we employ our analytical model to build regressions over our data, and examine how strong the influence of various types of ISP security efforts have been on the number of Mirai-like infections within their networks. As stated before, we consider two types of effort: (i) narrowing the attack surface, and (ii) improvement of network hygiene through the remediation of other types of non-Mirai malware.

Collected Data

For this study we have collected several datasets that can be described as belonging to one of three data categories: (i) *network mapping*, (ii) *infection*, and (iii) *network scan* data. In short, these datasets allow us to empirically observe factors of our analytical model by proxy, and approximate their effects on infection levels. The first category of data allows us to map the networks of different broadband ISPs. The second and third categories have manifold purposes. Namely allowing us to indirectly observe ISP exposure, their network hygiene, and their security efforts. Table 1 provides an overview of our datasets along with corresponding time frames, and each dataset's original source. We first discuss what our collected datasets capture in more detail, how they relate to our analytical model, and what their purposes are in our study. Note that Figure 1 also conveys how these datasets relate to each of the factors of our analytical model.

Network Mapping Data. In our study we use commercial telecommunications marketing data from Telegeography (TG) to enumerate major broadband ISPs across the globe. This data contains detailed marketing information on telecommunication companies in 173 countries including their market shares, subscriber numbers and their deployed networking technologies. TG data covers both major and small broadband ISPs within its tracked list of country markets, covering on average 89% of each country's cumulative market share in terms of broadband.

This data has been used in prior empirical studies [54–57], most notably in studies by Asghari et al. to construct mappings of broadband ISP networks in 83 countries in their effort to examine ISP responses to spam botnets [56] in countries from which most of the global spam originates. This study leverages the network mapping and techniques constructed by these authors. We faithfully follow Asghari et al.'s methodology to combine the most recent version of the Telegeography data with BGP routing data [58], IP geolocation data from Maxmind [59], as well as CAIDA's AS to organization mapping [60] to construct an updated mapping for this study. In short, we manually segment and map the AS (Autonomous System) space of each country on to the broadband ISPs contained within the TG datatset, based on cross referencing the data contained within the TG, BGP routing, IP location data, and the network mapping data constructed in the aforementioned study, as well as the AS-to-organization data from CAIDA which is derived from WHOIS information.

We employ the resulting broadband network map of 83 countries to map our other datasets (discussed shortly hereafter) onto the networks of broadband ISPs identified by our mapping, as well as to filter our data to only those network segments pertaining to broadband ISP networks rather than other types of networks for instance hosting, educational, or governmental networks.

Infection Data. We make use of data from two different sources to track various malware infections and botnets across broadband networks: (i) A large darknet, primarily used to track Mirai infections; and (ii) data from Spamhaus, used to track a wider range of IoT and non-IoT

Table 1: Overview of datasets used in this study, periods covered by data in addition to the sources of the data

Period(s) Covered	Description	Source	Type
2015, 2019	Broadband ISP network mapping and statistical data	www.telegeography.com, [55]	Commercial/Marketing
2016-04 ↔ 2020-04	Probes of various TCP services (TCP/21,23,53,2323,7547)	Rapid7 Project Sonar (TCP Scans + National Exposure Scans)	ZMAP scans of IPv4 space
2016-01 ↔ 2020-05	IPs with Mirai-like infections	NICT (Japan)	Darknet data
2016-10 ↔ 2020-05	IPs with non-IoT botnet infections	Spamhaus (CBL)	Anti-Abuse / Threat-Intel Feed
2016-10 ↔ 2020-05	IPs with non-Mirai IoT malware infections	Spamhaus (CBL)	Anti-Abuse / Threat-Intel Feed

malware infections worldwide. Together, these datasets render indirect observations of ISP network hygiene, their security efforts, as well as an indirect view of attacker behavior as certain botnets becomes larger, shrink, or get replaced by competing strands.

Darknets, which are unused but reachable IP address ranges utilized to passively monitor incoming traffic, capture a variety of packets that in principle should not be destined for their IP ranges. As such, packets captured by darknets indicate symptoms, either network misconfigurations, or malicious activity such as probing for vulnerable devices which is typical of worms and other malicious automated Internet scanning software.

Our darknet data, which we use to track Mirai infections with, spans the duration of 2016 to May 2020. As such, it captures the entire lifespan of Mirai infections up to current day levels of its derived variants. This data has been made available to us via NICT (Japan) and collected from a darknet with approximately 300,000 IP addresses operated from 40 different networks across 15 countries.

To track Mirai infections using this data, we leverage techniques from prior work which have also primarily employed darknet data for this purpose [1, 14]. Mirai's scanner component probes random IP addresses typically also probing darknet ranges and thereby leaving its fingerprint behind in darknet traffic. As such, we process our darknet data by searching for Mirai's distinct scanning signature: TCP SYN packets with sequence numbers equal to the packet's destination IP. A notable difference in our approach with that of the previous studies is that our matching criterion does not specify particular ports to which packets are destined, whereas those previous studies searched for only commonly probed ports by the original Mirai malware. We use a less restrictive search criterion to account for the fact that later Mirai variants, were found to have expanded the set of ports for which they probe [26]. As such, we track a broader set of '*Mirai-like*' malware variants with arguably little impact on the accuracy of our data. First, because of the extremely low probability of TCP SYN probes having sequence numbers equal to an IP address destination. And second, due to cross comparisons that we undertake with other Mirai tracking data from a second data source discussed later in section 4. We discuss this independent data, which has been collected by Spamhaus, next.

In conjunction with our darknet data, we also collect data from the Spamhaus Composite Blocking List (CBL). CBL captures data on botnet infrastructures and infected machines sending spam, acting as open relays, or participating in DDoS attacks [13]. It lists IP addresses exhibiting signs of infection and captures data on both IoT and non-IoT botnets which Spamhaus labels accordingly, and it includes labels indicating the type of malware identified to be behind each observed infection.

We have collected CBL data since October 2016 up to and including May 2020 and use its non-IoT botnet infection portion to primarily track non-IoT malware infections as a proxy for understanding broadband ISP security practices with respect to other types of botnets and the relative hygiene of their network. Spamhaus tracks a variety of non-IoT (primarily Windows-based) malware, including but not limited to things such as Zeus, Conficker, Gamarue, Gozi, Zeroaccess, Kelihos, Ramnit, Cutwail, and a variety of at least 40 different malware families which they include labels for in their data.

At the same time, we use the IoT infection portion of the CBL data to track various IoT malware infections, and ISP responses. CBL data tracks IoT malware such as Mirai, Wopbot, Qsnatch, Satori, as well a range of other generically labeled IoT malware. This data may also be leveraged to understand competition between various IoT malware strands as certain infections grow and replace other IoT malware infections within ISP networks.

Finally, as the CBL also contains independently collected data on Mirai infections, we use its Mirai infection portion of the data for cross examination with our primary Mirai infection data source.

Network Scan Data. We also leverage data on periodic TCP SYN probes of the entire IPv4 address space carried out by Rapid7 as part of their 'Project Sonar' [61]. This data is freely available through Rapid7's open data platform.

We use this to quantify the attack surface size of an ISP network over time with respect to services and ports that are commonly misused by Mirai variant bots to propagate. Security literature has indeed demonstrated that concentrations of openly accessible vulnerable services, for instance insecure Telnet and FTP services, correlate with high concentrations of abuse within networks [53] establishing a link between the attack surface and the amount of security incidents visible within a network. As such, we leverage the collected network scan data to examine the effects of countermeasures aimed at narrowing the attack surface of each broadband network. Think for instance of countermeasures that temporarily or by default and thus more permanently block ports used by Mirai variants to propagate. Such countermeasures have been adopted by certain ISPs in the past on by blocking port 25 as a countermeasure to prevent the sending of spam emails. Countermeasures could of course be implemented at the ISP network level, or customer network level through security efforts to push default configurations to customer routers that block typical Mirai propagation ports by default. Irrespective of where such countermeasures are implemented, our network scan data only allows us to observe the effective change to size of an

ISP's attack surface with respect to common propagation ports. And therefore we interpret an empirically observed reduction of the attack surface size as a proxy for such countermeasures without making any assumption of how and where countermeasures may have been implemented.

The subsets of Project Sonar's data relevant to our study are released by Rapid7 under the headings of 'National Exposure Scans' [62] and 'TCP Scans' [63] which we have collected over the duration of 2016 to 2020. The former provides irregular panel data on scans of common TCP services within IPv4 space during 2016-2018, and the latter more regular monthly data on similar measurements over the span of 2017-2020 (May). Combined, these sets provide a rough historical record of the volume and evolution of common TCP services on the Internet over a 4.5 year time span.

From this data we analyze only a subset of the common TCP services probed by Rapid7, namely: TCP/23+2323 (Telnet), TCP/21 (FTP), and TCP/7547 (TR-069); services commonly targeted by many IoT malware strands including Mirai variants, in addition to probe data on TCP/53 (DNS) services, which indicate the presence of open DNS resolvers potentially exploitable as DDoS attack amplifiers.

4. Mirai-like Infections in Broadband ISPs

We begin our analysis with a cross examination of our Mirai-like malware infection data from our two data sources: (1) the darknet data, and (2) the subset of Mirai data from CBL. The results of our cross examination are illustrated in Figure 2 which plots the number of unique IP addresses detected to have been infected with some Mirai variant on a daily basis. While we do see some noticeable differences between the measurements based on each dataset, most notably in the volume of captured infections, we see relatively consistent results in peaks and trends. The observed trends are also consistent with prior work on tracking and understand Mirai [1], which in our study is extended to a period of more than 4 years.

The larger volume of infections captured by the Spamhaus CBL data may be attributed to its larger operation compared to the relatively smaller darknet dataset. Some of the observed differences in volumes may also be attributed to the documented fact that Spamhaus accumulates its data over short intervals only to automatically remove stale infection data when an expiry interval is reached, thus resulting in a certain level of out-dated infection data in its dataset on a daily basis. Nevertheless, it appears that both of our datasets show relatively consistent peaks and a steady decline in Mirai infections to a long tail of lingering infections, which providers confidence that the collected data is robust enough to capture Mirai-like infection dynamics and to examine broadband ISP responses to Mirai-like malware.

For two reasons we have opted to base our analysis of Mirai variant infections solely on the darknet data. First, that it has a more transparent collection and processing methodology, and second is its more consistent effort in tracking and collection of data. Figure 2 clearly illustrates that. The Spamhaus dataset misses two significant time spans of the growth and decay phases of Mirai-like malware infections, first in 2016, and next in transitioning

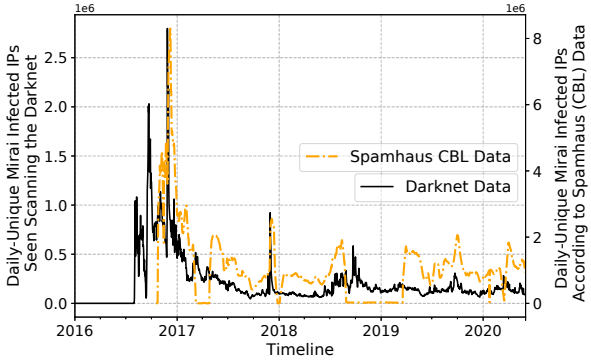


Figure 2: Mirai family infection numbers.

from 2018 to 2019. Note that the missing data is not due to lapses in our CBL data collection efforts, but rather it seems due to Spamhaus not having tracked Mirai during these phases. Our raw data files for these periods do not contain Mirai related entries during these periods.

Note that in all our analyses, we process infection data on a daily basis to limit the impact of IP churn and to avoid over counting the number of infections. This choice is grounded in the fact that typically IPv4 addresses churn at a longer frequency. Therefore in Figure 2, as well as other subsequent analysis, we typically report the number of daily unique IP addresses we have seen to be scanning the darknet with a Mirai signature.

The observed decline in Mirai infections, especially since it largest peak in late 2016, begs the question of the role of ISPs and whether/which security efforts may have contributed to this decline. Other factors, such as a rise in competing IoT malware, could also be behind such declines, yet our main interest is to understand to what extent the decline may be explained through ISPs actions. Note that in this context we may also place less emphasis on other factors such as manufacturers improving the security of their largely marketed insecure IoT devices. In many cases, consumers are left with no option to change default passwords on their IoT devices or update their firmware, if available at all, even if they technically know how to do so.

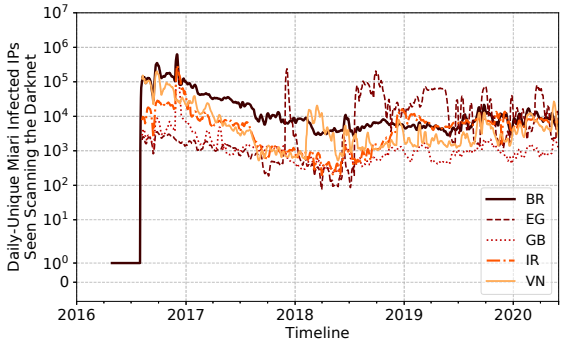


Figure 3: Top 5 countries with highest Mirai infection peaks

To continue our analysis we next map each infection to its respective location using historical MaxMind IP geolocation data and compare infection volumes across different countries. At the level of countries, we generally see comparable and consistent trends in the peaks and a decline in infections comparable to the one observed at the global level, albeit with significant differences across countries. Figure 3 for example, plots the infection data

for the five countries with the highest infection peaks: *Brazil*, *Egypt*, *Great Britain*, *Iran* and *Vietnam*. Note the logarithmic scale of the plot's y-axis applied for ease of comparison and an improved readability. While the infection peaks within most countries appear to have occurred in late 2016 according to our data, within a few months of Mirai's emergence, we see peak infection levels at much later stages in several countries. For example, the infections in Egypt have peaked almost a full year later towards the end of 2017. This may be explained through newer Mirai variants having been the driving force behind these infections. For each country, peak infection levels, average infection numbers (taken over the timeline of our data), as well as the overall number of lingering infections within each country towards the end of our study, vary significantly exhibiting several order of magnitude difference even among the top five countries that experienced comparable infection peak magnitudes around similar times. These observed differences are partially explainable through institutional macro economic factors such as the ICT development Index of the country suggesting that institutional factors indeed influence the differences seen among countries and the ISPs operating in those countries. Prior work has already shown that different countries also exhibit substantial differences in the composition of IoT devices available within their markets which could also explain the differences that we observe here [10].

Next, we map our Mirai infection data, to corresponding Autonomous Systems from which infections originate. We do so by using historical BGP routing data and the 'pyasn' python package [58]. We subsequently use our constructed broadband ISP network mapping (Section 3) to map our infection data to corresponding broadband ISP networks of ISPs within the 83 countries which we have mapped. This results in detailed time series data on Mirai infection levels across 342 broadband ISPs which again show markedly different levels of infection. In later sections we illustrate and discuss some specific examples (e.g., see Figure 5 discussed in Section 5)

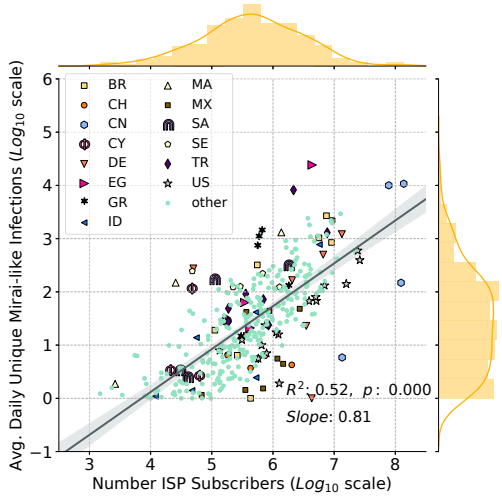


Figure 4: Average Mirai family infection rates across Broadband ISPs (2018-2020, second half of study timeline).

To compare ISPs, we first cross examine their average Mirai infection levels. Given the large decline in global infection levels since the malware family's initial peak however, we draw multiple comparisons based on averaging infection counts across shorter time spans of our data. Figure 4 which for instance plots the average number of daily unique Mirai infections against each ISP's respective number of subscribers, is based on averaging the infection counts from the second half of our data timeline (i.e., from 2018 onward) where global infection numbers appear to have more or less declined to a stable level. The countries depicted in the figure are those with the top 15 ISPs showing the most deviation from the regression line.

Figure 4 not only clearly shows that despite Mirai variants reaching relatively steady global infection levels, there are still stark differences among entire countries, which was also observed previously in Figure 3, but also significant differences among the ISPs within the same country, as well as ISPs with comparable subscriber numbers in different countries. At first glance, the figure demonstrates that for a large part, differences in infection levels across broadband ISPs may be explained by differences in their respective subscriber numbers. This may be observable through the linear regression results depicted in the figure which shows an R^2 value of 0.55. This suggests that 55% of the variation among ISPs infection levels may be purely explained through differences in their subscriber numbers. Differences in subscriber numbers may also be viewed as a proxy for differences in the potential attack surface of each ISP.

Note that similar comparisons of average infection levels across the first half of our data, including periods of peak global Mirai-like infections, show closely similar results. For all comparisons among ISPs that we have drawn subscriber numbers strongly explain the differences among ISP Mirai-like malware infection levels.

Nevertheless, Figure 4 also shows that significant differences exist, at times several orders of magnitude large, among ISPs of comparable size around the regression line which remain to be explained. In fact 45% of the variations observed in Figure 4 remain to be explained through other factors. Some of the observed differences may relate to differences in ISP security practices or their specific security countermeasures, but as stated before in the introduction, we know very little about what ISP security practices may be effective to stem the tide of IoT Malware.

In the following section we examine the relationship between Mirai variant infection levels across ISPs and two specific countermeasures: (i) reducing an ISP's attack surface size by for instance closing ports through which Mirai variants have been known to propagate, and (ii) security efforts to combat other forms of malware.

5. ISP Security Efforts

The differences that we observe among ISPs with respect to their Mirai infection levels in Figure 4 bring up the question of whether and to what extent these differences are driven by different levels of security effort.

To further examine why such differences exist, we take a closer look at our other datasets which could indicate different levels of broadband network operators security efforts. This data allows us to better understand how indicators of hygiene have evolved for each ISP since the emergence of Mirai and its variants, and how

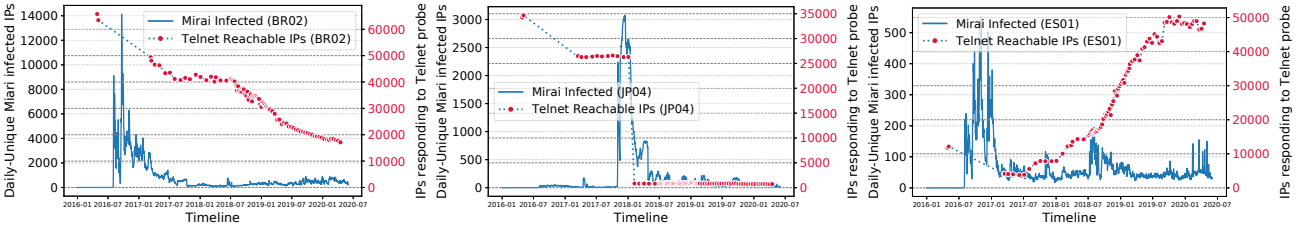


Figure 5: Number of telnet reachable devices in broadband ISP networks plotted along with the number of Mirai family infections seen in their network. (a), (b) and (c) are illustrative examples of ISPs from our data all showing declining Mirai-like infection numbers. (a) and (b) show decreasing telnet accessible services whereas (c) shows increasing numbers. (b) suggests implementation of strong countermeasure by ISP.

security efforts – seen through the lens of how hygiene has evolved over time – may play a role in inhibiting Mirai-like infection numbers within ISP networks.

In what follows we first examine our longitudinal data on reachable TCP services within ISP networks, for instance the number of publicly reachable Telnet and FTP services which are both propagation vectors for Mirai malware variants. We examine whether their numbers have dropped and how the changes to their numbers relate to Mirai infections. We interpret a drop in the number of such services as a narrowing of the attack surface and a proxy indicator for ISP security effort. Of course our data does not allow us to directly observe exactly who, i.e., the ISP or its customers, are driving the reduction. Nevertheless we do empirically observe the manifestation of their joint effort.

In the subsection thereafter, we examine the inhibiting effect of alternative ISP security efforts observed through the lens of a second set of data, namely our data on non-IoT and other IoT botnet infections within an ISP's network. We explore how their changing numbers, similarly signal about the joint security efforts of ISPs and their customers to cleanup other types of malware, and how this relates to Mirai infection numbers. Our rationale as before is that a reduction in the number of other botnet infections, also indicate ISP security effort by proxy.

Reducing the Attack Surface of Mirai

Our collected Rapid7 data (see Section 3) allow us to reconstruct a timeline of how the number of a set of reachable TCP services has evolved over time within each broadband ISP network. Specifically, we have processed the subsets of Rapid7's data on Telnet (TCP/23 and TCP/2323), FTP (TCP/21), and TR069 (TCP/7457) services which have been routinely targeted and exploited by Mirai variants as well as other IoT malware as vectors to compromise IoT devices and propagate. While Telnet and FTP are respectively protocols for text based communication and file transfer, TR069 is a protocol for the remote management of CPEs (Customer Premises Equipment), for instance routers which ISPs can remotely configure using this protocol. TR069 was first exploited in 2016 by a Mirai variant in the now well known example of Deutsche Telekom where large parts of its network were disrupted due the malware attempting to exploit this service on their CPEs [64]. Security advice with respect to this protocol is for ISPs to close this port to anyone outside of their own network for instance.

With respect to this data it is important to note that Mirai-like malware have been known to terminate services like Telnet and FTP upon infecting victim IoT devices and closing the ports on which these services communicate. Therefore, reductions in the number of such services in ISP networks may instead be driven by the Mirai malware itself rather than the security efforts of ISPs. If that were the case however, it should by necessity be accompanied by an increase in the number of Mirai infections within a network and therefore distinguishable as such in conjunction with infection data. For this reason, we interpret changes in the numbers of these exploited services in conjunction with changes in the numbers of Mirai-like infections to avoid misinterpreting reductions as the result of ISP security efforts.

We first provide a few illustrative examples of our data in Figure 5. The figure depicts three ISPs for which their Mirai infection numbers are contrasted against the number of publicly accessible Telnet services in their networks over the timeline of our study. For two of the ISPs, we observe a decline in the number of Telnet services, which for the most part of the timeline are NOT accompanied with increasing Mirai-like infections. This indicates that the decline in telnet services is not driven by more and more Mirai infections but rather by security efforts. For the first two depicted ISPs we also see that infections and the number of Telnet services change in the same declining direction. The third ISP on the other hand, exhibits a decline in infections despite an apparent increase in the number of publicly accessible telnet services within its network. This goes to show that our indicators are not to be interpreted as causal drivers of infection but rather imperfect approximations for the security efforts of providers which given the multi-causal nature of the phenomenon that we are investigating may or may not influence Mirai infection numbers. So at best we are seeing correlations between reduction in number of telnet services and a reduction in Mirai-like malware infections.

In these illustrative examples, the second ISP (Figure 5b) paints a distinctively interesting an unparalleled picture noteworthy of mentioning, as the decline in its infections is almost perfectly aligned with that of the decline in its telnet service statistics. Our detailed investigation of the timeline of events around the sudden drop in infections in the networks of this Japanese ISP revealed a series of informative events. We have been further informed of these details by the National Institute of Information and Communications Technology in Japan (NICT).

In October 2017, NICT had noticed an increase in Mirai-like infections scanning its darknet ranges which it traced to infected IP addresses located within the networks of the Japanese ISP from Figure 5b. Analysis of malware

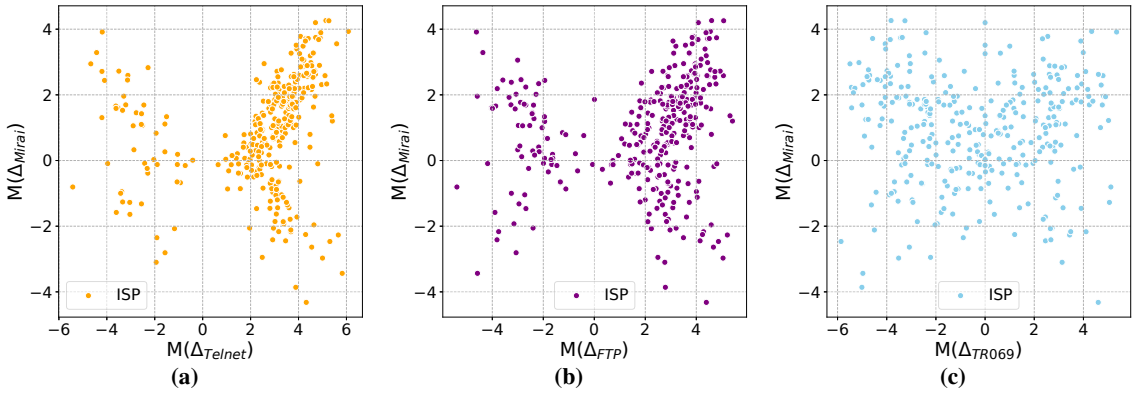


Figure 6: Change in number of reachable Telnet (a), FTP (b), and TR069 (c) services in ISP versus change in Mirai variant infections. $\Delta_{Telnet}, \Delta_{FTP}, \Delta_{DNS}, \Delta_{Mirai}$ calculated by averaging and subtracting the first and second half our timeline of measurements. Positive Δ values represent a decrease in the corresponding value since the start of our measurements and negative values an increase. Note the logarithmic transformation $M(...)$ applied to the scales: $M(x)$ is $Log_{10}^{\|x\|}$ if $x > 0$, $-Log_{10}^{\|x\|}$ if $x < 0$, 0 otherwise.

samples captured via honeypots then point to a Mirai variant exploiting a specific vulnerability (`CVE-2014-8361`). After notifications from NICT later in November 2017, analysis of the affected IP addresses by the ISP unearthed a substantial number of unpatched devices which lacked an automatic update mechanism from a popular manufacturer of IoT devices in Japan. By mid December 2019 multiple organizations including the manufacturer itself had issued press releases alerting the public about the affected devices including the Japanese CERT [65, 66]. By February 2018, NICT darknet data show the number of Mirai infections to have dramatically decreased. The Mirai variant behind this incident was dubbed Okiru [67]. While we have no confirmed information of the specific countermeasures implemented by the ISP, the dramatic reduction in number of telnet accessible devices coinciding with the large fall in Mirai infections suggest strict steps taken by the ISP to restrict access to telnet services in response to the malware. These restrictions may have remained in place to date according to our data.

The example of the Japanese ISP is of course not a common occurrence in our data. Instead, our data suggests that the relationship between infection numbers and the numbers of devices responding to the potentially exploitable protocols (Telnet, FTP, TR069) much more closely resembles the other two examples depicted in Figure 5 with the some fewer ISP cases showing growing trends either in terms of infections, the attack surface, or both. In other words, while there do appear to be indicators of narrowed attack surface correlating with reduced Mirai-like infections for a large portion of the ISPs, the relationship is also more complex. Figure 6 depicts this relationship in more condensed form for all ISPs in our data by plotting the change in Mirai infections against reductions in their attack surface. Note the logarithmic-transformed scale. Changes are depicted through delta (Δ) variables, and calculated by averaging and then subtracting the first and second half of our data for each ISP represented as one point in these plots. Note that positive Δs indicate a reduction in numbers and negative values an increase. As such, positive Δs may be interpreted as improvements and negative Δs as a turn for the worse.

Based on the analysis depicted in Figure 6 we see that the majority of ISPs in our data improve both in terms of infection numbers and attack surface (top right quadrant of the plots) with a smaller number falling in quadrants of the plot that indicate either a worsening of exposure, infections, or both. These results weakly suggests that security efforts to reduce exposure through reducing attack surface may have an impact on reducing Mirai-like infections as the variables correlate in the expected direction for the majority of ISPs. For a smaller number of ISPs we do not see a clear correlation, which is to be expected as the phenomenon is driven by multiple causal factors but this also leaves it unclear as to how strong the effect may be especially when considering that other factors also contribute to changes in infections levels. Later in Section 6 we use the $M(\Delta_*)$ values plotted in the figure as approximations for this type of attack surface reducing security effort, and construct statistical models of our data to estimate their effects along with the other factors from our analytical model.

General ISP Security Efforts

Beyond taking actions to reduce the attack surface for Mirai-like malware, ISPs can take security actions that are not specifically targeted at Mirai but rather aimed at improving the hygiene of their network in other respects, or actions aimed at remediating other types of infections. These efforts may also have inhibiting effects on Mirai as some security practices, for instance notifying customers of infections or quarantining, are much more 'general' and effective against a broader set of infections. We refer to such security practices as general ISP security efforts.

These general efforts may be observed indirectly through the lense of our Spamhaus CBL data (see Section 3) for instance. Data from this feed for instance allows us to track changes in the number of both non-IoT, as well as IoT botnets other than the Mirai family within each broadband ISP's network.

The CBL feed contains data on more than 50 malware strands such as Zeus, Conficker, Gamarue, Gozi, Zeroaccess, Kelihos and Ramnit to name a few examples, mostly windows based malware and thus related to *non-IoT* botnets, each individually tracked and labeled accordingly by Spamhaus. In processing the CBL data we combine all non-IoT labeled infection data into a single larger encompassing '*non-IoT*' botnet category. We then map infections to their respective networks. We then count the number of unique IP addresses infected by any of the botnet families falling into the larger umbrella category

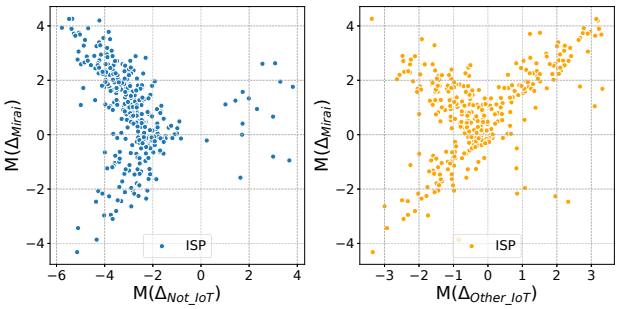
within each ISP's network to construct detailed time series data of non-IoT botnet infection volumes on a daily basis.

The same CBL data feed also contains data on IoT malware infections unrelated to Mirai, for instance `Qsnatch`, `Wopbot`, as well as several generically labeled IoT malware stands detected by Spamhaus, which we similarly use to track and count '*other IoT*' botnet infections within broadband networks. As before, we combine all IoT botnet data of the CBL feed (excluding its Mirai-like malware data), into a larger category of '*Other IoT*' botnets when processing this data.

As we have done before when examining the attack surface, reductions to the number of infections; whether from non-IoT or other-IoT category, are interpreted as proxy indicators for increased security efforts to combat other types of infections. We are of course not able to discern exactly what security measures have been taken based on our data and observe their manifestation.

In the same vein as Figure 6, Figure 7 contrasts the observed changes in infection numbers, this time plotting the change in number of non-IoT (7a) or other-IoT infections (7b), against observed changes in the number of Mirai infections per ISP. The results of this analysis are supposed to help us better understand the relationship between general ISP security efforts to combat other types of malware and combating IoT malware such as Mirai. Here, for non-IoT malware, we see patterns that closely resemble ones seen before in Figure 6 when examining changes in attack surface. We see the majority of ISPs exhibiting a reduction in the number of non-IoT malware infections which has a moderate correlation with a reduction in the number of Mirai variant infections. This potentially indicates efforts to combat other IoT infections to also have an inhibiting effect on Mirai infection levels. In the case of non-IoT malware infections however, the pattern breaks. We see a large number of ISPs exhibiting a negative reduction (an increase) in the number of non-IoT malware infections in their networks while surprisingly at the same time a large portion also exhibiting reductions in Mirai variant infections. These patterns are difficult to explain purely based on these plots since what we see may be the result of several conflated factors.

In addition to the CBL data a small subset of the Rapid7 scan data not discussed previously, namely its data on the changing numbers of open DNS resolvers (TCP/53) in ISP networks, may also be used as an indicator for ISP security efforts to improve network hygiene in conjunction with the CBL data. Open DNS resolvers have been used in previous studies to signify network hygiene [68]. Established initiatives such as the open resolver project (www.openresolverproject.org) and organizations like shadowserver https://scan.shadowserver.org/dns/ have also attempted to remediate open DNS resolvers globally through notifying network operators in the past. And while such resolvers are not directly targeted as an infection vector by Mirai malware per se but rather as a means to amplify DDoS attacks, reductions in their volume could similarly be used to signify ISP security efforts, but of the more geneal type which we have distinguished from direct efforts to reduce the attack surface for Mirai malware. Our results with respect to analysing this data, briefly stated, show similar patterns to those observed in Figure 6 when correlating changes to the number of open DNS resolvers and Mirai infections.



(a) Non-IoT Infections **(b)** Other IoT Infections

Figure 7: Changes in number of (a) non-IoT and (b) other non-Mirai IoT infections in ISP networks versus change in number of Mirai family infections. Δ variables calculated by averaging and subtracting the first and second half of our timeline of measurements. Positive Δs signify reductions in numbers. Note the logarithmic transformation $M(...)$ applied to the scales. $M(x)$ is $Log_{10}^{\|x\|}$ if $x > 0$, $-Log_{10}^{\|x\|}$ if $x < 0$, 0 otherwise.

To better understand how our factors of interest relate to Mirai infection levels, we argue that it is better to construct statistical models of our data through which factor effects may be disentangled and ceteris paribus inferences drawn about their effects. This is due to the multi-causal nature of the phenomenon we are trying to understand. Among the factors, ISP security efforts to reduce Mirai's attack surface, and efforts to improve network hygiene or remediate other types of infections (general security efforts) are our main factors of interest. Our goal is therefore to better understand what security countermeasures ISPs can implement to combat the tide of IoT malware.

6. Modeling Infection Levels

To better understand how and to what extent our factors of interest explain different Mirai infection levels across ISPs, we construct several linear regression models on top of our data. Overall, the factors that we model include institutional factors characterizing the surrounding environment of an ISP, its attack surface and network hygiene, its security efforts to reduce attack surface size, and its security efforts to improve hygiene and combat other types of infections within its networks as potentially influencing factors (Figure 1). Recall that the latter two factors are the main factors of interest which help us understand what other roles ISPs can play to combat IoT malware (than for instance quarantining infected customer networks [6]).

The regression model we built, model the average number of infections over the second half of our study timeline (2018-2020) for each ISP. In other words, their dependent variables are these average infection statistics. The models help answer the following question: To what extent are the differences in average Mirai infection levels related to and explained by differences in the characteristics of ISPs captured by our factors of interest.

We construct our models via a standard step-wise process, starting with a naive base line model including only a constant as an explanatory variable. Note that the naive baseline model ($Model_0$) has been omitted from the table in the interest of space. Next, models 1 to 5 are

Table 2: Linear Regression Models

		Dependent variable:				
		Current Nr. Mirai-Like Infections in ISP (Log_{10}) (Avg over 2018-2020)				
Factors ⇓	**Models** ↦ Proxy Variables ⇓	(1)	(2)	(3)	(4)	(5)
	Constant	1.82*** (0.21)	−1.88*** (0.26)	−1.23*** (0.23)	−1.35*** (0.24)	−0.81*** (0.23)
Institutional	**ICT Development Index** (of Country where ISP operates)	−0.04 (0.03)	**−0.06*** (0.02)	**0.07*** (0.02)	**0.07*** (0.02)	**0.04**** (0.02)
Exposure (ISP Characteristics)	**Nr. ISP Subscribers** (Log_{10})		**0.43*** (0.06)	0.05 (0.05)	0.06 (0.05)	−0.01 (0.05)
	Init. Nr. Reachable FTP IPs (TCP/21) (Log_{10}) (Avg over 2016-2018)		−0.01 (0.08)	0.05 (0.06)	0.09 (0.06)	0.07 (0.06)
	Init. Nr. Reachable Telnet IPs (TCP/23) (Log_{10}) (Avg over 2016-2018)		**0.23*** (0.08)	−0.04 (0.06)	−0.05 (0.06)	−0.01 (0.06)
	Init. Nr. Reachable Telnet IPs (TCP/2323) (Log_{10}) (Avg over 2016-2018)		0.03 (0.05)	0.01 (0.04)	−0.02 (0.04)	−0.02 (0.04)
	Init. Nr. DNS Responding IPs (TCP/53) (Log_{10}) (Avg over 2016-2018)		0.10* (0.06)	0.04 (0.05)	0.09* (0.05)	0.07 (0.05)
	Init. Nr. TR069 Responding IPs (TCP/7547)(Log_{10}) (Avg over 2016-2018)		**0.06**** (0.02)	0.02 (0.02)	0.02 (0.02)	0.002 (0.02)
Network Hygiene (ISP Char.)	**Peak Nr. Mirai-like Infections** (Log_{10}) (Avg over 2016-2018)			**0.23*** (0.04)	**0.23*** (0.04)	**0.19*** (0.04)
	Init. Nr. *Non-IoT* Botnet Infections (Log_{10}) (Avg over 2016-2018)			**0.26*** (0.06)	**0.24*** (0.06)	**0.17*** (0.05)
	Init. Nr. *Other* IoT Botnet Infections (Log_{10}) (Avg over 2016-2018)			**0.28*** (0.06)	**0.28*** (0.06)	**0.47*** (0.06)
Security Effort (Exposure Reduction)	**Reduction in Nr. FTP Reachable IPs** $M(\Delta_{FTP})$				−0.01 (0.01)	−0.01 (0.01)
	Reduction in Nr. Telnet Reachable IPs $M(\Delta_{Telnet})$				0.005 (0.01)	−0.01 (0.01)
	Reduction in Nr. Telnet$_{TCP/2323}$ Reachable IPs $M(\Delta_{Telnet_{2323}})$				0.003 (0.02)	0.01 (0.01)
	Reduction in Nr. TR069 Responding IPs $M(\Delta_{TR069})$				0.002 (0.01)	0.000 (0.01)
Generic. Sec. Effort	**Reduction in Nr. DNS Responding IPs** $M(\Delta_{DNS})$					**−0.02**** (0.01)
	Reduction in Nr. *Non-IoT* Infections $M(\Delta_{Non\_IoT})$					**−0.07***** (0.02)
	Reduction in Nr. *Other* IoT Botnet Infections $M(\Delta_{Other\_IoT})$					**−0.14***** (0.02)
	Observations	318	318	318	318	318
	R^2	0.01	0.76	0.79	0.80	0.84
	Adjusted R^2	0.004	0.75	0.78	0.79	0.83
	Residual Std. Error	0.85 (df = 316)	0.42 (df = 309)	0.39 (df = 307)	0.39 (df = 302)	0.35 (df = 300)
	F Statistic	2.17 (df = 1; 316)	121.01*** (df = 8; 309)	115.95*** (df = 10; 307)	78.48*** (df = 15; 302)	92.42*** (df = 17; 300)

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

$M(x)$ is $Log_{10}^{\|x\|}$ if $x > 0$, $-Log_{10}^{\|x\|}$ if $x < 0$, 0 otherwise

Δ_X is $Avg_{2016/2018}(X) - Avg_{2018/2020}(X)$

constructed by the step wise inclusion of an additional factor to each previous model, and $Model_5$ constitutes our complete model. Table 2 illustrates the models that we have built and summarizes their results in terms of coefficients that capture the directions and effect size of each factor on the dependent variable.

We will examine our complete model ($model_5$) and its results shortly. But first, we take a closer look at our more general findings by examining the goodness-of-fit statistics reported in Table 2 below each model, mainly the R^2 values. We examine what they suggest about the models themselves, in addition to our individual factors of interest (See the Appendix section for additional GoF indicators of our model which suggest that a linear regression model is reasonably suited to model our data). As our models include more factors the increasing R^2 values reported in the table indicate that our models are successively better capable of explaining the variance observed among ISPs in terms of their different Mirai infection levels. We find that $Model_5$ explains 84% of the observed variation.

Next by comparing each model with its preceding model we may observe how much of the variation in infection levels may be explained by the inclusion of each additional factor, or more precisely stated, by the inclusion of the collection of proxy variables that approximate the factor. The largest explanatory factor by far seems to be ISP *exposure* i.e., its attack surface size. This may be observed by comparing $Model_2$, for which the R^2 statistic

shows an increase of 75% by the inclusion of ISP exposure factors, against $Model_1$ in which only institutional factors have been taken into account. The increase of 75% in R^2 values from $Model_1$ to $Model_2$ is the largest increase among the successive models. The remaining step-wise additions of factors contribute only an additional 8% to explaining the observed variation among ISPs ($Model_2$ vs $Model_5$). We also see that 16% of variation remains unexplained. These results suggests that although not all of the variation in our data may be explained purely through the considered factors (and therefore that other factors may also be quite influential), nevertheless, our complete model reasonably fits our data and that we may reasonably explain a large part of the variation among ISPs through our set of explanatory factors.

In addition, we may observe that our models are reasonably robust through cross examining the reported model coefficients as more factors are added in succession. In this respect, we see three factors that consistently exhibit a significant effect on Mirai infection levels, namely, institutional factors, Initial network hygiene conditions, and general ISP security efforts to combat other types of malware. We see the proxy variables approximating each factor, to show both significant effects, as well as a consistent approximation of the direction in which they affect the dependent variable as more factors are taken into account. These are respectively observable via the reported significance level and signs of the reported model coefficients. Positive coefficient signs here signify an increasing effect where as negative signs a decreasing one. Note that exposure, the factor which we discussed as the factor with the largest explanatory power, looses its significance and consistency in terms of the reported coefficients as more factors are added to the models. This does not invalidate our previous statement regarding its explanatory power, as the explanatory effect still exists but rather the effect of its proxy indicators over shadowed and carried by the other significant proxy variables which co-vary with the variables that approximate exposure.

So what have we learned from our complete model: i.e., $Model_5$? To answer this question we interpret and explain what its significant factors, proxy variables, and their corresponding coefficients entail. Starting from the institutional factors, we observe a significant positive relation between ISPs operating in environments of relatively higher ICT infrastructure development, and having more Mirai-like infections. This comes as no surprise as it is more likely that IoT devices, among them vulnerable devices, have a higher market penetration in such environments. The effect is reported to have a positive coefficient value of 0.04 in $Model_5$ which its effect size may be interpreted as follows. When all else is constant, moving towards greater ICT development by one unit in the scale of the variable we should also expect a 0.04 increase to the dependent variable scale. Since our dependent variable is base-10 logarithmic scaled, this means that we should expect the average number of Mirai-like malware infections to be multiplied by $10^{0.04} = 1.09$, which is a 9% increase. In other words if two ISPs had the same number of subscribers, exposed services, etc, etc, and their only difference were that they operated in different institutional environments that differ in their ICT development by 1 unit (as measured by the ICT development index), we expect one to have 9% more Mirai infection on average than the other.

We also see a positive correlation between poor initial network hygiene conditions, i.e., having greater average numbers of non-IoT, other-IoT, or peak Mirai-like malware infections during the first half of our study period, and ISPs having more Mirai infections towards the end of our study. We may interpret these effects as "bad" network hygiene leading to more infections. In other words, when all else is the same, as in similar institutional environment, similar exposure, etc, etc, ISPs with worse initial hygiene have more Mirai family infections on average at the end of the study. To understand the effect sizes here, coefficient values need to be interpreted differently due to the logarithmic transformation applied to the independent variables. We illustrate by example. Consider the 'initial Non-IoT botnet infection numbers' for instance, which has a reported coefficient of 0.17. Our model suggests that for an x% increase in this variable, we expect to see the number of Mirai-like infections increase by a factor of $(1+\frac{x}{100})^{0.17}$. Therefore a 10% increase in initial average non-IoT botnet infections for instance, is expected to correlate with a $1.1^{0.17} = 1.016$ or 1.6% increase in Mirai-like infections. Other coefficients for logarithmic transformed variables may be interpreted similarly.

With respect to the two types of ISP security countermeasures that we are interested in, (i) security efforts to reduce attack surface size, and (ii) general security efforts for instance to combat other types of malware, we find mixed results.

On one hand, we find no clear evidence of efforts to reduce attack surface size to have a significant effect. For instance our model does NOT find any significant correlation between reducing the number of Telnet or FTP reachable customer networks and a reduction in the number of Mirai variant infections in an ISP's network. The same results are echoed for the other TCP services that we have examined, TR069 and Telnet on non-standard ports such as TCP/2323.

On the other hand, we find significant correlations between general ISP security practices, for instance efforts to reduce the number of other types of IoT malware infections, non-IoT malware infections, as well as open DNS resolvers, to correlate with reductions in the number of Mirai variant infections in an ISP's network. The negative coefficient signs of the corresponding independent variables in $model_5$ indicate the direction of this effect. Despite the slightly different transformations applied to these proxy variables, interpreting their effect size analytically works out to be similar to previous cases when we interpreted variables with logarithmic transformations. In short, we see the following effect sizes based on the coefficient values of the model. For a 10% additional reduction in the number of *non-IoT* malware infections, we expect to see a reduction of $1.1^{-0.07} = 0.99$ or a 1% in the number of Mirai variant infection average of the ISP. Stated differently, these effects suggest that ISPs in the top 25% of non-IoT malware reduction have at least 29% lower Mirai infection average rates than ISPs in the bottom 25%. As for a 10% additional reduction in the number of *other IoT* botnet infections, our model suggests that we should expect a $1.1^{-0.14} = 0.98$, or 2% reduction in the number of Mirai variant infection averages to go

with it. Stated differently, ISPs in the top 25% of other-IoT malware reduction have at least 48% lower Mirai infection average rates than ISPs in the bottom 25%.

So what to these effect sizes mean? In sum, our findings suggest that ISP security efforts to combat infections have a demonstrable empirical effect on reducing Mirai-like malware infections. While the effect sizes are not remarkably large, nevertheless we do see an effect in a desirable direction, while also keeping in mind that our models are imperfect and based on imperfect variables approximating the effects of security efforts. Moreover, our results yet again confirm previous empirical findings that poor network hygiene conditions lead to more infections and a concentration of abuse [53]. This means that ISPs could indeed play a role in combating the rising tide of IoT malware, although it appears that specific security efforts to reduce attack surface size may not have the desired effect when other factors are also taken into account. Another findings is that the institutional environment does appear to have a small but significant effect on infection levels: with more developed ICT infrastructure also comes more IoT infections. Finally, we also see that the factor explaining the largest proportion of variation among ISP infection levels is their exposure. And while this is consistent with many empirical studies from security literature it also suggests that the largest part of the complexity of IoT malware lies at the level of individuals and the IoT devices which they connect to the Internet. This interpretation is in line with recent empirical studies that find the types, models and manufactures of IoT devices significantly vary worldwide [10].

7. Discussion

In the face of ever larger IoT security problems, our study empirically examines the role of ISPs in combating IoT malware. We collect data on and examine ISP reactions to Mirai-like malware infections in their networks and find evidence to support that ISPs may, through standard security practices, combat and reduce IoT malware infection levels. There are several noteworthy limitations to our study which we will discuss first, before summarizing our results and discussing their broader implications.

Limitations. An important limitation of our study is that it relies on third-party data sources. Each of these data sources have their own data collection methodology and idiosyncrasies that are not fully understood and which may impact our results in unanticipated ways.

Another limitation of our study is that it examines just one malware family, albeit still a dominant one. While one reason for our focus on Mirai and its variants is certainly their relative ease of tracking, which provides us with a rare longitudinal view, a key reason for our study's Mirai focus is the malware family's lack of persistence and the relative ease with which it may be removed from infected devices, i.e., restarting and resetting the password of infected devices in many cases solves the problem. This provides a suitable basis for measuring if ISPs have had an impact in combating IoT malware. If the impact cannot be observed here, then it is unlikely to have been present for more complex IoT malware families.

We still do caution that our results and the conclusions have to be balanced against the limitations that have we discussed here.

Results and Implications. In our study we find evidence to suggest that the Mirai problem is worse within countries with more developed ICT infrastructure. ICT development likely correlates with higher market penetration of consumer IoT. We also find evidence that many ISPs and their users reduced the attack surface for Mirai and its variants by blocking certain ports used for propagation. A surprising result is that these efforts seem to have had no effect on infection rates in the long run. This might be explained by the fact that newer Mirai variants have moved on to exploiting and propagating through an expansive set of non-standard port/protocol combinations, culminating in a whack-a-mole phenomenon. The targeted countermeasures of ISPs have therefore been rendered less effective over time and given that ISP are unlikely to block all propagation ports. These results however do not suggest attack surface reducing efforts to be a lost endeavor: had these countermeasures not been implemented Mirai infection numbers might have been higher than the levels that we have observed in our data.

And while these results may be unsatisfying, the overall conclusion seems to be positive, as we do find evidence to support that ISPs may play a significant role in combating IoT malware. We find that broadband networks that have poorer network hygiene and abuse mitigation (as measured by proxy of DDoS amplifiers and non-Mirai botnet infections in their networks) also have higher infection rates for Mirai. The implication is that the best practices for general botnet mitigation appear to also be relevant for IoT malware [11, 12]. This suggests that 'generic' best practices prescribed to combat a broad range of malware, are also somewhat effective against IoT malware.

In sum, it appears that ISPs have several countermeasures at their disposal against IoT malware. At the level of their customers' networks they may set more secure default configurations on their customers' router/modem equipment to reduce attack surface: ports that are closed by default, initial passwords that are stronger, basic firewall rules that prevent mass scale port scanning. Other solutions also exist, for instance notifying infected customers, or quarantining infected customer networks (which have been shown to have positive causal remediation effects with respect to Mirai infections in prior work cf. [6]), or even updating equipment and their firmware.

Our findings provide lessons from the first half decade of IoT malware mitigation. That being said, the role of ISPs in mitigation should not obscure the need to develop policies to tackle the root cause of this problem: poor security practices of IoT manufacturers. Different policies in this direction have been discussed in section 2. For instance, minimum security standard requirements for the manufacturing and marketing of IoT devices, e.g., guaranteed software update mechanisms, and requiring strong, unique and modifiable administrative credentials for each device. There are also calls to strengthen consumer rights to return and replace substandard IoT devices. Yet, these solutions will take time to be developed and implemented, let alone become an effective barrier against the influx of insecure devices. Until then, this problem is squarely in the hands of ISPs and their users, for better or worse.

References

[1] Manos Antonakakis, Tim April, Michael Bailey, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J Alex Halderman, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, Yi Zhou, Manos Antonakakis Tim April, Matthew Bernhard Elie Bursztein, Jaime J Cochran Zakir Durumeric Alex Halderman Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever Zane Ma, Joshua Mason, and Nick Sullivan Kurt Thomas. "Understanding the Mirai Botnet". In: *USENIX Security*. 2017.

[2] Netscout. *Dawn of the terrorbit era*. 2018. URL: https://docs.broadcom.com/doc/istr-24-2019-en (visited on 02/18/2021).

[3] Kaspersky. *New Mirai botnet is targeting enterprise IoT — Kaspersky official blog*. 2019. URL: https://www.kaspersky.com/blog/mirai-enterprise/26032/ (visited on 02/18/2021).

[4] OWASP Fundation. *OWASP Internet of Things Project - OWASP*. 2018. URL: https://wiki.owasp.org/index.php/OWASP_Internet_of_Things_Project#tab=IoT_Top_10 (visited on 02/18/2021).

[5] Sara Boddy, Justin Shattuck, Debbie Walkowski, and David Warburton. *The Hunt for IoT: Multi-Purpose Attack Thingbots Threaten Internet Stability and Human Life*. Tech. rep. 2018. URL: https://www.f5.com/labs/articles/threat-intelligence/the-hunt-for-iot--multi-purpose-attack-thingbots-threaten-intern.

[6] Orcun Cetin, Carlos Ganan, Lisette Altena, Takahiro Kasama, Daisuke Inoue, Kazuki Tamiya, Ying Tie, Katsunari Yoshioka, and Michel van Eeten. "Cleaning Up the Internet of Evil Things: Real-World Evidence on ISP and Consumer Efforts to Remove Mirai". In: *Proceedings 2019 Network and Distributed System Security Symposium*. Reston, VA: Internet Society, 2019.

[7] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. "Exploring how privacy and security factor into IoT device purchase behavior". In: *Conference on Human Factors in Computing Systems - Proceedings* (2019), pp. 1–12.

[8] Christopher McDermott, John Isaacs, and Andrei Petrovski. "Evaluating Awareness and Perception of Botnet Activity within Consumer Internet-of-Things (IoT) Networks". In: *Informatics* 6.1 (2019), p. 8.

[9] Eric Zeng, Shrirang Mare, Franziska Roesner, Santa Clara, Eric Zeng, Shrirang Mare, and Franziska Roesner. "End User Security and Privacy Concerns with Smart Homes This paper is included in the Proceedings of the End User Security & Privacy Concerns with Smart Homes". In: *Thirteenth Symposium on Usable Privacy and Security (SOUPS)* (2017), pp. 65–80.

[10] Deepak Kumar, Kelly Shen, Benton Case, Deepali Garg, Galina Alperovich, Dmitry Kuznetsov, Rajarshi Gupta, and Zakir Durumeric. "All Things Considered: An Analysis of IoT Devices on Home Networks". In: *USENIX Security Symposium*. 2019.

[11] MAAWG. *Abuse Desk Common Practices*. 2007. URL: https://www.m3aawg.org/sites/default/files/document/MAAWG_Abuse_Desk_Common_Practices.pdf.

[12] J. Livingood, N. Mody, and M. O'Reirdan. *Recommendations for the Remediation of Bots in ISP Networks (RFC 6561)*. 2012. URL: https://tools.ietf.org/html/rfc6561.

[13] Spamhaus. *Composite Block List (CBL)*. URL: https://www.abuseat.org/.

[14] Orçun Çetin, Lisette Altena, Carlos Gañán, and Michel Van Eeten. "Let me out! Evaluating the effectiveness of quarantining compromised users in walled gardens". In: *Proceedings of the 14th Symposium on Usable Privacy and Security, SOUPS 2018*. 2019, pp. 251–263.

[15] Georgios Kambourakis, Constantinos Kolias, and Angelos Stavrou. "The Mirai botnet and the IoT Zombie Armies". In: *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*. Vol. 2017-Octob. IEEE, 2017, pp. 267–272.

[16] Cloudflare. *Inside the infamous Mirai IoT Botnet: A Retrospective Analysis*. 2017. URL: https://blog.cloudflare.com/inside-mirai-the-infamous-iot-botnet-a-retrospective-analysis/.

[17] Ya Liu and Hui Wang. "Tracking Mirai variants". In: *Virus Bulletin*. October. 2018, pp. 1–18.

[18] Jason Sattler. *IoT threats: Explosion of 'smart' devices filling up homes leads to increasing risks*. 2019. URL: https://blog.f-secure.com/iot-threats/.

[19] Natalija Vlajic and Daiwei Zhou. "IoT as a Land of Opportunity for DDoS Hackers". In: *Computer* 51.7 (2018), pp. 26–34.

[20] Netlab 360. *Miraci C2 Data*. URL: https://data.netlab.360.com/mirai-c2/.

[21] Shadowserver. *Darknet Report*. URL: https://www.shadowserver.org/what-we-do/network-reporting/darknet-report/.

[22] Ang Cui and Salvatore J. Stolfo. "A quantitative analysis of the insecurity of embedded network devices". In: *Proceedings of the 26th Annual Computer Security Applications Conference on - ACSAC '10*. Vol. 95. 22. New York, New York, USA: ACM Press, 2010, p. 97.

[23] Andrei Costin and Jonas Zaddach. "IoT Malware: Comprehensive Survey, Analysis Framework and Case Studies". In: *BlackHat USA*. 2018.

[24] Artur Marzano, David Alexander, Osvaldo Fonseca, Elverton Fazzion, Cristine Hoepers, Klaus Steding-Jessen, Marcelo H. P. C. Chaves, Italo Cunha, Dorgival Guedes, and Wagner Meira. "The Evolution of Bashlite and Mirai IoT Botnets". In: *2018 IEEE Symposium on Computers and Communications (ISCC)*. 2018, pp. 00813–00818.

[25] Brian Krebs. *New Charges, Sentencing in Satori IoT Botnet Conspiracy*. 2020. URL: https://krebsonsecurity.com/2020/06/new-charges-sentencing-in-satori-iot-botnet-conspiracy/.

[26] Sadegh Torabi, Elias Bou-Harb, Chadi Assi, El Mouatez Billah Karbab, Amine Boukhtouta, and Mourad Debbabi. "Inferring and Investigating IoT-Generated Scanning Campaigns Targeting A Large Network Telescope". In: *IEEE Transactions on Dependable and Secure Computing* (2020). Early access, pp. 1–17.

[27] Stephen Herwig, Katura Harvey, George Hughey, Richard Roberts, and Dave Levin. "Measurement and Analysis of Hajime, a Peer-to-peer IoT Botnet". In: *Proceedings 2019 Network and Distributed System Security Symposium*. Reston, VA: Internet Society, 2019.

[28] Sergei Shevchenko. *"VPNFilter" botnet: a SophosLabs analysis*. Tech. rep. May. 2018. URL: https://www.ibm.com/support/knowledgecenter/SSEPGG_9.5.0/com.ibm.db2.luw.sql.rtn.doc/doc/.

[29] Farooq Shaikh, Elias Bou-Harb, Nataliia Neshenko, Andrea P. Wright, and Nasir Ghani. "Internet of Malicious Things: Correlating Active and Passive Measurements for Inferring and Characterizing Internet-Scale Unsolicited IoT Devices". In: *IEEE Communications Magazine* 56.9 (2018), pp. 170–177.

[30] Nataliia Neshenko, Martin Husak, Elias Bou-Harb, Pavel Celeda, Sameera Al-Mulla, and Claude Fachkha. "Data-Driven Intelligence for Characterizing Internet-Scale IoT Exploitations". In: *IEEE Globecom Workshops*. IEEE, 2018, pp. 1–7.

[31] Yin Minn Pa Pa, Suzuki, Shogo, Katsunari Yoshioka, Tsutomu Matsumoto, Takahiro Kasama, and Christian Rossow. "IoTPOT: Analysing the Rise of IoT Compromises". In: *USENIX WOOT*. Vol. 29. 6. 2015, pp. 775–792.

[32] Daniel Ramirez, José Ignacio Uribe, Luca Francaviglia, Pablo Romero-Gomez, Anna i Morral, and Franklin Jaramillo. "IoTCandyJar: Towards an Intelligent-Interaction Honeypot for IoT Devices". In: *BlackHat USA* (2017).

[33] Amit Tambe, Yan Lin Aung, Ragav Sridharan, Martin Ochoa, Nils Ole Tippenhauer, Asaf Shabtai, and Yuval Elovici. "Detection of Threats to IoT Devices using Scalable VPN-forwarded Honeypots". In: *ACM CODASPY*. 2019.

[34] Alexander Vetterl. *Honeypots in the age of universal attacks and the Internet of Things*. Tech. rep. 944. UCAM-CL-TR-944, Univeristy of Cambridge, 2020.

[35] Markus Miettinen, Samuel Marchal, Ibbad Hafeez, Tommaso Frassetto, N. Asokan, Ahmad Reza Sadeghi, and Sasu Tarkoma. "IoT Sentinel Demo: Automated Device-Type Identification for Security Enforcement in IoT". In: *ICDCS*. 2017, pp. 2511–2514.

[36] Xuan Feng, Qiang Li, Haining Wang, and Limin Sun. "Acquisitional rule-based engine for discovering Internet-of-Thing devices". In: *Proceedings of the 27th USENIX Security Symposium*. 2018, pp. 327–341.

[37] Talha Javed, Muhammad Haseeb, Muhammad Abdullah, and Mobin Javed. "Using application layer banner data to automatically identify IoT devices". In: *Computer Communication Review* 50.3 (2020), pp. 23–29.

[38] Dan Yu, Lilong Zhang, Yongle Chen, Yao Ma, and Junjie Chen. "Large-Scale IoT Devices Firmware Identification Based on Weak Password". In: *IEEE Access* 8 (2020), pp. 7981–7992.

[39] Gordon Chu, Noah Apthorpe, and Nick Feamster. "Security and Privacy Analyses of Internet of Things Toys". In: *CoRR* abs/1805.0 (2018).

[40] Omar Alrawi, Chaz Lever, Manos Antonakakis, and Fabian Monrose. "SoK: Security Evaluation of Home-Based IoT Deployments". In: *2019 IEEE Symposium on Security and Privacy (S&P)*. IEEE, 2019.

[41] Roberto Perdisci, Thomas Papastergiou, Omar Alrawi, and Manos Antonakakis. "IoTFinder : Efficient Large-Scale Identification of IoT Devices via Passive DNS Traffic Analysis". In: *ACM IMC*. 2020.

[42] Yair Meidan, Vinay Sachidananda, Hongyi Peng, Racheli Sagron, Yuval Elovici, and Asaf Shabtai. "A novel approach for detecting vulnerable IoT devices connected behind a home NAT". In: *Computers and Security* 97 (2020), p. 101968.

[43] Said Jawad Saidi, Anna Maria Mandalari, Roman Kolcun, Daniel J. Dubois, David Choffnes, Georgios Smaragdakis, and Anja Feldmann. "A Haystack Full of Needles: Scalable Detection of IoT Devices in the Wild". In: *ACM IMC*. 2020.

[44] Shane D. Johnson, John M. Blythe, Matthew Manning, and Gabriel T. W. Wong. "The impact of IoT security labelling on consumer product choice and willingness to pay". In: *PLOS ONE* 15.1 (2020). Ed. by Muhammad Khurram Khan, e0227800.

[45] Philipp Morgner, Christoph Mai, Nicole Koschate-Fischer, Felix Freiling, and Zinaida Benenson. "Security Update Labels: Establishing Economic Incentives for Security Patching of IoT Consumer Products". In: *2020 IEEE Symposium on Security and Privacy (S&P)*. IEEE, 2020, pp. 429–446.

[46] Eireann Leverett, Richard Clayton, and Ross Anderson. "Standardisation and Certification of the 'Internet of Things'". In: *Workshop on the Economics of Information Security (WEIS)*. 2017.

[47] Bruce Schneier. *Securing the International IoT Supply Chain*. 2020. URL: https://www.schneier.com/blog/archives/2020/07/securing_the_in_1.html.

[48] Alan Butler. "Products Liability and the Internet of (Insecure) Things: Should Manufacturers Be Liable for Damage Caused by Hacked Devices?" In: *University of Michigan Journal of Law Reform* 50.4 (2017).

[49] *IoT Cybersecurity Improvement Act of 2020*. 2020. URL: https://www.congress.gov/bill/116th-congress/house-bill/1668/text.

[50] *Dutch regulator sets requirements for IoT device manufacturers*. 2020. URL: https://www.telecompaper.com/news/dutch-regulator-sets-requirements-for-iot-device-manufacturers--1351664.

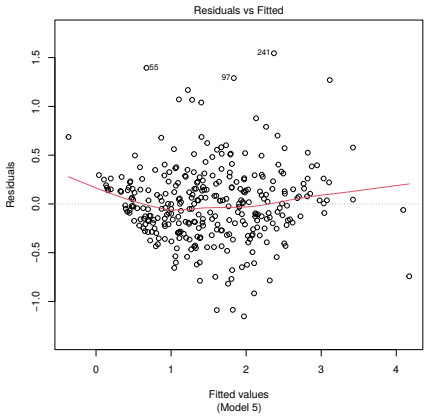
[51] Pieter Meulenhoff, Sjoerd Langkemper, and Willem Westerhof. *Essential requirements for securing IoT consumer devices*. Tech. rep. 2020. URL: https://www.agentschaptelecom.nl/documenten/rapporten/2020/08/26/onderzoeksrapport-essential-requirements-for-securing-iot-consumer-devices.

[52] Davey Winder. *Weak Gadget Passwords Could Be Illegal In 2021, Says U.K. Government*. 2020. URL: https://www.forbes.com/sites/daveywinder/

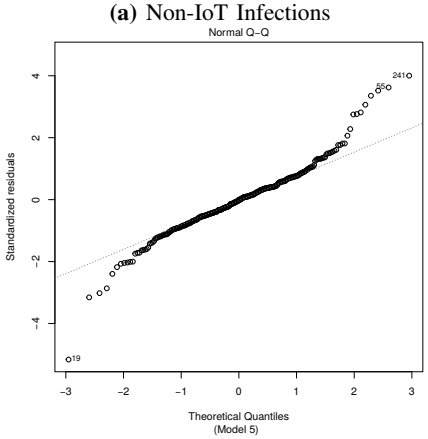
2020/07/19/weak-gadget-passwords-could-be-illegal-in-2021-says-uk-government-iot-hacking-cybercrime-law/.
[53] J Zhang, Z Durumeric, and M Bailey. “On the Mismanagement and Maliciousness of Networks”. In: *NDSS*. 2014.
[54] Arman Noroozian, Michael Ciere, Maciej Korczynski, Samaneh Tajalizadehkhoob, and Michel Van Eeten. “Inferring the Security Performance of Providers from Noisy and Heterogenous Abuse Datasets”. In: *WEIS*. 2017.
[55] Hadi Asghari, Michael Ciere, and Michel J G Van Eeten. “Post-Mortem of a Zombie: Conficker Cleanup After Six Years”. In: *USENIX Security*. 2015.
[56] Hadi Asghari, Michel J.G. van Eeten, and Johannes M. Bauer. “Economics of Fighting Botnets: Lessons from a Decade of Mitigation”. In: *IEEE Security & Privacy* 13.5 (2015), pp. 16–23.
[57] Arman Noroozian, Maciej Korczyński, Carlos Hernandez Gañan, Daisuke Makita, Katsunari Yoshioka, and Michel Van Eeten. “Who Gets the Boot? Analyzing Victimization by DDoS-as-a-Service”. In: *Proc. of RAID*. 2016.
[58] Hadi Ashghari. *PyASN*. URL: https://github.com/hadiasghari/pyasn.
[59] *Maxmind GeoIP2 DB*. URL: https://www.maxmind.com/en/geoip2-isp-database.
[60] CAIDA. *Inferred AS to Organization Mapping Dataset*. URL: https://www.caida.org/data/as-organizations/.
[61] RAPID7. *Rapid7 Project Sonar (open data)*. URL: https://opendata.rapid7.com/.
[62] RAPID7. *Project Sonar National Exposure (open data)*. URL: https://opendata.rapid7.com/sonar.national_exposure/.
[63] RAPID7. *Project Sonar TCP Scans (open data)*. URL: https://opendata.rapid7.com/sonar.tcp/.
[64] Radware. *Deutsche Telekom Mirai Botnet Takeover Attempt*. 2016. URL: https://security.radware.com/ddos-threats-attacks/threat-advisories-attack-reports/deutsche-telekom-routers-takeover/.
[65] Logitec. *Notice and Instructions for Updating Router Firmware*. 2017. URL: https://www.logitec.co.jp/info/2017/1219.html.
[66] JPCERT. *Alert Regarding Mirai Variant Infections*. 2017. URL: https://www.jpcert.or.jp/english/at/2017/at170049.html.
[67] Fortinet. *Rise of One More Mirai Worm Variant*. 2017. URL: https://www.fortinet.com/blog/threat-research/rise-of-one-more-mirai-worm-variant.
[68] Marc Kührer, Thomas Hupperich, Christian Rossow, and Thorsten Holz. “Exit from Hell ? Reducing the Impact of Amplification DDoS Attacks”. In: *USENIX Security*. 2014.

Appendix

Figure 8 plots several goodness-of-fit measures for $model_5$ reported in paper which demonstrate a reasonable fit of the linear model to the underlying data.



(a) Non-IoT Infections



(b) Other IoT Infections

Figure 8: Residuals Plot and QQ-plot for $model_5$ reported in Table 2.

Acknowledgments

The authors would like to thank our anonymous reviewers for their feedback and suggestions to improve the quality of our manuscript. We would like to thank the Dutch Ministry of Economic Affairs for supporting our research as well as NICT, Rapid7 and Spamhaus for providing access to the underlying data of this study. This publication is part of the MINIONS project (number 628.001.033) of the “Joint U.S.-Netherlands Cyber Security Research Programme” which is (partly) financed by the Dutch Research Council (NWO); and of the related MINIONS-TLD project, which is financed by SIDN, the .nl registry.