# Cleaning Up the Internet of Evil Things

## Real-World Evidence on ISP and Consumer Efforts to Remove Mirai

Cetin, Orcun; Hernandez Ganan, Carlos; Altena, Lisette; Kasama, Takahiro; Inoue, Daisuke; Tamiya, Kazuki; Tie, Ying; Yoshioka, Katsunari; van Eeten, Michel

**DOI**
10.14722/ndss.2019.23438

**Publication date**
2019

**Document Version**
Final published version

**Published in**
Network and Distributed System Security Symposium (NDSS) 2019

**Citation (APA)**
Cetin, O., Hernandez Ganan, C., Altena, L., Kasama, T., Inoue, D., Tamiya, K., Tie, Y., Yoshioka, K., & van Eeten, M. (2019). Cleaning Up the Internet of Evil Things: Real-World Evidence on ISP and Consumer Efforts to Remove Mirai. In *Network and Distributed System Security Symposium (NDSS) 2019* (26th Annual Network and Distributed System Security Symposium, NDSS 2019). https://doi.org/10.14722/ndss.2019.23438

# Cleaning Up the Internet of Evil Things: Real-World Evidence on ISP and Consumer Efforts to Remove Mirai

Orçun Çetin[1], Carlos Gañán[1], Lisette Altena[1], Takahiro Kasama[3], Daisuke Inoue[3], Kazuki Tamiya[2], Ying Tie[2], Katsunari Yoshioka[2], and Michel van Eeten[1]

[1]Delft University of Technology, Email: {f.o.cetin, c.hernandezganan, e.m.altena, m.j.g.vaneeten}@tudelft.nl
[2]Yokohama National University, Email: {tamiya-kazuki-gj, tie-ying-fc, yoshioka}@ynu.ac.jp
[3]National Institute of Information and Communications Technology, Email: {dai, kasama}@nict.go.jp

***Abstract*—With the rise of IoT botnets, the remediation of infected devices has become a critical task. As over 87% of these devices reside in broadband networks, this task will fall primarily to consumers and the Internet Service Providers. We present the first empirical study of IoT malware cleanup in the wild – more specifically, of removing Mirai infections in the network of a medium-sized ISP. To measure remediation rates, we combine data from an observational study and a randomized controlled trial involving 220 consumers who suffered a Mirai infection together with data from honeypots and darknets. We find that quarantining and notifying infected customers via a walled garden, a best practice from ISP botnet mitigation for conventional malware, remediates 92% of the infections within 14 days. Email-only notifications have no observable impact compared to a control group where no notifications were sent. We also measure surprisingly high natural remediation rates of 58-74% for this control group and for two reference networks where users were also not notified. Even more surprising, reinfection rates are low. Only 5% of the customers who remediated suffered another infection in the five months after our first study. This stands in contrast to our lab tests, which observed reinfection of real IoT devices within minutes – a discrepancy for which we explore various different possible explanations, but find no satisfactory answer. We gather data on customer experiences and actions via 76 phone interviews and the communications logs of the ISP. Remediation succeeds even though many users are operating from the wrong mental model – e.g., they run anti-virus software on their PC to solve the infection of an IoT device. While quarantining infected devices is clearly highly effective, future work will have to resolve several remaining mysteries. Furthermore, it will be hard to scale up the walled garden solution because of the weak incentives of the ISPs.**

## I. Introduction

Events of the past two years have made it abundantly clear that Internet of Things (IoT) devices are being compromised at scale, especially in the consumer space. It is also clear that this situation will not improve in the short term. Due to lack of effective regulations, poorly-secured devices will keep flooding the market. Given the life cycle of the existing and new devices, this means we will be confronted with IoT botnets for years to come.

All this presents us with a critical challenge: how can we remediate the population of vulnerable and compromised IoT devices? Since most of the compromised devices are consumer products, this implies overcoming a number of unsolved problems. A recent study into Mirai [1] identified three critical challenges. First, there is no public information to identify the owner of the device. Second, there is no established communication channel to reach the owner. Third, where owners are reachable, we do not know how to provide them with an actionable notification. There is often no clear and simple remediation path. In fact, in many cases we cannot even state exactly which of the owner's devices is actually affected.

For the first two problems, identifying and contacting owners, we can turn to an existing arrangement: botnet mitigation by Internet Service Providers. Many of the devices are in access networks, so ISPs can identify and contact the customers who own them. For regular PC-based malware, botnet mitigation by ISPs is widely accepted and has met with some success [3]. However, cleaning up infected devices is still an open problem, even when considering conventional malware. Years of usability research have shown just how hard it is to support end users with little technical expertise in protecting and remediating their personal computers [15]. In the IoT space, all of this becomes much harder. User intuitions ('folk models' [34]) about security are even less aligned with the IoT environment. Furthermore, the actions users need to take are different across devices, vendors and local configurations. Finally, contrary to conventional malware, there are no automated tools to support users in protecting and remediating infected devices. In short, we have no clue whether owners can act at all effectively on the kind of notifications that we can currently provide them with.

We present the first empirical study of the cleanup of compromised IoT in the wild. For this, we collaborate with a mid-sized ISP that notifies Mirai-infected customers via email or by placing their connection in a quarantine network – a so-called 'walled garden'. We measured the remediation rate and speed of 220 users in an observational study and a randomized

Network and Distributed Systems Security (NDSS) Symposium 2019
24-27 February 2019, San Diego, CA, USA
ISBN 1-891562-55-X
https://dx.doi.org/10.14722/ndss.2019.23438
www.ndss-symposium.org

controlled experiment by tracking the infections in darknet, honeypot and abuse reporting data. We combined this with additional scan data to identify the type of devices that are affected. Next, we studied the user experience by conducting 76 phone interviews and analyzing the logs of the users' communications with the ISP. Finally, we also conducted lab tests with real IoT devices to observe the effectiveness of removal actions and to measure reinfection speed.

In short, we make the following contributions:

- We show that over 87% of all Mirai-infected IoT devices reside in broadband access networks, underlining the critical role of ISPs in IoT botnet mitigation.
- We provide the first real-world measurement of remediation rates for Mirai-infected devices and find that quarantining and notifying affected customers remediates 92% of the infections.
- We find very high natural remediation rates of 58-74% in the control group and in two reference networks where no notifications were sent, probably reflecting the non-persistent nature of the malware.
- We find a remarkably low reinfection rate. Only 5% of the customers who remediated suffered another infection in the five months after our first study. This highlights the effectiveness of the countermeasures taken by the infected customers but stands in contrast to our lab tests, which found very fast reinfections of real IoT devices.
- Remediation succeeds even though customer interviews and communications show that many users are operating from the wrong mental model – e.g., they run anti-virus software on their PC to solve the infection of an IoT device.

Combining insights on the location of compromised IoT devices, effectiveness of different treatments and the experience of real-world users, we contribute scientific evidence for establishing industry best practices around the remediation of compromised IoT.

## II. ISP BOTNET MITIGATION

Cleaning up infected IoT devices can be seen as the next phase of a long-standing challenge: fighting botnets. Over the past decade, mitigation of PC-based malware has consisted of two complementary approaches: taking down the command-and-control infrastructure and cleaning up the infected hosts. Cleanup is an arduous process that demands efforts from different actors, such as operating system vendors, anti-virus vendors, ISPs and the affected end users. As most infected machines reside in consumer broadband networks [3], the role of ISPs has become more salient over time. A range of best practices and codes of conduct have been published by leading industry associations [21], [23], public-private initiatives [18], [11] and governmental entities [17], [12]. These documents share a common set of recommendations for ISPs around educating customers, detecting infections, notifying customers, and remediating infections.

While the existing mitigation practices of ISP are exclusively focused on PC-based malware, they might still provide a good starting point for the remediation of compromised IoT. This assumes, however, that the bulk of the devices reside in the networks of broadband consumer ISPs. To test this assumption, we analyzed the location of compromised devices.

First, following the approach of Antonakakis *et al.* [1], we used darknet data to observe the location of devices infected with a version of Mirai. Darknets, also known as network telescopes, are routed but unused IP address ranges. They passively monitor all arriving traffic at these ranges. We leverage observations from a darknet of approximately $300,000$ IPv4 addresses, spanning 40 networks in 15 countries. As Mirai malware displays worm-like behavior, actively scanning the Internet for spreading itself, we can track its presence in the darknet data. We use data collected in the period January 2016 to April 2018.

We measured per protocol –i.e., per destination port– how many IP addresses were scanning at any point in time. To distinguish Mirai traffic from backscatter traffic and other scanning activity, we uniquely fingerprinted Mirai probes based on an artifact of Mirai's stateless scanning, where every probe has a TCP sequence number – normally a random 32-bit integer – equal to the destination IP address. We observed over 96 million IP addresses. Figure 1 shows how they are distributed over six protocols: 23/TCP (Telnet), 2323/TCP (Telnet), 5358/TCP (Telnet), 5555/TCP (TR-069/TR-064), 6789/TCP (Telnet), 7547/TCP (TR-069/TR-064), 23231/TCP (Telnet), 37777/TCP (UPnP), 22/TCP (SSH), 2222/TCP (SSH), 80/TCP (HTTP), 81/TCP (HTTP), 88/TCP (HTTP), 8000/TCP (HTTP), 8080/TCP (HTTP), and 53869/TCP (Realtek SDK Miniigd). Since Mirai's source code was publicly released, it expanded from targeting telnet to other ports. While port 23 is the second most targeted port, HTTP-related ports have become the main vector – i.e., IoT devices with default credentials for HTTP-related services.

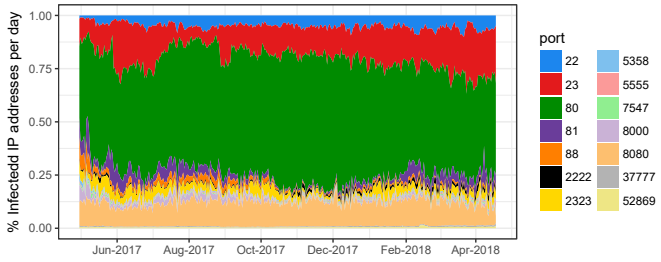



Figure 1: Percentage of Mirai-infected IP addresses per port

Next, we mapped these IP addresses to broadband consumer ISPs and other network types. We use the same approach as a study on 'regular' botnet mitigation by broadband ISPs, where a mapping had been developed to identify the Autonomous System Number (ASN) of broadband ISPs in 82 countries [3]. The mapping is organized around ground truth data in the form of a highly accurate commercial database; *TeleGeography Globalcomms* [29], containing market data on the broadband ISPs in 211 countries. In total, $2,050$ ASNs have been labeled manually as belonging to one of the consumer broadband ISPs or to another category: mobile provider, another type of ISP (e.g., business provider), hosting, governmental, educational and other types of networks. Table I

summarizes the percentage of infected IP addresses in each of the network types. The overwhelming majority of these devices (87.61%) are located in ISP broadband networks, while less than 1% reside in other types of networks including hosting, education or governmental networks.

Table I: Distribution of infected hosts across different markets as captured by the darknet (Jan 2016 - April 2018)

| | | | |
|---|---|---|---|
| #Countries | | 232 | |
| #ASNs | | 21,196 | |
| #IP addresses: | | | |
| | ISP-broadband | 78,885,434 | (87.61 %) |
| | ISP-mobile | 6,888,640 | (7.65 %) |
| | ISP-other | 3,380,164 | (3.75 %) |
| | Hosting | 196,123 | (0.22 %) |
| | Educational | 30,765 | (0.03 %) |
| | Governmental | 313 | (0.01 %) |
| | Others | 655,753 | (0.72 %) |
| | Total | 96,041,559 | (100 %) |

## III. Partner ISP Remediation Process

Now that we have established that ISPs are in a crucial position to remediate IoT botnets, even more so than for PC-based botnets, the question becomes: what can they realistically do? To answer this question, we have collaborated closely with a medium-sized European ISP with several million customers. The ISP decided to include an abuse feed with Mirai-infected hosts, reported by Shadowserver, in their existing botnet notification and remediation process.

At the heart of the ISP's process is an industry best practice: placing an infected machine into a quarantine network, a so-called walled garden [24]. There are different ways of implementing walled gardens to fight malware infections. RFC6561 [21] describes two types: *leaky*, an implementation that permits access to all Internet resources, except those that are deemed malicious; and *strict*, an implementation that restricts almost all services, except those on a whitelist. Our partner ISP has implemented a strict version for its consumer broadband subscribers. The walled garden only allows access to 41 white-listed domains, which provide cleanup tools, anti-virus solutions, Microsoft updates, webmail, online banking and a forum for elderly people.

Besides keeping the infected users safely in quarantine, the walled garden also plays an important role in notifying the user. When the user tries to browse the Web, she or he will be redirected to a landing page with a notification about the infection and advice on how to clean it up. The same information is also sent by email to the customers. Whereas emails with the same content can be ignored relatively easily, the walled garden notification cannot.

Next to its own brand, the ISP also provides services to broadband consumers via a subsidiary brand that is targeting the cheaper end of the market. Customers of the subsidiary brand are not quarantined. Notifications are less common and conducted only via email. The ISP also sells subscriptions in the business and mobile service networks. These customers are never quarantined and do not receive IoT related security notifications.

The notification and remediation process starts when an infection is reported in one of the trusted abuse feeds that the ISP receives. For IoT malware, the ISP uses the daily Shadowserver Drone feeds [26]. These include infections labeled as Mirai. The infected machines are discovered through a range of methods, including monitoring sinkhole traffic and malicious scans to honeypots. If an IP address in the report belongs to one of its consumer broadband subscribers, then the ISP places the connection of that customer in the walled garden. It also sends an accompanying email with the same information. Occasionally, e.g., when the walled garden is full, the ISP sends an email-only notification about the infection.

Once customers are notified via the walled garden, they have three ways of getting out of the quarantine environment. First, they can release themselves by filling out the form on the landing page and report how they have fixed the problem. Submitting the form immediately restores the connection. This option is revoked after two subsequent quarantine events within 30 days, to avoid customers using this route without making an effort to clean up. The second release option is to ask for assistance from the ISP abuse staff to restore the connection. Customers might end up in assisted release because they no longer have the self-release option or because they have contacted the ISP for help. Quarantined customers can contact abuse desk members via email, via the walled garden form, or they can call the regular help desk. The third option is to get a time-out release. After 30 days, customers are automatically released, even if they have not contacted the ISP.

## IV. Study Design

Aiming at understanding the impact of the notifications on the remediation process of Mirai-infected devices, we designed a study which consisted of two stages: (i) an observational study on walled garden notifications that the ISP conducted during 4 months; and (ii) a randomized controlled experiment to assess the impact of an improved notification tailored to IoT infection remediation. Figure 2 shows the timeline of both studies. Furthermore, to understand Mirai infection dynamics, we also conducted a battery of tests with real vulnerable devices.

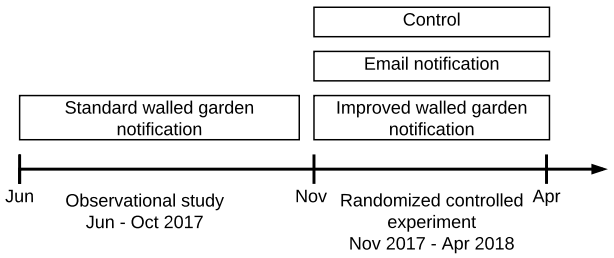



Figure 2: Timeline of the experiment

### A. Data sources

To identify and track the infected Mirai devices in the ISP network, we leveraged a combination of several data sources. Table II provides a high-level summary.

*1) Daily Shadowserver abuse feeds:* The Shadowserver Foundation is a non-profit security organization that gathers and distributes data on abused Internet resources, most notably compromised machines. It provides network operators with a

Table II: Data Sources – We used various data sources to analyze the remediation rate of infected ISP subscribers

| Role | Data Source | Collection Period | Data Volume |
|---|---|---|---|
| Detecting infections | Shadowserver drone feed | 01/06/2017-18/04/2018 | 658 IP addresses |
| | IoT honeypot | 01/06/2017-18/04/2018 | 512 IP addresses |
| Tracking infections | Darknet | 01/06/2017-18/04/2018 | 349 IP addresses |
| | Shadowserver drone feed | 01/06/2017-18/04/2018 | 349 IP addresses |
| | IoT honeypot | 01/06/2017-18/04/2018 | 281 IP addresses |
| Device identification | Censys scans | 02/05/2017-16/04/2018 | 49 Internet-wide scans |
| | Nmap scans | 01/06/2017-18/04/2018 | 349 port scans |
| Customer experience | Phone interviews | 10/10/2017-18/04/2018 | 76 subscribers |
| | Walled garden forms | 01/06/2017-18/04/2018 | 159 forms |
| | Communication logs | 01/06/2017-18/04/2018 | 521 tickets |

daily report on compromised hosts in their networks (Botnet-Drone feed [26]). We use the daily reports sent to our partner ISP, in combination with other datasources, to detect and track Mirai-infected users. During the study period, 658 IP addresses that belong to one of the ISP's networks were detected as infected with Mirai. We selected 349 of these IP addresses for the purpose of our study (see Section IV-C for the specifics of the selection process). These 349 IP addresses correspond to 343 different subscribers, i.e., there are 6 subscribers whose IP addresses were not completely static during the study period.

*2) IoT Honeypot:* An additional data source for detecting and tracking infected devices are the daily log files of a low-interaction honeypot running the open-source IoTPOT software [25]. This IoT-specific honeypot emulates various well-known vulnerable network services by implementing specific IoT architectures. These emulated services include Telnet protocol, IoT devices' HTTP front-ends, the CPE WAN Management Protocol (CWMP) and the remote access setup service of several types of IP cameras. To capture infected IoT devices, the honeypot has been deployed over 738 IP addresses distributed across three countries, including the country in which the partner ISP operates.

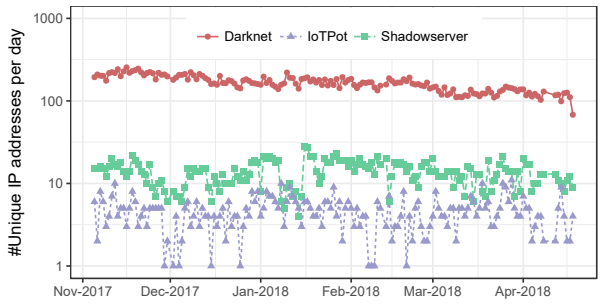



Figure 3: Number of unique IP addresses per day of Mirai-infected hosts in the consumer broadband network of the ISP, as detected by Shadowserver, darknet, and honeypot (log-scale)

During the study period, the honeypot captured 512 different IP addresses that belonged to the partner ISP. As the ISP only relies on Shadowserver feeds, we did not use these IP addresses for notification purposes –note that 54.9% (281 IP addresses) of them overlapped with the IP addresses captured by Shadowserver– but instead we used them to track the infections together with the darknet.

*3) Darknet:* A third data source for detection and tracking is the darknet mentioned in Section II. We have monitored 16 protocols that are known to be abused by Mirai botnets for the network ranges operated by the partner ISP. The darknet data is much more granular than the honeypot and Shadowserver data, so we mostly rely on this data for measuring the time to remediation.

Figure 3 shows the number of unique IP addresses seen each day in each of the data sources. The darknet has the best coverage, with around 150 unique IP addresses seen every day. The honeypot and the Shadowserver observe only around 10% of these hosts. It is important to note that the ISP's abuse handling process only works with the Shadowserver feed. We use the darknet and IoT honeypot sources only for tracking the infected hosts that entered the ISP abuse handling process.

*4) Censys Scans:* Censys [9] is a platform that scans the IPv4 space and aggregates application layer data about hosts on the Internet. We obtained the raw scan data for 49 Internet-wide scans, including HTML code and banner information, for each IP address of the ISP where an infected host was observed. We focused our analysis on scans of CWMP (port 7547), FTP (port 21), HTTP (port 80 and 8080), HTTPS (port 443), SSH (port 22) and Telnet (port 23 and 2323) between May 01, 2017 and April 31, 2018.

*5) NMAP Scans:* We used the Nmap network scanner tool [22] to enrich the dataset used for the device identification. Once a device was identified as infected with Mirai, we obtained a list of the open ports as well as banner information. In total, we scanned 349 IP addresses, though 67 of these were already off-line at the time of the scan.

### B. Treatment variables

Our studies are designed to determine the impact of different notification mechanisms on remediation. For this purpose, we compare two experimental treatments using a different notification method (walled garden and email-only) to a control group where no notifications were made during the experiment period. While preparing the experiment, we also improved the standard ISP notification message so as to provide more actionable advice to users. We assess the impact of the improved message via comparing the remediation rate and speed for the new walled garden notification to those measured in the observational study, where the ISP was still using the standard walled garden notification. Figure 2 summarizes the different treatment groups that we compare across the two studies. We now take a closer look at the two main treatment variables: notification method and notification content.

*1) Notification method:* ISPs have various methods to notify end users for malware infections, such as email, phone calls, SMS, postal mail and a walled garden. However, the efficacy of these methods has rarely been studied, let alone for

IoT malware cleanup. In the experimental study, we compare two common methods: email and walled garden.

***Email:*** This method is commonly used by ISPs as it is cheap and easy to scale. However, a major drawback is that it cannot be assured that the email is read in a timely manner, or whether it is read at all. A user might use a different primary email address than the one provided by or to the ISP. The user's email service might also classify the notification as spam. In short, while email is a convenient method, it is unclear how effective this is in terms of promoting IoT malware cleanup.

***Walled garden:*** Walled garden notifications – i.e., the landing page in the quarantine environment – are much more likely to be read by a user. Furthermore, the quarantining provides a strong incentive for the user to remediate. That being said, remediation is not assured. The option of self-release does provide an option to leave the walled garden without any action. Also, when the ISP staff provides an assisted-release, it cannot actually see whether the user successfully remediated. Only when a later Shadowserver report flags the same user again, might the ISP conclude that cleanup failed.

*2) Notification content:* Crafting usable security notifications for end users is a difficult challenge. A range of previous studies have focused on how different abuse and vulnerability notification contents can expedite remediation of the security issues [32], [19], [7]. However, such work has not been conducted on remediating IoT malware nor with consumers in real-world broadband networks.

We discussed with the partner ISP the standard notification content that they were using (see Appendix A). We noticed it used technical jargon that is probably unfamiliar to most consumers (e.g., Telnet, SSH). Also, the steps that customers were supposed to take were somewhat buried in the overall message. In collaboration with the ISP, we drafted an improved version which avoided certain technical terms and organized the remediation in a numbered series of steps, which we hoped would be more actionable for users. We also added steps to reset the router, as this would close all ports as well as disable the demilitarized zone (DMZ) and universal plug and play (UPnP) (See Appendix B).

### C. Study procedure

As shown in Figure 2, our study consisted of two stages. The first stage was an observational study of the effectiveness of the existing ISP walled garden mechanism. In the period from June 2017 to the end of October 2017, the ISP quarantined 97 customers and informed them via the standard walled garden notification. All of these users were reported by Shadowserver as having a Mirai infection. We looked up customer IDs and the set of IP addresses associated with each customer over the period of the study. (Most users retained the same IP address.) We then checked these IP addresses against our three sources of infection data: Shadowserver, honeypot, and darknet. We also logged how long each customer had spent in quarantine, during which they would not be observed in the infection data, of course. By combining these datasets, we could measure remediation success and speed for each user.

Once the first stage of the study was done, we continued with the randomized controlled experiment. To determine the total sample size, in other words how many users needed to be notified, we completed a power calculation for the main outcome variable, remediation rate. We estimated power for an 90% level and used a 10.95 standard deviation based on prior studies [6]. Differences in mean fourteen-day cleanup time of about 10 hours between conditions can be detected with 90% power in two-tailed tests with 95% confidence, based on a sample of 40 Mirai-infected users in each treatment group. This resulted in a total sample size of 120 Mirai-infected users.

The experiment was conducted from the first week of November 2017 to early April 2018. Throughout this period, we followed the procedure summarized in Figure 4. First, for each IP address in the Shadowserver report, we identified the customer ID. Then, we checked whether this customer was notified before for Mirai. As prior experience with the notification procedure and remediation actions might influence the remediation time, we discarded a handful of cases that had been notified previously. All others were randomly assigned: 40 users ended up in the walled garden treatment group, 40 in the email-only group, and 43 in the control group. To establish a baseline, the control group was notified later than the treatment groups. For ethical reasons, this delay has to be limited, so as not to expose the customers to unnecessary risks. In collaboration with the ISP, this delay was set at 14 days. After these 14 days, if these customers were still reported in the Shadowserver feeds, then they would be assigned to the walled garden treatment group. When customers in either of the treatment groups were seen again in the Shadowserver feed within this period of 14 days, we would repeat the treatment. In practice this means that some users got multiple e-mails or were quarantined more than once. This study design means the comparison of the treatments will be conducted over these 14 days, though we did keep track of infections and reinfections after this period, as well will explain below.

In parallel to the experiment (November 2017–April 2018), we also collected data on the remediation of infected customers in two additional networks that belong to two different networks of the partner ISP: (i) business services and; (ii) a subsidiary operating under another consumer brand offering broadband services. Customers in these networks do not receive any IoT malware notifications from the ISP. During the experiment period, the business network had 62 infected customers and the subsidiary network had 61 infected customers. We used the same methodology as in the observational study to estimate the remediation rates and compared these to the control group of the consumer network.

Finally, we conducted tests in a lab setup to observe Mirai's infection, cleanup and reinfection process with real vulnerable IoT devices. By infecting these devices with the malware captured with the honeypot, we could test certain assumptions about removal and reinfection.

### D. Tracking the infected hosts

Remotely assessing the cleanup status of an IoT device is daunting as passive data sources only allow us to corroborate infections, not cleanup. In this sense, the fact that IP addresses disappear from the infection data (Shadowserver Mirai feed, IoTPOT and darknet) do not necessarily mean the device is clean. We could also be missing observations. It is quite

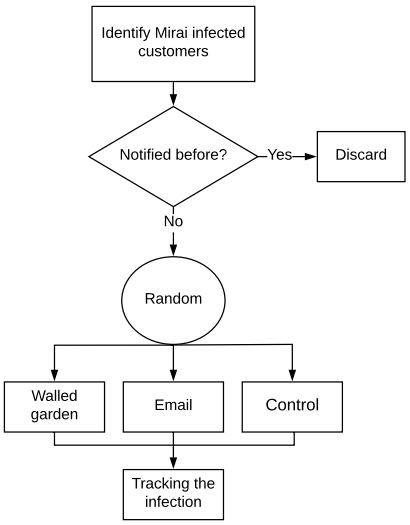



Figure 4: Diagram of the randomized controlled experiment

possible for an infected device to not be seen for a few days in the Shadowserver, IoTPOT and darknet data. This can be caused by a range of reasons, including temporary network disruptions, behavior of the malware or the infected device. (We discuss these limitations in Section IX.)

Without additional safeguards, the missing observations during the 14-day period that we track the infections could easily lead us to overestimate the remediation rate. To mitigate this issue, we include a safeguard. After the 14 days, we monitor the infection data sources for an additional 21 days for recurring observations of the customers that were in the experiment. If we see a customer again in this period, we will assume that he has not remediated during the 14 day period. For 34 (15%) of all customers, we collected one or more infection observations in the 21 day period. We therefore set their status to *not remediated* – i.e., still infected – at the end of the 14 days.

Our conservative approach has one downside: within the period of 35 days (14+21), we treat every observation in the Shadowserver, IoTPOT and darknet data as evidence that the infection persists. In reality, some of these cases will be reinfections of devices that had been clean for a short period, rather than continuously infected. In other words, within this period of 35 days we cannot distinguish between infection and reinfection. To reliably measure reinfection rates, we therefore turned to the customers from the observational study. We continuously monitored our data sources for the IP addresses associated with these customers for five months after the observational study period ended in October 2017. If at any point between November 2017 and early April 2018 we saw these customers reappear in the Shadowserver, IoTPOT or darknet data, we would count these cases as reinfections.

## V. Results

We can now evaluate the effectiveness of the Mirai notifications. As can be seen in Figure 5, the total number of Mirai-infected customers was reduced from around 150 to less than 80 infected customers per day at the end of the experiment.

To further understand the impact of the experiment, we will first compare the results for the different treatments (improved

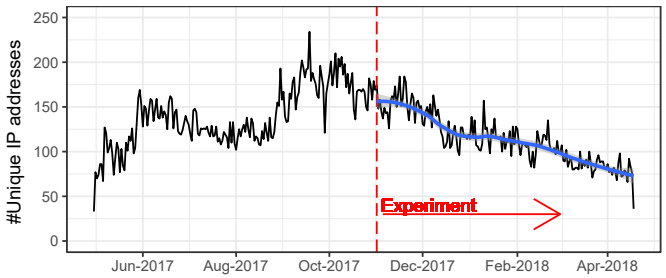



Figure 5: Number of infected devices on the ISP's consumer market before and after the notification experiment

walled garden and email-only notification) to both the control group (no notification) and group of the observational study (standard walled garden notification). Next, we will dive into the high remediation rates for the control group. We find similar results in the two reference networks (business and subsidiary brand) where no notifications were issued. We will then discuss the issue of reinfection and long term efficacy of remediation as well as the influence of device type on cleanup. We will end with discussing the results from lab experiments with remediation and reinfection of real IoT devices.

### A. Impact of notification mechanism

We first determined the impact of notifications on remediation by comparing the experimental groups. The top of Table III shows the percentage of IoT devices that were remediated 14 days after the initial notification. It also includes the median infection time for each group. The control group achieved the lowest cleanup rate (74%), closely followed by email-only treatment group (77%). Remarkably, the email-only treatment seemed to have no effect, displaying no statistically significant difference with the control group. The remediation rate of the email-only group is a bit higher, but the median time is a bit longer. The results were significantly better for the customers who received the improved walled garden notifications: 92% of the infected devices were remediated after 14 days. The median infection time is substantially shortened as well: 26 hours, less than half of the 66 hours for the control group.

We also plotted the survival probabilities for the different groups (see Figure 6a). The groups are quite close one day after the notification, but by day five we see notable differences in the cleanup rates. For instance, 60% of the infected devices in the control group are cleaned within 5 days, compared to 55% of those receiving an email notification and 88% of those receiving improved walled garden notifications.

The log-rank test shows that the difference between the control group and the improved walled garden treatment group is significant ($\chi^2 = 4.4$, $p = 0.0359$). In short, these results provide evidence that quarantining is effective, while email-only notifications are not.

### B. Impact of notification content

To investigate if the improved notification content made a difference, we compared the remediation rates of the walled garden group in the experiment to that in the observational study. Remember, the customers in the observational study

were notified with the standard message. Table III shows a slightly higher cleanup rate and a shorter the median infection time for the improved walled garden treatment group compared to the standard walled garden treatment group. This difference, however, does not pass the log-rank significance test ($\chi^2 = 1.7$, $p = 0.197$). Either the effect is too small to be visible with our sample size or there is no effect. We should also note that this comparison is hampered by the fact that the studies were conducted at different periods in time. In any case, we cannot observe a clear impact of the more actionable walled garden content.

### C. Natural remediation

As we have seen in section V-A, the control group showed remarkably high remediation rates, even though they were not notified.

To confirm the presence of this natural remediation in other networks, we randomly selected 4 other ISPs within the same country where our partner ISP operates and investigated the remediation rates during the period of the observational study. Though we do not control for the potential causes of remediation, figure 7 shows that all 5 ISPs actually experience some degree of remediation in their networks even though not all of them issue notifications regarding Mirai-infections. This suggests the pervasive presence of a natural remediation process across different networks. We have investigated potential explanations for this result.

We can rule out three sources of error. First: DHCP churn. Churn often affects measurements that use IP addresses as identifiers for hosts or users. This greatly complicates external tracking, as devices might be assigned new IP addresses during the measurement period. Our results are immune to this problem, as we knew the ISP's customer ID for each user in the study. The ISP's DHCP logs gave us ground truth on the different IP addresses that were assigned to each customer ID over time. Second source of error: additional notifications. If customers in the control group were to receive some other security notification during the experiment, this might trigger remediation actions that could also affect the Mirai infection. Our design, however, ensured that customers in the control group would not receive any other notifications during the 14-day period.

A third source of error we investigated was whether our ability to track infections deteriorated over time. We speculated that perhaps cleaned devices would get reinfected with new Mirai variants or other IoT malware families that we could not observe in the darknet data using Mirai's TCP sequence number artifact. While theoretically we cannot rule this out, we do observe that overall Mirai infection levels remained more or less constant in the darknet data. So the Mirai variants that produced the initial infections were still very active. There was even an increase in command-and-control servers reported during that period [16]. Also, we saw none of the affected customers reappear in the other two datasets: Shadowserver and IoTPOT.

One explanation that can explain, at least partially, natural remediation is the fact that Mirai infections are reported to be non-persistent [31]. We also confirmed this ourselves (see section V-G). This means that every power cycle or unplugging action leads to cleanup. High natural cleanup might thus be driven by users who turn off devices or otherwise disconnect them, rather than use them continually. Indeed, many of these infections are very short-lived. Around 37% of the infections in the control group are seen only once or twice and disappear from the darknet data within one hour. These transient infections might also reflect volatile usage patterns specific to certain IoT devices. Think of a NAS device that is temporarily connected to another network, perhaps at a friends house. It gets infected there, but then is removed again from the network.

Now, these devices might get cleaned naturally because of usage patterns, but wouldn't they quickly get reinfected again when they are turned back on? In the experiment, we cannot distinguish between infection and reinfection (see Section IV-D), so this might happen. However, all the devices that we counted as clean were not seen again for 21 days after the experimental treatment ended. This suggests that reinfection stopped at some point. Something must have changed, beyond a mere reboot. We take a closer look at the issue of reinfection in section V-E.

### D. Natural remediation in other networks

To investigate whether the high natural remediation rate in the control group was an idiosyncratic result specific to this network or customer base, we also analyzed the infection data for two other networks of the same ISP: their business services network and the network of a subsidiary brand offering consumer broadband on the cheaper end of the market. We compared the remediation rate of the control group from the experiment to the rates for the two other networks. As with our control group, the customers in the two other networks did not receive any notifications for IoT infections from the ISP. This makes them very relevant points of reference.

As shown in Table III, the other networks also display high natural remediation rates. The rate in the business network (55%) was lower compared to the control group (74%) and the subsidiary (74%). Remediation in the two consumer groups (control and subsidiary), however, are virtually the same. Figure 6b also shows this pattern. The log-rank test reports a significant difference between customers with business service subscription and the control group (log-rank test, $\chi^2 = 5.4$ with $p-value = 0.0196$) and business network and subsidiary network (log-rank test, $\chi^2 = 4.9$ with $p - value = 0.0268$).

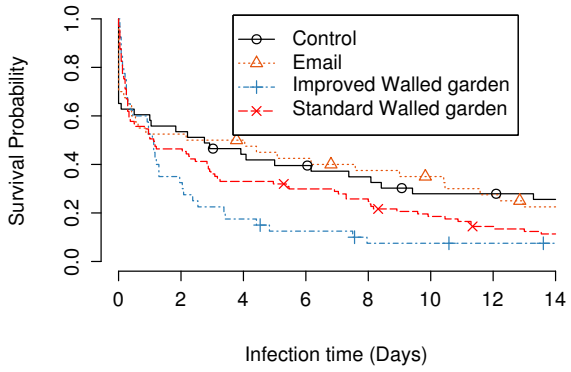
The median infection time for the business network was also significantly longer compared to the other networks. One hypothesis for this finding is that for business continuity reasons, business customers are less likely to reboot or power off their devices as often as consumers. Related to this different usage pattern, we would also expect the composition of IoT device types to be different from the two consumer groups. As we will discuss in section V-F, this is in fact the case. Taking these factors into account, we find very consistent natural remediation rates across the different networks, increasing our confidence in the results of the experiment.

### E. Long-term efficacy

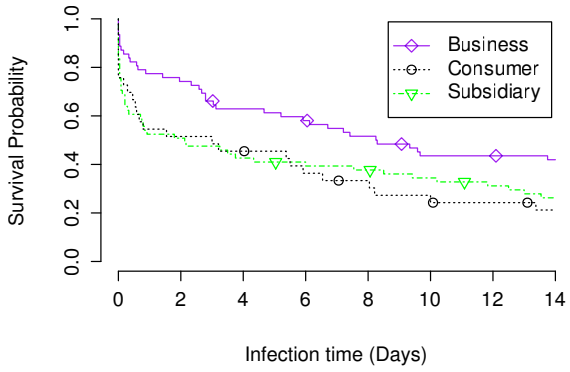
The non-persistent nature of Mirai means that rebooting, shutting down or unplugging an infected device would cause

Table III: Summary statistics of Mirai remediation

| Groups | Sample Size | % clean | Median infection time | Standard deviation |
|---|---|---|---|---|
| Control (Experimental study) | 43 | 74% | 66 Hours | 142.51 |
| Email (Experimental study) | 40 | 77% | 74 hours | 144.18 |
| Walled garden: improved (Experimental study) | 40 | 92% | 26 Hours | 91.64 |
| Walled garden: standard (Observational study) | 97 | 88% | 27 Hours | 121.63 |
| Subsidiary network (Observational study) | 61 | 74% | 51 Hours | 148.02 |
| Business network (Observational study) | 62 | 58% | 198 Hours | 141.64 |




(a) Infection rates for the different treatment variables used during the study




(b) Infection rates across different networks. The consumer network data includes only the control group.

Figure 6: Survival curves of the Mirai infections

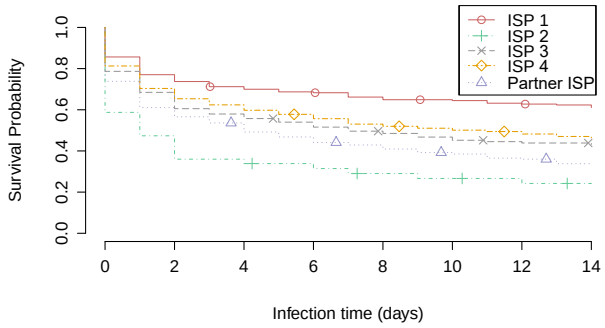



Figure 7: Cleanup rates for 4 randomly chosen ISPs within the country where the partner ISP operates

it to be removed. This fact seems to be an important driver of the high natural remediation rate we observed during the experiment. However, merely rebooting the device does not fix the underlying problem as the device remains vulnerable to infections once it comes back online. To put it differently, the high remediation rates we observed in our experimental and observational study might be Pyrrhic victories if the devices are simply reinfected again soon thereafter. Removing the underlying problem would require affected users to take other actions, such as changing default passwords, updating the firmware or changing router settings – measures that are much more complicated than a mere reboot.

To get a sense of reinfection rates and the long-term efficacy of remediation efforts, we looked at the 97 customers in the observational study. We investigated reinfection rates for this group during a five-month period after the initial 35 days tracking period. We find that only 5 of these customers (5%) were seen again at some point during those five months in the Shadowserver, IoTPOT, or darknet data. In other words: not only is short-term remediation very high, the longer-term reinfection rate also is surprisingly low. This strongly suggests that whatever action the customer took, it was more than a mere reboot of the device. We have asked users about the actions they took and discuss the results in Section VI

On the other hand, intentional action by the user cannot explain the whole story. This is what the high natural remediation rate in the control group tells us. The high remediation rate also contains a signal about low reinfection rates. Remember that to conservatively count them as clean, we tracked the customer IP addresses for an additional 21 days. We did not see these devices again, which clearly means they stopped getting reinfected at some point. In other words, while we might explain the quick removal of Mirai from the combination of non-persistence and device usage patterns, this does not explain why most devices are never seen again. In short, while the low reinfection rate is a positive finding, it is also one for which we have no explanation.

### F. *Impact of device type*

So far we have encountered a number of surprisingly positive results: high remediation rates across all groups, even in the control group, the two reference groups, and low reinfection rates in the months thereafter. To understand if these results are somehow the result of a peculiar composition of device types in these networks, we take a closer look at the affected devices. Is there anything special about them in terms of the cleanup actions or usage patterns?

Following a similar methodology as Antonakakis *et al.* in [1], we have used Censys [9] to determine the device types. We analyzed the banner information obtained through Censys scans and were able to label 88 devices (28%). These devices were mainly network cameras/DVRs (11%), storage units (7.44%) and routers (3.83%). However, the Censys scans

did not allow us to label 72% of the infected devices due to the lack of banner information. In order to increase the number of identified devices, we further conducted port scans on the unidentified devices suing the Network Mapper (Nmap). With this active scanning we gathered banner information of additional ports, i.e, port 5000 (UPnP), 8443 (alternative HTTPS), 32400 (Plex media) and 37777 (QSee DVRs). This allowed us to label 36 additional devices.

Table IV: Type of infected devices per service

| Service | Device type | Amount of Devices | |
|---|---|---|---|
| FTP | NAS | 20 | (7.81 %) |
| | Router | 13 | (5.08 %) |
| | Server | 3 | (1.17 %) |
| | Set top box | 1 | (0.39 %) |
| Telnet | Set top box | 6 | (3.04 %) |
| | DVR | 4 | (1.74 %) |
| HTTP | Camera | 13 | (5.26 %) |
| | DVR | 5 | (2.02 %) |
| | Printer | 4 | (1.79 %) |
| | NAS | 3 | (1.21 %) |
| | Media streamer | 2 | (0.81 %) |
| | Server | 1 | (0.40 %) |
| HTTPS | Media streamer | 3 | (1.35 %) |
| UPnP | NAS | 9 | (3.95 %) |
| Alt. HTTP | Camera | 18 | (7.53 %) |
| | Media streamer | 1 | (0.42 %) |
| | Firewall | 1 | (0.42 %) |
| Alt. HTTPS | Router | 11 | (4.78 %) |
| Plex | Media streamer | 1 | (0.45 %) |
| QSee DVR | DVR | 3 | (1.35 %) |
| | Total identified | 124 | (36.15 %) |
| | Unknown | 219 | (64.04 %) |

Table IV shows the types of devices identified by port. The devices we identified were primarily network-attached storage (NAS) appliances, home routers, cameras, DVRs, printers, and media streamers. This composition of device types is consistent with the composition reported in an earlier study on global Mirai infection [1], suggesting our findings are not driven by selection bias in the types of devices that were affected and remediated.

Device type does seem to influence the infection time. Figure 8 shows the survival curves for the top 5 most common types of devices in our study. The results show that around 50% of the DVRs and cameras remain still infected, while only 20% of the infected routers and NAS appliances were infected after 14 days. While these overall remediation rates per device type seem to indicate that some devices are easier to clean, the survival curves did not show significant differences (log-rank test, $\chi^2 = 7.1$, $p - value = 0.1$).

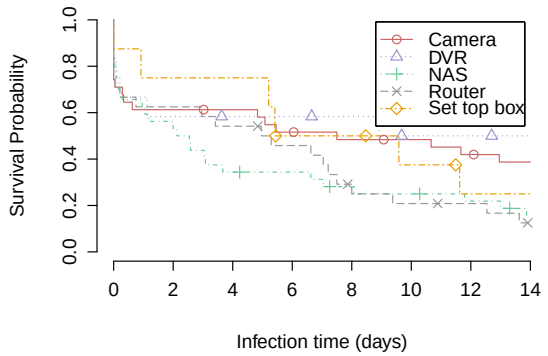



Figure 8: Cleanup rates for the top 5 device types

Interestingly, the composition of device types was different for the business network compared to the two consumer networks (see Figure 9). Routers, security cameras and video-conferencing hardware were more common in the business networks, while storage units and DVR were mainly present in the customer and subsidiary networks. This supports our earlier speculation that the natural remediation rate is indeed tied to the usage patterns of the devices. Remember that the natural remediation rate in the business network was lower. We now see that indeed this concerns a different device population. More of these devices are likely to be always-on for business continuity reasons. If rebooting or unplugging occurs less frequently, there is also less opportunity for natural remediation to occur.

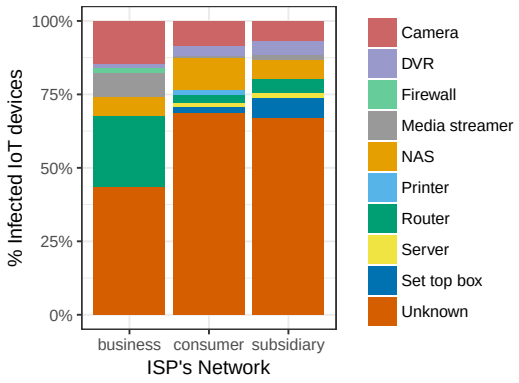



Figure 9: Distribution of device types per network

### G. Lab testing of cleanup and reinfection

In addition to the observational study and the randomized controlled experiment, we also conducted a series of in-lab tests with actual vulnerable devices. These simple tests aim to test the assumption that Mirai malware was indeed not persistent and to also shed some light on reinfection.

The test environment consisted of 7 vulnerable devices (1 IP camera, 1 printer, 1 home router, 3 network storage units, and 1 satellite TV receiver) in their default state (i.e., with their network ports open, and able to accept default credentials). We infected them with a Mirai binary captured by the honeypot. Once infected, we connected the devices to the public Internet and logged all the incoming/outgoing traffic. After malicious outgoing traffic was observed in the infected devices, we rebooted them. Our results showed that after the restart there were no signs of infection in any of these devices: (i) no suspicious process was running after the reboot; and (ii) no malicious communication traffic was observed. However, even though the binary was not running in any of the devices, we did find it in the file system of one of the devices as this device was using a non-volatile storage and the presence of the malware file survived the reboot.

These results are in line with previous studies [31] which also demonstrated the non-persistent nature of Mirai infections. (While [1] did report some persistence, this appears to be related to binaries for X86-64, so non-IoT.) In general, our findings confirm the advice to consumers to reboot the device, though this alone does not resolve the underlying vulnerability. As long as non-persistence is the norm, rebooting will remain effective. As recent as May 2018, the FBI issued a global

alert with the same advice [14] for dealing with a massive population of devices compromised with VPNFilter. Of course all of this, including the high remediation rates we reported earlier in this section, will change when attackers find a way to gain a more persistent foothold on the devices. There are early signs that this is happening [5].

Next, we investigated the reinfection rate, i.e., the time it takes to infect a device, that was cleaned, again. To this end, we connected the devices back to the Internet after rebooting them and monitored the outbound traffic to see whether they get reinfected. We conducted the same procedure three times for each device. Table V shows the average reinfection speed per device. Five out of six devices got reinfected within an hour after being rebooted. This high reinfection rate is consistent with the aggressive scanning behavior of Mirai. One vulnerable device did not get reinfected. A closer analysis of the traffic showed that indeed there were infection attempts but the implementation of the telnet service denied any login attempt for 30 minutes after an unsuccessful login attempt. The aggressive scanning behavior together with the timeout of the telnet service served as an impediment to reinfection.

Table V: Reinfection rate per device type

| Device type | Mean time to reinfection |
|---|---|
| IP camera | No infection for 48 hours |
| Printer | 19min 0sec |
| Router | 1min 50sec |
| NAS 1 | 14min 35sec |
| NAS 2 | 47min 9sec |
| NAS 3 | 37min 47sec |
| Satellite TV Receiver | 5min 35sec |

These results have two implications for our study. First, it underlines the validity of the conservative approach that we took in measuring remediation. Our tracking methodology did not allow us to measure reinfections on a granularity of minutes. This means it is not feasible to distinguish infection from reinfection. It makes more sense to collate the different infection observations over time into a more or less persistent status of being infected.

Second, and more important, this aggressive reinfection behavior means that if we do not see a device for 21 or more consecutive days (our extended tracking period, see Section IV-D), then some remediation action was taken that goes beyond a mere reboot. No vulnerable device with a direct connection to the Internet would survive that long without reinfection.

## VI. User Experiences

Our experimental results show remarkably high remediation rates, especially for the improved walled garden notification. While this is a hopeful result, it is also truly puzzling. We know from prior work that remediation is difficult for end users, even for the more conventional scenario of cleaning up PC-based malware (see related work, Section VII). In this scenario, it is easier for the user to identify the offending device and the ISP can tell the user more precisely what steps she or he needs to take and point to readily available tools to automatically detect and remove the infection. In other words, the notification is much more actionable for the user.

Compared to the conventional scenario, remediating IoT malware seems much more difficult for users. Even in our improved notification we cannot tell the user which of their devices is affected or even what type of device they should look for and disinfect. Next, there are no tools available for disinfection. Finally, remediation actions vary greatly per device type, vendor, local configuration, etc. Absent all of this information, the notification is limited to describing several rather generic actions.

And yet, we find very high clean-up rates – higher, in fact, than the rate for PC infections. We have a direct point of comparison from a prior study conducted recently also at a European mid-sized ISP [6].

The high remediation rate puts a premium on better understanding how users responded to the notification. In this section, we analyze data on the user experience of IoT cleanup collected via phone interviews and the communication logs of the ISP.

### A. Phone interviews

We called 173 customers to invite them to participate in a short telephone interview. This includes all customers in the observational study and the experimental study, except for the customers in the control group and 4 customers who had terminated their contract in the time between the treatment and the interview.

In total, 76 (44%) of the customers accepted the invitation. The response rate was nearly the same in each treatment group. The non-response consisted of customers who did not want to participate (20, 12%), or who could not be reached by phone within several attempts (77, 44%).

Table VI: Respondents receiving and reading the notification

| Experimental group | Total | Received | | Read | | Distrust | |
|---|---|---|---|---|---|---|---|
| Email-only | 16 | 8 | (50.00 %) | 6 | (37.50 %) | 2 | (12.50 %) |
| Walled garden (improved) | 18 | 18 | (100 %) | 18 | (100 %) | 0 | (0.00 %) |
| Walled garden (standard) | 42 | 40 | (95.20 %) | 36 | (85.70 %) | 6 | (14.80 %) |

We first asked participants if they remembered receiving the notification and, if so, if they remembered reading it. Nearly all customers in the walled garden groups remembered receiving it, compared to just around half of the customers in the email-only group. For those customers who did not remember receiving the notifications, we checked whether we used the correct email address. All confirmed it was correct. In other words, the emails likely reached their inbox, but were overlooked (or perhaps got caught in the spam filter). Most of the customers who remembered receiving the message also remembered reading it (See Table VI). Some of the customers who did not read it mentioned that they did not trust the message and wondered whether it was a phishing mail. (One interviewee also did not trust our phone interview and thought it was a Microsoft scam call).

We then asked the 60 customers who remembered reading the notification if he or she took any action and, if so, what action. Four respondents (6.7%) said they did nothing. A further seven (11.7%) said they had called an IT repair service and did not know what this person had done exactly. All others listed doing one or more of the steps mentioned in the notification,

most often mentioning their attempts to identify the offending device. Furthermore, 22 customers (36.7%) specifically stated they had disconnected a device like a camera, DVR or NAS device from the network. One even claimed to have thrown the device in the trash. Also, 22 (36.7%) people mentioned changing the password for one or more devices and 23 (38.3%) said they reset one or more devices. One customer mentioned conducting a firmware update. Four customers reported that they had run an anti-virus scanner. This latter answer signals a misunderstanding of the nature of the infection. We encountered this more frequently in the communication logs, which we discuss below.

Next, we asked whether the customer sought additional help for the problem. Thirteen people (21.7%) mentioned seeking help from another person, such as their relatives or calling the ISP's help desk. Ten people (17%) asked the ISP to send a paid repair person and one person contacted another repair service. Another form of additional help is searching the web. Five people (8.3%) used Google to find additional information and one person mentioned that they consulted the website of the manufacturer of the offending device.

76 respondents were asked how confident they felt in their ability to solve computer security issues like this. Surprisingly, the largest number of people reported to be very confident (34%) or fairly confident (29%). Some of these respondents elaborated on their answer by stating that they had competent people in their environment who they could turn to. On the other end of the spectrum, 17% ranked themselves as not very confident and 18% stated having no confidence at all and little to no knowledge about these issues. Several of these people said they always ask someone else for help. Some of the interviewees stated that they considered themselves too old for these types of problems. We analyzed the correlation between confidence level and cleanup success and found no relationship. It seems confidence, or lack thereof, does not predict remediation outcomes.

We ended the interview by asking all customers how the ISP can improve its communication about these issues with customers. This question revealed wildly different experiences. On the positive side, 17 respondents (22%) explicitly stated being satisfied or even very satisfied with how the ISP handled the situation. A few suggested sending prior warnings before quarantining the connection and to provide more specific information on what to do and what the offending device is. Another suggestion was to provide an option to contact abuse staff during evenings or weekends for customers who cannot self-release from the walled garden. On the negative side, nine respondents (12%) expressed dissatisfaction or anger about the process. The most vocal critics said that they had incurred economic losses as they were running small businesses on their consumer subscription which were interrupted by the quarantine event.

### B. Communication logs

Additional insights into the user experience of IoT cleanup were extracted from the communication logs between the help desk and the customers in the study, except for those in the control group. In total, we found one or more messages for 92% of these customers in the ISP's logs. We investigated 159 walled garden contact forms (from 90 unique customers), 404 emails (from 106 unique customers) and 117 help desk logs (from 68 unique customers).

First, we explored the distribution of messages across the different treatment groups (See Table VII). We found that about a third of the customers replied to the email notification and only 3 customers contacted the help desk. This rate is much higher for the walled garden groups: around 50% of the quarantined customers called the help desk. While less communication is cheaper and improves the incentives of ISPs for cleanup, it seems that the rate of seeking help is related to action on the side of the customers. As we saw in Section V-A, the remediation rate of the email-only group was indistinguishable from the control group. The walled garden groups did take action and this is also associated with more communication with the ISP.

Table VII: Communication channel used by customers in different groups

| Experimental group | n | email | contact form | helpdesk |
|---|---|---|---|---|
| Email-only | 40 | 16 (40.0%) | – | 3 ( 7.5%) |
| Walled garden (improved) | 40 | 23 (57.5%) | 31 (77.5%) | 21 (52.5%) |
| Walled garden (standard) | 97 | 67 (69.1%) | 59 (60.8%) | 44 (45.4%) |

Next, we read a sample of about 20% of messages in each category and created labels for recurring themes. We then read all messages and manually labeled each one as to whether a certain theme was present in it or not. Table VIII presents the results aggregated over all customers, i.e., whether a theme was present in one of the messages of a customer. The general pattern confirms what we found during the phone interviews. Some issues are more salient, though. In the walled garden treatments, about one in three customers states that they have run an anti-virus scanner on their PC to remediate the problem. This underlines, even more than the phone interviews, that a significant portion of affected population does not understand the basic properties of IoT malware, even when they have actually seen and read the notification. We found a weak correlation with remediation: customers who mention running anti-virus remediated more slowly. Around 60% was clean after five days, whereas 60% of the other customers was clean within little more than one day. That being said, both groups reached 90% remediation in two weeks.

While a significant portion of the users is working from an incorrect mental model ('folk theory' [34]) of the problem, they do seem to be able to remediate in the end. Of the 51 customers that mentioned running a virus scanner, 23 also mentioned disconnecting a device. Proportionally, this rate is actually a bit higher than for the people who did not mention running anti-virus. Overall, around 40% of the customers in the walled garden groups mention that they disconnected a device, compared to just 7.5% for those who received the email-only notification.

In the improved walled garden group, dissatisfaction or frustration is substantially lower than in the standard walled garden group. We are not sure how to explain this. It might be that the improved message is more helpful. We should note, however, that the improved walled garden notifications were issued several months later in time than the standard notifications. By that time, more people might have seen

reports in the media about IoT compromise and they might thus be more accepting of the need to take countermeasures.

Table VIII: Themes of user experience in communication with the ISP

| | Email-only | Walled garden (improved) | Walled garden (standard) |
|---|---|---|---|
| | n=40 | n=40 | n=97 |
| Runs a virus scanner | 7 (17.5%) | 12 (30.0%) | 32 (33.0%) |
| Identifies IoT device | 9 (22.5%) | 17 (42.5%) | 58 (59.8%) |
| Requests additional help | 2 ( 5.0%) | 8 (20.0%) | 41 (42.3%) |
| Wants possibility to call the abuse team | 0 ( 0.0%) | 2 (5.0%) | 16 (16.5%) |
| Requests paid technician | 0 ( 0.0%) | 4 (10.0%) | 11 (11.3%) |
| Disconnects device | 3 ( 7.5%) | 15 (37.5%) | 42 (43.3%) |
| Cannot work due to quarantine | 0 (0%) | 4 (10.0%) | 18 (18.6%) |
| Complaints over disruption of service | 0 (0%) | 1 (2.5%) | 13 (13.4%) |
| Threatens to terminate contract | 0 (0%) | 1 (2.5%) | 5 (5.2%) |

All in all, the customer experience data helps us to make sense of the high remediation rates for the walled garden groups. While users might not grasp the technical foundations of the infection, as signaled by running AV on a PC in their network, they still end up taking effective action. Disconnecting devices is an intuitive countermeasure, after all, even if it is also costly on the side of the customer – in the sense of not being able to use the device.

It is tempting to speculate about how these customer responses might help explain the remarkably low reinfection rate of the the standard walled garden group (see Section V-E). One might reason, for example, that these users either keep the devices disconnected over a longer period or that they reconnect them differently than before, leaving them no longer exposed to the public Internet. Another explanation is that they factory reset their router, which for certain models means closing open ports and disabling the DMZ and uPnP. This leaves the user in a less vulnerable state.

In the end, though, these speculations seem somewhat beside the point. Remember, even the control group had a low reinfection rate, in the sense that most of the customers in that group were not seen again for at least 21 days after their initial infection disappeared (see Section V). Whatever the explanation is for this result, it could very well also explain the bulk of the low reinfection rates in the other groups in the study, rather than intentional remediation actions on the side of the users in those groups. For now, we are stuck with a mystery that future work will have to resolve.

## VII. Related Work

In this section, we briefly review three related areas of work. We survey studies on botnet mitigation by ISPs, efficacy of abuse notifications and end user security behavior.

**Botnet mitigation by ISPs:** Various studies have looked into the role of ISPs in the fight against botnets mitigation and remediation. Most notably, [3] empirically confirmed the point that ISPs are indeed critical control points for botnet mitigation and that infection levels are very different across ISPs, even when they operate in the same country and market, demonstrating they have leeway to act. Work on Conficker cleanup [2] found no clear impact of national initiatives to mitigate botnets.

Additionally, industry groups and international organizations have published ISP best practices that explain how to contact and clean up infected customer's machines. RFC 6561 describes various methods that can be used by ISPs to notify end users about a security problem [21]. Some of the described methods include postal mail, email, phone or walled garden notifications. On the other hand, the effectiveness of the methods is not discussed in detail. A report outlined by M3AAWG identifies best practices for walled garden notification. However, ITUs Anti-Botnet Toolkit raised potential issues that may result from ISP notifications [17].

In an earlier study, Çetin *et al.* investigated the usability of walled garden notifications [6]. This study mainly focused on regular malware infections of PCs while presenting a simple comparison on remediation rates per malware type which also includes Mirai. This study was purely observational and as such the authors could not measure the effectiveness of walled garden notifications, but instead analyzed users' behavior while in quarantine. Nevertheless, the authors reported overall remediation rates as observed for the whole system which cannot be solely attributed to the effect of the walled garden. Authors found that were roughly over 85% of Mirai-infected machines were cleaned after 2 weeks period, which is a bit lower than both standard and improved walled garden notifications for Mirai-infected customers observed in the current study. On the contrary, we focused on analyzing the actual impact of the walled garden by designing an experiment with a control group which allowed us to estimate the efficacy of the walled garden notifications on their own. Moreover, our study is specific to Mirai-infected devices which allowed us to customize the content of the notifications with IoT-specific cleanup instructions.

**Efficacy of abuse and vulnerability notifications:** A large body of research focuses on efficacy of email notifications on large scale vulnerability notifications. For instance, Li *et al.* issued various types of notifications to CERTs and operators of networks [19]. They concluded that detailed notifications to operators made the highest impact. On the other hand, their results suggested that overall vulnerability remediation was marginal, even with detailed notifications to operators. Similar to this work, Stock *et al.* studied the feasibility of large-scale vulnerability notification and found that notified parties achieved higher remediation rates than the ones that received no notifications [28]. Additionally, in another study on large scale Cetin *et al.* demonstrated the poor deliverability of email-based notifications and proposed searching for other mechanisms to deliver notifications [7]. Stock *et al.* evaluated the effectiveness of other mechanisms to delivery vulnerability information such as postal mail, social media, and phone and reported slightly higher remediation rates for these mechanisms [27]. On the other hand, they stated that slightly higher remediation do not justify their costs and additional work put into issuing them. Majority of these studies used email to reach affected parties. Because it scales reasonably well. Conversely, many emails bounced before even reaching the affected parties. Moreover, the ones that reached often triggered no follow-up actions.

Another series of studies explored the efficacy of email notifications on abuse remediation. These notifications are sent to the affected owners of the site or to their hosting

provider. Li *et al.* assess the influence of abuse notifications for $761,935$ infected websites detected by Google Safe Browsing and Search Quality [20]. Direct notifications to webmasters increased the likelihood of cleanup by over 50%. Vasek *et al.* found that verbose notifications to webmasters and hosting providers were the most effective [32]. Cetin *et al.* studied the effect of reputation of the sender of the abuse notification on cleanup rates. While notifications in general improved cleanup, there was no observable effect of the sender reputation [8].

Results of these website cleanup studies indicate a much lower remediation rate than that observed in this study. This could be partly because of Mirai's non-persistent nature.

**End user security behavior:** A large body of work has studied the challenges of end users in obtaining and following security advice. A study on end user perceptions on automated software updates concluded that most users do not correctly understand the automatic update settings on their computer and thus cannot manage to update as they intend to [33]. Fagan *et al.* [13] investigated user motivations regarding their decisions on following common security advice. They reported that the majority of users follow the usability/security trade-off. Forget *et al.* collected data on users' behavior and their machine configurations and highlighted the importance of content, presentation, and functionality of security notifications provided to users who have different expertise, expectations, and computer security engagement [15]. This work demonstrated the importance of effective communication between customers and the ISP. This can help to ensure a better understanding of the notifications and a higher rate of remediation.

## VIII. Ethical Considerations

This study leveraged passively collected datasets and a small number of active scans that were carried out following the guidelines of the Menlo report [4]. All raw data and statistics generated during the study were anonymized, and only the partner ISP's employees knew what customer corresponded to which infection. We always followed the policies of the ISP and notified all the infected subscribers accordingly. We only added the experimental design of random assignment and the observation in abuse feed and darknet data of the infected devices. The latter is not regarded as human subject research by our IRB and thus out of scope. For the purpose of the experiment, the customers in the control group received the notification with a delay of 14 days. Moreover, during the phone interviews, interviewees were provided with an opt-out option. Throughout the interview process, only 20 interviewees asked to be excluded from the interviews.

## IX. Limitations

Our study faces three key limitations. First, detecting and tracking infections is difficult. No method detects all infected machines and when tracking a detected infection there will also be missing observations complicating inferences about cleanup. The former issue is less of a problem for our study, as our design is not based on capturing all infections. The latter issue we mitigated by adopting a very conservative approach in measuring cleanup. If we saw the same customer again within 21 days after the experiment, we would assume they were not cleaned up, irrespective of the missing observations in between. This gives us a lower-bound estimate of cleanup.

Second, the external validity of this research project is open to discussion. On the one hand, the study is conducted in a real-world setting within normal business processes. In addition, the ISP is the second largest in the country and has several million broadband customers. They represent a wide variety of people in terms of demographics. Therefore, we have no reason to assume that our findings are particular to this ISP. On the other hand, it is impossible to know to what extent a walled garden mechanism at another ISP would get the same results until follow-up experiments are conducted.

Last, the dynamic nature of malware limits the generalizability of our findings. Our results are based on Mirai. During the study period, new Mirai versions and other IoT malware families were still non-persistent. This greatly increases natural cleanup via rebooting of devices and it also facilitates cleanup by end users. As IoT malware becomes more sophisticated, it seems a matter of time before they are able to establish a more permanent foothold on the device. Indeed, a recent study reported the first persistent IoT malware [5]. We expect this to cause lower remediation rates.

## X. Conclusion

We have presented the first empirical study on the cleanup of IoT malware in the wild. We found that quarantining and notifying infected customers via a walled garden remediates 92% of the infections within 14 days. Email-only notifications have no observable impact. We also found high natural remediation rates and low reinfection rates. We have no good explanation for the low reinfection rate, though we are quite confident the result itself is correct. While quarantining infected devices is clearly highly effective, future work will have to resolve these remaining mysteries.

At first glance, the implications of our study for industry seem clear. First, ISPs have a critical role to play as more than 87% of the infections reside in their networks. Second, walled garden notifications work and are feasible, even though the actual usability of the notification and cleanup advice is currently rather poor. Third, since walled gardens are a recognized best practice for 'regular' botnet mitigation by ISPs, we can leverage the existing mitigation structures and practices to also help mitigate IoT botnets, rather than having to go through the time-consuming path of setting up new organizational structures and agreements.

There is a 'but', however. A significant one. The economic incentives for ISPs to adopt walled garden solutions are rather weak, as evidenced by the fact that only a fraction of the ISPs currently have them. Setting up and operating a walled garden, or operating any effective abuse management process in general, is a cost center for the ISP. Further eroding the incentives is the fear of customer pushback. Our analysis of customer experiences did indeed uncover a small but vocal minority that was angry or frustrated. Given the high cost of customer acquisition in these saturated markets, this fear might be enough to dissuade ISPs from quarantining infections. Overcoming this incentive problem might require a governmental measure to assign intermediate liability to ISPs. Soft versions hereof – e.g., a so-called 'duty of care' – already exist in many jurisdictions [10], [30]. While calling upon ISPs to take on this task, we can point out that their actions will have much higher

chance of success than educating millions of end users about IoT security. Also, we can point to the fact a non-trivial portion of customers was pleased to be notified via the walled garden. As more people will become aware of the threats to their IoT devices, ISP mitigation might become more accepted – or even expected.

## ACKNOWLEDGMENT

This publication was supported by a grant from the Netherlands Organisation for Scientific Research (NWO), under project number 628.001.022, 628.001.033 and by WarpDrive project funded by NICT, Japan. Also, we would like to thank the anonymous reviewers, Calvin Brierley, Jamie Pont, Darren Hurley-Smith, Folkert Visser, Dennis van Beusekom and Burcu Açar for their helpful comments and support.

## REFERENCES

[1] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the Mirai Botnet," in *26th USENIX Security Symposium (USENIX Security 17)*. Vancouver, BC: USENIX Association, 2017, pp. 1093–1110.

[2] H. Asghari, M. Ciere, and M. J. Van Eeten, "Post-mortem of a zombie: conficker cleanup after six years," in *USENIX Security Symposium*. USENIX Association, 2015, pp. 1–16.

[3] H. Asghari, M. J. van Eeten, and J. M. Bauer, "Economics of fighting botnets: Lessons from a decade of mitigation," *IEEE Security & Privacy*, vol. 13, no. 5, pp. 16–23, 2015.

[4] M. Bailey, D. Dittrich, E. Kenneally, and D. Maughan, "The Menlo Report," *IEEE Security & Privacy*, vol. 10, no. 2, pp. 71–75, 2012.

[5] B. Botezatu, "Hide and Seek IoT Botnet resurfaces with new tricks, persistence," Bitdefender Labs, 2018. [Online]. Available: https://labs.bitdefender.com/2018/05/hide-and-seek-iot-botnet-resurfaces-with-new-tricks-persistence/

[6] O. Çetin, C. Gañán, L. Altena, and M. van Eeten, "Let me out! evaluating the effectiveness of quarantining compromised users in walled gardens," in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, 2018.

[7] O. Cetin, C. Ganán, M. Korczynski, and M. van Eeten, "Make notifications great again: learning how to notify in the age of large-scale vulnerability scanning," in *16th Workshop on the Economics of Information Security (WEIS 2017)*, 2017.

[8] O. Cetin, M. Hanif Jhaveri, C. Gañán, M. van Eeten, and T. Moore, "Understanding the role of sender reputation in abuse reporting and cleanup," *Journal of Cybersecurity*, vol. 2, no. 1, pp. 83–98, 2016.

[9] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, "A Search Engine Backed by Internet-Wide Scanning," in *22nd ACM Conference on Computer and Communications Security*, Oct. 2015.

[10] K. K. e Silva, "How industry can help us fight against botnets: notes on regulating private-sector intervention," *International Review of Law, Computers & Technology*, vol. 31, no. 1, pp. 105–130, 2017.

[11] ECO Internet industry association, "Botfree," 2013. [Online]. Available: https://www.botfree.eu/en/aboutus/information.html

[12] European Network and Information Security Agency (ENISA), "Involving Intermediaries in Cyber-security Awareness Raising," 2012. [Online]. Available: https://www.enisa.europa.eu/publications/involving-intermediaries-in-cyber-security-awareness-raising

[13] M. Fagan and M. M. H. Khan, "Why do they do what they do?: A study of what motivates users to (not) follow computer security advice," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, 2016, pp. 59–75.

[14] Federal Bureau of Investigation (FBI), "Foreign cyber actors target home and office routers and networked devices worldwide," 2018, https://www.ic3.gov/media/2018/180525.aspx.

[15] A. Forget, S. Pearman, J. Thomas, A. Acquisti, N. Christin, L. F. Cranor, S. Egelman, M. Harbach, and R. Telang, "Do or do not, there is no try: user engagement may not improve security outcomes," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, 2016, pp. 97–111.

[16] D. Holmes, "The Mirai Botnet Is Attacking Again)," 2018. [Online]. Available: https://www.darkreading.com/partner-perspectives/f5/the-mirai-botnet-is-attacking-again/a/d-id/1331031

[17] International Telecommunication Union (ITU), "ITU Botnet Mitigation Toolkit," https://www.itu.int/ITU-D/cyb/cybersecurity/projects/botnet.html, 2018.

[18] U. Jilani, "The ACMA and Internet providers working together to combat malware," 2015. [Online]. Available: https://www.acma.gov.au/theACMA/engage-blogs/engage-blogs/Cybersecurity/The-ACMA-and-internet-providers-working-together-to-combat-malware

[19] F. Li, Z. Durumeric, J. Czyz, M. Karami, M. Bailey, D. McCoy, S. Savage, and V. Paxson, "You've got vulnerability: Exploring effective vulnerability notifications," in *25th USENIX Security Symposium (USENIX Security 16)*. Austin, TX: USENIX Association, 2016, pp. 1033–1050.

[20] F. Li, G. Ho, E. Kuan, Y. Niu, L. Ballard, K. Thomas, E. Bursztein, and V. Paxson, "Remedying Web Hijacking: Notification Effectiveness and Webmaster Comprehension," in *Proceedings of the 25th Int. Conference on WWW*, 2016, pp. 1009–1019.

[21] J. Livingood, N. Mody, and M. OReirdan, "Recommendations for the Remediation of Bots in ISP Networks (RFC 6561)," *Internet Eng. Task Force*, 2012.

[22] G. F. Lyon, *Nmap network scanning: The official Nmap project guide to network discovery and security scanning*. Insecure, 2009.

[23] Messaging Anti-Abuse Working Group and others, "Abuse Desk Common Practices," 2007. [Online]. Available: https://www.m3aawg.org/sites/default/files/document/MAAWG_Abuse_Desk_Common_Practices.pdf

[24] ——, "M3AAWG best practices for the use of a walled garden," 2015. [Online]. Available: https://www.m3aawg.org/documents/en/m3aawg-best-common-practices-use-walled-garden-version-20

[25] Y. M. P. Pa, S. Suzuki, K. Yoshioka, T. Matsumoto, T. Kasama, and C. Rossow, "IoTPOT: Analysing the Rise of IoT Compromises," in *9th USENIX Workshop on Offensive Technologies (WOOT 15)*. Washington, D.C.: USENIX Association, 2015.

[26] Shadowserver Foundation, "Shadowserver reports," 2018. [Online]. Available: https://www.shadowserver.org/wiki/pmwiki.php/Services/Reports

[27] B. Stock, G. Pellegrino, F. Li, M. Backes, and C. Rossow, "Didnt You Hear Me?–Towards More Successful Web Vulnerability Notifications," in *The Network and Distributed System Security Symposium (NDSS)*, 2018.

[28] B. Stock, G. Pellegrino, C. Rossow, M. Johns, and M. Backes, "Hey, you have a problem: On the feasibility of large-scale web vulnerability notification," in *USENIX Security Symposium (Aug. 2016)*, 2016.

[29] TeleGeography, "Telegeography Globalcomms Data," 2017. [Online]. Available: http://shop.telegeography.com/products/globalcomms-database

[30] E. Tjong Tjin Tai, B.-J. Koops, D. Op Heij, K. Silva, and I. Škorvánek, "Duties of care and diligence against cybercrime," 2017.

[31] J. B. Ullrich, "An Update On DVR Malware: A DVR Torture Chamber," SANS Technology Institute, Tech. Rep., 2017. [Online]. Available: https://isc.sans.edu/forums/diary/An+Update+On+DVR+Malware+A+DVR+Torture+Chamber/22762/

[32] M. Vasek and T. Moore, "Do malware reports expedite cleanup? an experimental study," in *5th Workshop on Cyber Security Experimentation and Test, CSET*, 2012.

[33] R. Wash, E. Rader, K. Vaniea, and M. Rizor, "Out of the loop: How automated software updates cause unintended security consequences," in *Symposium on Usable Privacy and Security (SOUPS)*, 2014, pp. 89–104.

[34] R. Wash and E. J. Rader, "Too Much Knowledge? Security Beliefs and Protective Behaviors Among United States Internet Users," in *Eleventh Symposium On Usable Privacy and Security, SOUPS 2015, Ottawa, Canada, July 22-24, 2015.*, 2015, pp. 309–325.

## APPENDIX A
## STANDARD WALLED GARDEN NOTIFICATION CONTENT

**Secure environment**

A safe Internet is in everyones interest. We strongly care about protecting your (confidential) information.

We have received information from one of our partners that a security issue has been detected on your Internet connection. You probably have not noticed anything yet.

Dont worry. To protect you against the security risks we have placed your Internet connection in our secure environment. In this environment you can safely solve the security issues. We are willing to help you to do so.

**What is the problem and how can you solve it?**

One or more devices connected to your Internet connections are infected with the Mirai-virus. This virus targets devices that make use of your Internet connection independently. In most cases IP Cameras or Digital TV decoders.

The infection probably occurred due to the use of a standard password username combination to access the device.

To solve this problem please reset all your devices to factory defaults. After the reset change all the passwords for accessing the devices to strong passwords.

In case the device can be reached by Telnet or SSH please also change these passwords.

**Necessary steps**

1. Take the measures stated above

2. Fill in our form (and restore your Internet connection)

**General security tips**

* Use an up-to-date virus scanner to keep out potential hazards

* Keep computer software, like your operating system, up to date

* Do not open messages and unknown files that you do not expect or trust

* Secure your wireless connection with a unique and strong password

## APPENDIX B
## IMPROVED WALLED GARDEN NOTIFICATION CONTENT

**Secure environment**

A safe Internet is in everyones interest. We strongly care about protecting your (confidential) information.

We have received information from one of our partners that a security issue has been detected on your Internet connection. You probably have not noticed anything yet.

Dont worry. To protect you against the security risks we have placed your Internet connection in our secure environment. In this environment you can safely solve the security issues. We are willing to help you to do so. What is the problem and how can you solve it?

One or more Internet connected devices in your home have been infected with the Mirai virus. We cannot detect which Internet connected device has been infected. Most likely it is a digital video recorder (DVR), security camera or a printer connected to the Internet rather than a computer, laptop, tablet or mobile phone.

What should you do to remove the Mirai virus and prevent future infections? Please follow the steps below. If you cannot complete a step, please proceed to the next one.

1. Determine which devices are connected to your Internet connection. Reminder: The Mirai virus mainly infects Internet connected devices such as a DVR, security camera or printer connected to the Internet.

2. Change the password of the Internet connected devices. Choose a password that is hard to guess. If you do not know the current password, please refer to the manual.

By following these steps, you have prevented future infections.

3. Restart the Internet connected devices by turning it off and on again. Hereafter, the Mirai virus has been removed from the memory of the devices. Now that your Internet connected devices are safe, the last steps are to protect your router/ modem.

4. Reset your modem/router to the factory settings. On https://address.com it is described how you do this for an Experia Box.

5. Set the password of your modem/router. On https://address.com it is described how you do this for an Experia box.

Warning! If remote access to a certain device is absolutely necessary, manually define port forwards in your router for this device. On https://address.com it is described how you do this for an Experia Box.

**Necessary steps**

1. Take the measures stated above

2. Fill in our form (and restore your Internet connection)

**General security tips**

* Use an up-to-date virus scanner to keep out potential hazards

* Keep computer software, like your operating system, up to date

* Do not open messages and unknown files that you do not expect or trust

* Secure your wireless connection with a unique and strong password