A VISIT TO THE CRIME SCENE

MASTER THESIS

BRENNEN BOUWMEESTER

A Visit to the Crime Scene

Monitoring end-users during the remediation process of Mirai infected Internet of Things devices

Master thesis submitted to Delft University of Technology
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE
in Engineering and Policy Analysis

Faculty of Technology, Policy and Management
to be defended publicly on September 10th 2020

by

Brennen Joel Bouwmeester
Student number: 4446461

Graduation committee:

Chairperson	Prof. M.J.G. (Michel) van Eeten, section Organisation & Governance
First supervisor	Ir. A. (Arman) Noroozian, section Organisation & Governance
Second supervisor	Dr. I. (Iulia) Lefter, section Systems Engineering and Simulation
Internal supervisor	Ir. E.R. (Elsa) Turcios Rodríguez, section Organisation & Governance
External supervisor	D.W.J. (Dennis) van Beusekom, Abuse Desk KPN

This page was intentionally left blank

Acknowledgements

Only five years ago I stood at the doorstep of the faculty of Technology, Policy and Management, ready to take on the world. I learned to model this world to make it a better place and to think in ways that help me in life every single day. I am grateful for the opportunities that I got to gain such knowledge. Although my name is stated in the lonely middle of this final document that closes the door of an era for me, this is a result that carries many more names.

First of all, I am very grateful to my committee from Delft for allowing me to be part of their very interesting, constantly relevant and always educational research team. I could not have wished for a better team. Elsa, thank you for your infinite empathy, patience and friendliness. Arman, thank you for your everlasting support, wisdom and ability to relativize. Michel, thank you for sharing your great experience, creativity and calming trust. Iulia, thank you for putting things into perspective and your fresh look on my thesis journey.

I would like to thank my family and friends for having faith in me; you make me want to become the very best self I could be. Mom, dad, Rowin, you bring out the best in me. I would like to give a special thanks to my partner in crime for the past five years at TU Delft. Kevin, there is no more iconic duo. Also, my special thanks go to Marleen. You are a constant energy and motivation boost. I could not have done this without you.

Last, but certainly not least, I would like to thank the members of the KPN Abuse Desk for making this research possible, useful and above all memorable. I have often wondered things would have been like if I spent the past months with you at the headquarters, as the first few visits there were very valuable to me. However, even though this was not possible, you still did not fail to make me feel welcome and guided me through the good times and the bad. I had an awesome time and always started the day with a smile and a laugh during the daily start-up meetings, for which I am grateful. I want to thank Virgil and Raymond for being my first line of defence, who were always ready and excited to help me and I want to thank Dennis van Beusekom for being my guidance from day one. It has been a pleasure.

Summary

The increasingly important availability of online services is constantly threatened by malicious software such as botnets. Attackers have gained power through devices that are part of the rising Internet of Things (IoT), mostly through infections caused by Mirai. The botnets created by Mirai are used for the purpose of DDoS attacks, which can take away the availability of an online service. Although Mirai can be detected relatively easily due to its superficial signature, the remediation process of Mirai infected IoT devices runs far from smoothly.

As end-users often do not notice the presence of Mirai and manufacturers lack incentives to invest in better security or support, ISPs like KPN are amongst the few viable actors that could defend against botnets like Mirai. As ISPs are able to link infection feeds to their customers, they are able to send out notifications accompanied by protocols that can resolve Mirai infections when executed properly. Although research exists on the remediation rates, it is not clear what processes take place at end-users homes during the remediation process and what critical points of error exist throughout the phases of the anti-botnet cycle. As the remediation rate of Mirai infections is currently only 60 – 76 percent, it can be worth looking into the remediation processes to see where they could be improved.

The main research question is the following: "*What do we learn about how and to what extent Internet Service Providers can improve the remediation process of malware infected Internet of Things devices by monitoring end-users while they are cleaning their Mirai infections?*". To answer this question, we have closely followed 17 Mirai infected end-users over a period of 7 weeks, at the KPN Abuse Desk, after a 1 week pilot phase to test our email notification and think aloud protocol. We have prepared and analyzed all steps from identification of a Mirai infection until successful remediation took place. The lion's share of this experiment is about a virtual visit; a phone call with an option to upgrade to a video conference, in which infected end-users get advanced support in performing the 5 cleaning steps stated in the protocol they received. As the end-users thought aloud during the calls, we were able to follow them closely and pinpoint arising issues. Using a thematic content analysis, we synthesized the personal stories that end-users shared.

During the 7 week experiment, we saw 37 unique IP addresses infected with Mirai, of which 12 were excluded due to the ISP policy of not providing support during the weekends. Of the 25 remaining IP addresses, 3 could not be notified due to technical issues within KPN, 2 did not pick up the phone after being notified and 3 were not willing to take part in the experiment due to trust issues.

16 out of 17 participants that were responsible for the internet security were male, but their varying household sizes shows that this does not relate to men becoming infected more often. The age of the end-users was normally distributed between 21 and 80 and we found a household size of 1 to 6, excluding 3 small business locations that became victim of Mirai. End-users can often only identify 1 or 2 IoT devices in their network (13 out of 17) and are almost always able to pinpoint the infected device (16 out of 17). Many issues arose during the virtual visits, such as a lack of trust, not willing to spend effort, a lack of support by the manufacturer, or the idea that regular protection measures should have protected against Mirai.

Only 6 out of 17 end-users were able to perform all steps successfully. In most cases end-users failed to change the password of their device or performed a regular reset on their router instead of a factory reset. This caused 3 failing remediation efforts and 5 reinfections during the experiment phase. The remediation process has barriers in each phase that could be addressed. The biggest improvement can be made in the awareness of end-users, which would lead to higher prevention of infections. Prevention would keep the many potential issues in the remediation process from arising altogether.

Contents

List of tables

List of figures

List of abbreviations

CEFR – Common European Framework of Reference of Languages

CISO – Chief Information Security Office

DDoS attack – Distributed Denial-of-Service Attack

DMZ – Demilitarized Zone

DVR – Digital Video Recorder

IoT – Internet of Things

ISP – Internet Service Provider

NAS – Network Attached Storage

SDG – Sustainable Development Goals

SOC – Security Operation Centre

TAP – Think aloud protocol

TCA – Thematic Content Analysis

UPnP – Universal Plug and Play

URL – Uniform Resource Locator

VPN – Virtual Private Network

1. Introduction

1.1 Background

1.1.1 Attacking availability of services

The availability of services is becoming ever more important, as an increasing number of actions takes place online. In security, availability is part of the CIA (Confidentiality, Integrity and Availability) triad, that should be guaranteed at all times (Perrin, 2008). Distributed Denial-of-Service (DDoS) attacks are used to keep services from staying alive, by overloading the capacity of the victim's service. This then disrupts the regular traffic that takes place and make the service unavailable, often leading to profit loss. Next to monetary loss, also damage is done to the reputation of the victim and to customer satisfaction (Cardoso de Santanna, 2017; Chromik, Cardoso de Santanna, Sperotto, & Pras, 2015; Cheung, 2017). DDoS attacks are considered to be a main threat in the online world (Holl, 2015; Cheung, 2017), which is where about 50% of all crimes takes place (Blythe & Johnson, 2019).

Although DDoS attacks are not new, emerging services such as *Booters*, *DDoS-as-a-service* or *DDoS-for-hire* have revived and strengthened the security issue. As is stated by Karami and McCoy (2013), thousands of DDoS attacks are provided by DDoS as a service. These days, anyone can invoke an attack by paying an affordable price, for example using the so-called Booter websites, which are easily accessible, according to Cardoso de Santanna (2017). By removing the technical barrier of performing attacks, the launching of an attack has become available to a much wider audience, which increases the number of occurrences in a fast tempo (Chromik et al, 2015; Cardoso de Santanna, 2017). Next to the low barrier, also the power of DDoS attacks has increased due to amplification techniques (Cardoso de Santanna, 2017; Noroozian et al., 2016; Rossow, 2014).

1.1.2 The ever-increasing army

The main resource that is needed for attackers to be able to launch DDoS attacks exists of an army of soldiers that operates in a coordinated and effective manner. An attacker needs control over these so-called robot networks (botnets), which can then be asked to send requests to a certain domain collectively. Due to the 4th industrial revolution and rise of the Internet of Things (Evans, 2011), an increasing number of potential bots exists. According to Silverio-Fernández et al. (2018), a device is seen as a “smart device” or an Internet of Things device if:

- The device can connect to a network to share data with other devices (*connectivity*)
- The device can perceive information from its environment (*context-awareness)*
- The device can perform tasks without any user interference (*autonomy)*

In contrast to the former regular connected devices, such as desktop computers or laptops, IoT devices often lack in sufficient security (Alaba et al., 2017; European Union & Agency for Network and Information Security, 2017), which makes them easy targets for malicious actors (attackers). As only more and more devices will become “smart”, the size of existing botnets also potentially increases and thus the problem of DDoS attacks. One of the largest fish in the sea of botnets is Mirai (meaning “Future”), which took over the internet in late 2016 with a peak total number of infections of over 600,000 (Antonakakis et al., 2017). Despite of the simplicity of the way in which Mirai scans the internet for vulnerable devices and infects them, it has been a significant online threat for the past 4 years, which is unlikely to change anytime soon.

1.1.3 Homes as crime scenes

As botnets are the main resource of DDoS attacks, it could be stated that the source of the problem lies with the owners of the infected devices. However, these end-users often do not notice any negative effects of their infected devices themselves (Van Eeten & Bauer 2008). Therefore, they do not have any incentive to clean their devices. In essence, this makes them an accomplice to a crime without their consent or knowing.

Even when users are notified about their infected Internet of Things device, it is relatively unknown in what way these users react to this notification and to what extent they know how to resolve the problem (Altena, 2018; Verstegen, 2019). Internet service providers (ISPs), which are seen as the actors with the best position to intervene (Van Eeten & Bauer, 2008), experiment with different techniques to notify users, but are struggling to find the best way of doing so. Sending a notification has multiple variables that can be adjusted to optimize the effect, such as the sender of the notification or the channel through which the notification is sent. However, in finding the best way to notify, there are both ethical and technical boundaries.

1.1.4 Complex system

Although the problem at hand might seem trivial, the nature of the system complicates matters significantly. The existence of a botnet such as Mirai starts with the manufacturing of IoT devices, which are then shipped, bought by retailers and later by consumers. By then, product labels, retail shops, governments, Internet Service Providers and even customs could have put constraints and minimal security requirements on the products. After a device has been infected, it is also unclear who carries the responsibility for cleaning the device and whether having an infected device should be punishable in the first place. Therefore, incentives for solving the problem do not align.

Next to the multi-actor nature of the system, also the technical domain and the social domain collide. As it is human interaction with the devices that creates weaknesses in devices and a lack of incentive that keeps these devices from being cleaned, it is important to map the system considering both social and technology aspects.

To make things even more complicated, there is an inherent of information asymmetry between the different parties in the arena. Attackers have knowledge about their bots, which can only be discovered in a later stage by whistle-blowers, who inform ISPs. On their turn, ISPs have information about commonly infected devices, but not about the present devices of their infected customers. This causes issues and time delays in finding out about infections in the first place, but also in moving the end-users towards cleaning the infected devices.

1.2 Problem statement

While some ISPs have a thought through setup for notifying end-users about their infected IoT devices, the remediation rate can still be improved. Also, there is still little knowledge about what processes actually take place in end-users' homes after receiving a notification. As ISPs have little information about these processes and about end-users' home network in general, it remains a challenge to supply the users with the right knowledge and information to be able to execute the needed steps for remediation. Although stated reactions from users have been researched before (Altena, 2018; Verstegen, 2019) by making use of the following 5 cleaning steps, a difference exists between this stated behaviour and the actual behaviour:

1. Identify the devices that are connected to the Internet
2. Reset the device(s)
3. Change the passwords of the device(s)
4. Reset the modem/router (back to factory settings)
5. Change the password of the modem/router.

It is important to find out exactly in which of the following stages of the protocol users tend to misunderstand or lack knowledge to perform these required steps. Only by gathering detailed data on their actions, ISPs can get a better view on what sort of information might be misleading or missing from the notification. Moreover, it can be useful to link the actions of users to their perceived actions. This would create knowledge about what users think they are doing versus what users are actually doing.

Next to the remediation that is invoked by human interaction, there is still a mystery around the natural remediation of infected IoT devices which should be understood better. Therefore, all actions in users' homes should be mapped to shed light on the reason for remediation or the lack of it. Next to end-users' ability to perform the needed steps, it is also unknown how they react to the notification mechanism in general. Verstegen (2019) has researched end-users' perception to different types of notifications, but knowledge lacks about their actual thoughts.

1.3 Research questions

In order to structure the problems addressed in section 1.2, it is necessary to scope the system at hand and follow a stepwise recipe to solve the problem. This research will focus on Mirai malware mainly, which still is one of the main threats for the abuse of the IoT. For Mirai, this research will look into private customers of KPN, which are easy to reach and comparable in contrast to business customers. The scope is in line with that of previous research in this field (Altena, 2018; Verstegen, 2019. The objectives stated in section 1.2 will be achieved through the following research question:

(RQ) What do we learn about how and to what extent Internet Service Providers can improve the remediation process of malware infected Internet of Things devices by monitoring end-users while they are cleaning their Mirai infections?

The main research question above will be supported by several sub-questions, which each hold a piece of the puzzle that is needed to answer the main research question. The following sections will elaborate on the defined sub-questions.

1.3.1 The crime scenes of Mirai

(SQ1) What does the crime scene of Mirai infected Internet of Things devices look like?

Before any remediation can take place, it is important to know what the situation is around the owners of infected Internet of Things devices. This includes the devices that are actually present at the homes of end-users and which of these devices are more prone to get infected. It is also about the households and characteristics of the owners. The findings in this research about the infected devices can be compared to prior research on commonly infected devices.

Moreover, the number of present devices can help in considering varying risks for end-users of getting infected. This question will be answered by asking end-users about their devices, their characteristics and by looking into open ports of end-users for information about the infected device (section 4.1.2). This data will show the variety of devices and can indicate the underlying awareness of users for example for security.

1.3.2 Reaction to the notification

(SQ2) How do end-users react to an email notification about their infected Internet of Things device(s)?

To achieve action by end-users, they should have the opportunity to act, but they should also be motivated to do so. This second sub-question looks into the reaction that end-users have when they receive an email notification about their infected Internet of Things device(s). If end-users for some reason do not make it to the remediation steps in the first place, there is no chance of succeeding in cleaning the infected device(s). It is therefore useful to look into how end-users react to the notification in the first place.

1.3.3 Following the right steps

(SQ3) To what extent are end-users able to perform the required actions to remediate their Mirai infected device(s)?

This sub-question is about the success rate of the clean-up efforts of end-users. As has been mentioned before, there is uncertainty about the actions that end-users perform after receiving a notification about their infected Internet of Things device. Verstegen (2019) has interviewed users about their actions, but was not able to verify the stated answers by the users. Therefore, there is still a remaining question about the actual actions that end-users perform. The answer to this question will shape the base for the advice that can help ISPs to send more useful notifications, as it can pinpoint critical points of error in the protocol where users seem to fail.

As has been researched by Verstegen (2019), there is a difference between the stated actions that users took in reaction to a notification and what actions they had actually taken. The reason for this misalignment is not yet clear. During the calls with end-users, they will be asked to perform the actions stated in the protocol, or recreate their actions if they already acted upon a notification. This will create two sets of data: one set of end-users who did not receive the protocol yet and will react to the notification during the call and another set of end-users that already received the protocol and should have already performed several steps before they are called. Figure 1 shows the possible timelines of events that indicate the different types of stages that end-users can enter. The variables that we consider are highlighted and coloured blue.

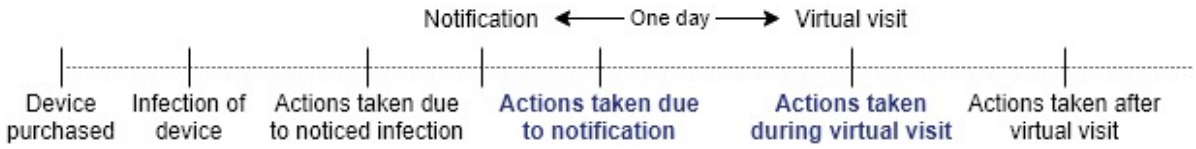


Figure 1. Timeline from purchase to supposed remediation.

Linking the actions that the users take to the actions stated in the protocol, can show critical points where users misunderstand the protocol or lack in skill to perform the required actions.

Next to the role of end-users in remediation of infected IoT devices, sometimes the devices return to a clean state naturally. As it is unknown why and how this natural remediation takes place, it is unclear whether users have a role in this process, or whether there are other causes for this phenomenon. It can be valuable to get a better understanding of how and why natural remediation takes place, as this might influence the way in which users are incentivized to act. The actions that users have taken will be analyzed to recognize unintentional cleaning. This will not cover all of the cases of natural remediation, but it will exclude some cases.

1.3.4 Barriers for successful process

(SQ4) What are possible improvements for the remediation process of Mirai infected Internet of Things devices?

This last sub-question partly summarizes findings in earlier sub-questions, by looking at all potential reasons why successful remediation does not take place. This covers the whole process of an infection taking place until the end-user has performed the required actions after being notified. The other part of the answer to this sub-question is to link these reasons to possible improvements and/or solutions.

1.4 Research approach

A variety of methods is used to answer the research questions. In a broad sense, the research is about capturing, processing and analysing both qualitative observed data as well as quantitative data on what devices are connected, on personal characteristics of the end-users, and on the remediation process during the virtual visits. As we are dealing with a complex socio-technical system, this method can interpret both the human side as well as the technical side of the system. According to Creswell (2009) and Denscombe (2008), this so-called mixed-methods approach can add value to research compared to either a completely qualitative or quantitative methodology.

Moreover, the research executes the capturing of both types of data during a single virtual visit. During a call with the user, both types of data can be captured. The analysis will then take place on both of the types of data combined, which indicates a convergent type of research (Creswell, 2009). Convergent research creates a result that is based on the combined findings of the methods used in the research.

Figure 2. shows the phases that can be distinguished from each other in the research. The figure also shows the serialism of the research, as one step cannot take place without the previous step. Over time, the research is divided into four separate phases, which will be elaborated upon in the following sections. Note that the research is mainly based around the live virtual visit to end-users.

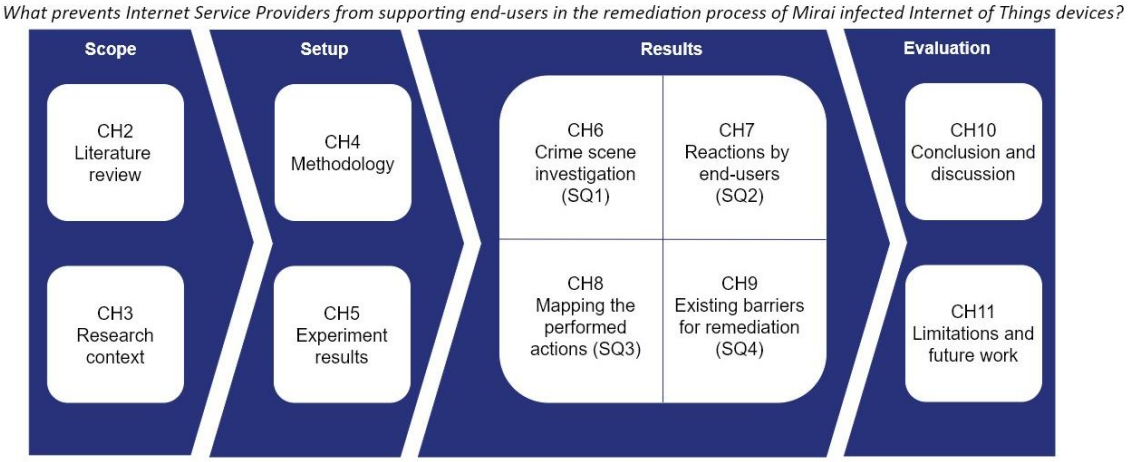


Figure 2. Research flow diagram.

1.4.1 Before the calls

The first phase is all about preparing the calls with the consumers. A literature review will identify the knowledge gaps that can be filled with this research. It will also provide a strong base on which this research can be build. Also, the current way of notifying the end-users will be displayed, as the advice that this research will create has an impact on how this notification mechanism should work.

Next to that, the preparation involves the protocol that will be used during the calls. As part of the data will be gathered during observation, it is important to make sure that the calls are comparable and thus that they take place in a similar fashion. Setting up a protocol will help with this consistency. As it is of great importance that end-users feel comfortable with sharing the steps they are performing while on the phone, the protocol includes ways to create this comfortable feeling for users. To make sure the protocol is usable, several test-runs take place to discover the pitfalls and potential issues. Thirdly, the target end-users will be identified, notified and asked for their consent. Only then can calls take place. The period before the calls can thus be found within the **scope** and **setup** of the research.

1.4.2 During the calls

After preparations have been made, the actual calls take place. Here the protocol is put to the test, and data is captured and recorded. This generates descriptive data about the end-users that are called and also on their network and capabilities of performing the required actions. Next to the data, useful information for both ISPs as well as future researchers in this area is generated, for example the pitfalls of the protocol or the do's and don'ts for approaching the customers. With this data, the last part of the **setup**, which is about the tracking results of the experiments and the **results** part of the research can be created.

1.4.3 After the calls

After the observation calls are finished, the gathered data is analyzed, merged and concluded upon. The answers to the sub-questions and thereby the answer to the main question will be displayed and limitations, recommendations and suggestions for future work can be stated. This happens in the **evaluation** part of the research.

1.5 Relevance of research

1.5.1 Scientific relevance

Although steps in identifying users' reactions have been taken, such as interviewing users after sending them different types of notifications (Altena, 2018; Verstegen, 2019), and finding that email notifications could have no effect, in contrast to walled gardens (Altena, 2018; Çetin et al., 2018), the remediation rate of infected IoT devices can still be improved from its current 60 - 76 percent (Çetin et al., 2019a; Durumeric et al., 2014). Next to the manual remediation of infected devices, a large part of the observed remediation supposedly happened without any human intervention (Verstegen, 2019). The exact cause of this so-called natural remediation, however, remains a question.

As has been stated before, knowledge is still lacking in the area of notifications in IoT abuse (Verstegen, 2019; Altena, 2018). Although several steps have been made to understand the end-users and to improve the notification mechanisms, there is still a demand for improvement of both the knowledge and the effectivity of notifications and processes at end-users' homes. This research fulfils this demand by investigating end-user home settings and occurring processes. Hypotheses exist on mental models of end-users in different security issues and their potential issues with the notification and remediation protocol, but these have not been put to the test in the area of IoT malware.

In a bigger picture, knowledge about end-users (re)actions is a piece of the puzzle of creating a more secure IoT environment. Research exists on vulnerable devices (Antonakakis et al., 2017; Çetin et al., 2019b), their vendors and also on different ways to notify end-users (Çetin et al., 2016; Durumeric et al., 2014; Li et al., 2016a; Vasek & Moore, 2012; Stock et al., 2018). If knowledge about end-users actions is added, the picture of the prevention and remediation of IoT devices will become more complete.

1.5.2 Societal relevance

Because of the complex multi-actor setting that IoT abuse finds itself in and the unclarity of who bears the responsibility for the issue, further insights in the processes of remediating devices from being Mirai infected can be useful for all involved actors. ISPs will benefit from the knowledge to create an improved notification mechanism. They will have a better understanding of what happens at their users' homes, which means their workload with regards to Mirai and likewise malware drops. The finding of Altena (2018) and Verstegen (2019) that walled garden techniques are too time-consuming strengthens the importance of improving less time-consuming methods for IPSs, such as a more effective notification mechanism.

Governments will have an increased oversight of the problem, which means their methods of informing and creating laws can become more effective. This can be useful for their current and future methods to create awareness. The Dutch government has not been idle since the rise of IoT and its corresponding issues, as their budget for this matter has increased and an awareness campaign with the name "update your devices" has been active since October 2019 (Ministerie van Economische Zaken, 2019). Another example is the effort of the Dutch version of the European Data Protection Board to create an overview of the suggested steps to maintain a safe internet environment at home (Autoriteit Persoonsgegevens, 2019).

End-users will be informed in a more proper way, which means the consequences of their infected devices will hurt less. Also, there will be more awareness of the problem of IoT malware in general, which supposedly increases security enhancing activities. Overall, the IoT environment should become safer if this research were to be used as input for policies.

In a broader view, internet security can be seen in the sustainable development goals (SDGs) defined by the United Nations (2020a). Although these goals are mainly focused on global problems such as hunger, climate and poverty, internet safety and connectivity are on the list too (United Nations, 2020b). Goal 9, which is about resilient infrastructures, mentions the importance of access to the internet for as many people as possible and goal 17 elaborates on the importance of this connectivity for partnership purposes. Moreover, goal 1 mentions the decrease of economic loss due to disasters taking place. Again, this is mainly focused on natural disasters, but problems such as DDoS attacks may be considered as well.

1.6 Research organization

Figure 2 gives an overview of the way in which this research report is organized. In chapter 2, a literature review will put this research into context with regards to the scientific domain. Moreover, chapter 3 will display the current setting at KPN to put the research into perspective considering the societal domain. Chapter 4 will elaborate upon the mixed methods methodology that is used in this research, including the altered notification mechanism that is used during the approaching of potential customers and the protocol that is used during the virtual visits. In chapter 5, an overview of the results is presented together with the observations that were made during the virtual visits. This chapter gives an answer to the first sub-question. Chapter 6 to 9 give answers to sub-questions 2 to 5, respectively. In chapter 10, main findings are recapped and combined to create conclusions. Based on these conclusions, several recommendations will be given. Chapter 11 ends the report by listing the limitations of the research and pointing out possibilities for future research.

2. Literature review

In this chapter, the existing literature on the problem at hand is summarized and discussed. The literature can be divided into four groups based on their subject:

The rise of the Internet of Things
The first part of the literature review is about the Internet of Things in general and about the security issues that arose with it. It also covers the reasons for the poor security that is currently inhabited in a large part of the IoT and why the incentives are not in the right place to improve this level of security.

The future of computer crime: Mirai malware
Mirai is one of the large threats to the Internet of things. The second part of the literature review will cover how this malware works and why it is such a big threat.

Existing remediation tactics and evaluation
Although there are many possible ways to improve the security of the IoT, these mitigation techniques have not been completely successful. The third part of the literature review covers different ways to improve the security level that have been tried in the past. It also covers reasons why these mitigation techniques have not been able to solve the problem fully.

End-users behaviour in IoT security
As the infected devices are purchased by end-users, these actors could play a large role in the remediation process. Therefore, their behaviour regarding security in general and their reactions to mitigation techniques that are performed by others is researched in the last part of the literature review.

Section 2.1 explains the methodology that is used for performing the literature review and summarizes the results of the literature search. Sections 2.2 to 2.5 discuss the literature that was found during the literature search and that was deemed useful for this research.

2.1 Methodology and literature search

The literature review is executed according to the best practices noted in van Wee & Banister (2016). According to van Wee & Banister (2016), the paper of Scheepers et al. (2014) should be seen as an example on the subject of literature reviews. Therefore, this literature research will focus on the paper of Scheepers et al. (2014). This means that the literature review includes the databases and languages that were considered, the time boundaries for research, the keywords that were used and the performed search strategy. For the search strategy, the systematic approach of Levy & Ellis (2006) is used. Their paper shows a stepwise approach to find the most valuable existing literature which exists of finding literature through using keywords in several databases, backwards snowballing through the cited articles in the identified articles and forward snowballing through the articles that have cited the identified articles.

The exact search strategy is depicted in Appendix A where the used databases, keywords, time boundaries and languages are described. In total, 43 reports were considered relevant. The title, main findings, introduced knowledge gaps and relevance for this research can also be found in appendix A. These 43 reports are supported by several news articles and books in the following sections.

2.2 The rise of the Internet of Things

At the end of 2020, twenty and a half billion “things” will supposedly be connected to the internet (Gartner, 2017). These widely adopted connected, context-aware and autonomous devices (Silverio-Fernández et al. 2018) are seen as an important part of the fourth industrial revolution. Since 1999, an exponential growth in the number of devices has taken place, and is still ongoing. One can no longer imagine today’s society without connected cameras, doorbells and thermostats; the Internet of Things is no longer at our doorstep, but everywhere inside our homes. Moreover, smart things can be found in more critical locations as well, such as smart cars.

Although life often becomes easier and more efficient with the aid of these smart devices, there are risks in sharing info through the multitude of devices. What increases these risks, is the poor security that a large part of the Internet of Things brings along (Alcaide, 2013; Abu Waraga et al., 2020; Lee et al., 2015; Sicari et al., 2015).

The poor security in many Internet of Things devices can be linked to several causes. First of all, the incentives to create well secured devices cannot be found within the supply chain, which means none of the physically involved parties benefits from investing in the security. Manufacturers and retail shops want cheap, fast produced products on the shelf and customers demand usable, working devices for a low price. This lack of incentive also holds for creating updates that could eliminate certain threats. Although users might encounter negative effects when a weakness in their IoT device is utilized by attackers, they often do not experience any negative impact from their infected device (Bauer & van Eeten, 2009). The lack of incentive causes customers to silently pass on taking security measures and performing updates.

Secondly, most IoT devices are focused on performing a single task, at a low price. This low price comes at the cost of the security investments by manufacturers. Also, as the devices have low computational power and can often perform but a single task, they are not compatible with the security measures that are implemented in regular computers such as laptops or smartphones (Batalla et al, 2017). Moreover, the operating systems often do not support user interfaces, which makes it hard for customers to change the standard passwords. Even when a supportive interface is present, users prefer having a working device as soon as possible over going through the installation steps (such as replacing the standard password) thoroughly (Kolias, 2017). The heterogeneity of devices and the way they work tops the effort it takes to take the basic security measures (Anthi et al., 2018).

2.3 The future of computer crime: Mirai malware

The IoT is large and poorly secured, which is an optimal hunting territory for attackers. Because of the small taskset that most devices can cover and the low market price, consumers acquire multiple devices to cover their demands. The size of this pool of easy prey makes it almost logical that attackers utilize it. As IoT devices inherently need to communicate constantly, there are many windows of opportunity for attackers to connect to the devices (Coulter & Pan, 2018). This creates possibilities for many types of malware to find their way into the devices (Coulter & Pan, 2018). Getting rid of the existing malware can be difficult, as it requires several steps to ensure that the malware has completely vanished. Figure 3 shows the necessary steps to disable the network, often starting with the identification of the malware; if the malware is not a known threat, it will be hard to mitigate it. Attackers naturally try to stay under the radar to prevent this first step from occurring. Therefore, it is unknown how many types of malware scavenge the internet at this moment in time. A successful type of malware that has been discovered in 2016 is Mirai (Kolias et al., 2017).

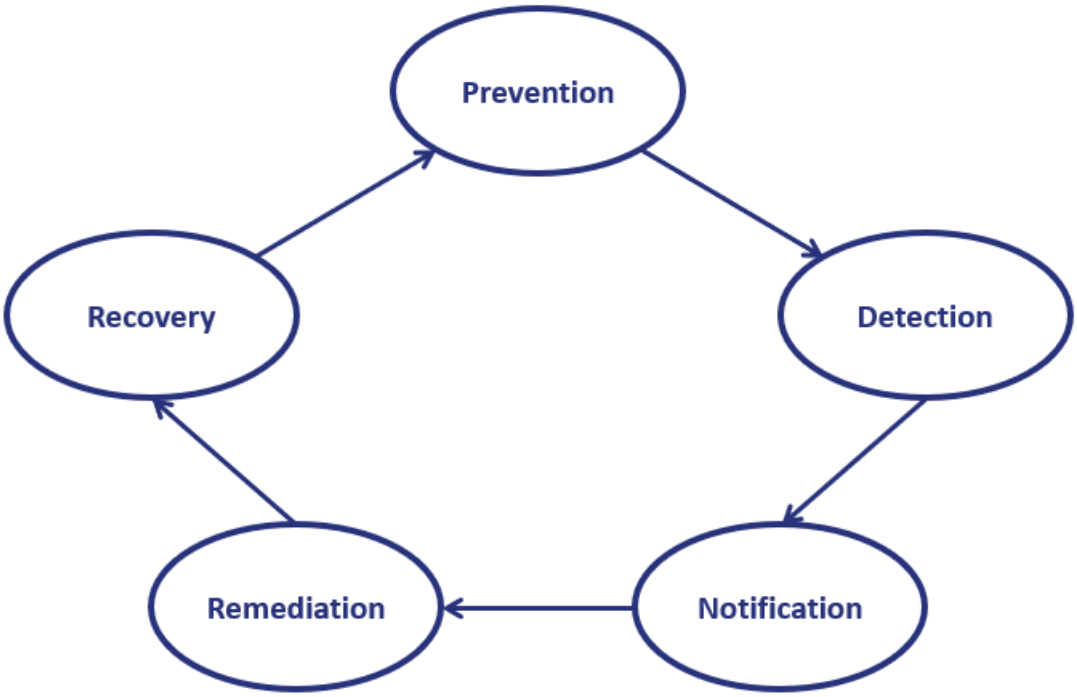


Figure 3. The anti-botnet cycle (Online Trust Alliance, 2013).

Mirai is one of the biggest predators in the field. It is a botnet malware that has been effective due to the large set of poorly secured devices, as it can spread easily and fast (Heartfield et al., 2018; Signes-Pont et al., 2018). Studies have measured that in 24 hours, Mirai is able to infect 400.000 new devices Safaei Pour et al., 2020). It is seen as the main threat in the IoT network (Dange & Chatterjee, 2019), as it creates opportunities for strong DDoS attacks, as illustrated in chapter 1 (Cardoso de Santanna, 2017; Cheung, Ryan, 2017; Chromik et al., 2015; Holl, 2015). Because the minds behind Mirai shared the code of the malware on GitHub, several months after it was discovered (Biggs, 2016), researchers have been able to investigate the working and the weaknesses of Mirai in the years thereafter.

Mirai exists of four components (Kolias, 2017): the **bot**, which is the malware that infects devices; the **command & control server**, which gives the attacker the possibility to manage their bots; the **loader,** which supports the spread of the executables and the **report server**, which holds info records on all infected devices. The first step of Mirai is scanning IPv4 addresses for vulnerable IoT devices. In this process, Transfer Control Protocol (TCP) ports 23 and 2323 are used for scanning, excluding known governmental IP addresses to stay under the radar.

When the bot infects a new, vulnerable device (mostly routers, cameras and digital video recorders; Antonakakis et al., 2017; Çetin et al., 2019b), the bot has four stages to go through (Antonakakis et al., 2017; Sinanović & Mrdovic, 2017; Kambourakis et al., 2017; Margolis et al., 2017; De Donne et al., 2018). The first stage is to perform a brute-force on the device, by using ten random entries of a list of standard known username/password combinations. If this brute-force succeeds, the newly infected device sends its IP and username/password combination to the attacker. In the third stage, the report server informs the loader, which is responsible for loading the correct binaries on the device. After the binaries have been executed successfully, they are deleted, and the device is now part of the botnet.

All bots in the botnet perform one of four possible actions:

1. Scanning random IPv4 addresses for new targets, which means, botnets will grow at a larger rate proportional to the number of bots in the botnet.
2. Killing any other traffic through ports 23 and 2323 to prevent other malware from taking over and to make sure that they are available as much as possible.
3. Waiting for further orders from the command & control server, which means the attacker can change the state of the bots, for example to turn into attack mode.
4. Attacking mode, where the attacker orders the bots to send requests to a given target, with the goal of denying the availability of the service of the target (DDoS).

Many studies have attempted creative, effective and novel ways to detect malware, including Mirai, which are deemed to be successful (Cid-Fuentes et al., 2018, Gill et al., 2020, Safaei Pour et al., 2020; Kumar & Lim, 2019). Mitigation of malware should begin in the prevention of the spreading (Dange & Chatterjee, 2019), which is also the case in the IoT, as the effects of malware that is able to flourish are proportional to the large size of the IoT. Still, identification of the existing malware is only the first step of mitigation.

2.4 Existing remediation tactics and evaluation

As can be seen in figure 3, the mitigation of botnets happens in 5 phases. Each of these phases provides opportunities for different actors to aid in the process of creating and maintaining security of the Internet of Things. As is described in section 2.2, the prevention of botnets from every existing is hard due to the mismatch between incentives and abilities. Although appointed by some literature as the responsible actors, (Kolias, 2018; Abu Waraga et al., 2020; Kambourakis et al., 2017; Mohsin et al., 2017) manufacturers and vendors are not likely to take significant action to improve the security.

Most vendors of IoT devices do not deliver a manual or support page with their product and even if they do, information about security is often absent or not adequate (Blythe et al., 2019; Furnell, 2007; Gibson et al., 2017). This means that even though consumers do care about security (Blythe et al., 2020; Nguyen et al., 2017; Rowe & Wood, 2013), the transaction costs of purchasing the most secure device are simply too high (Blythe & Johnson, 2018; Allen, 1999).

Moreover, there is no existing government party that can oversee the field of the IoT. As products travel across the world and the internet connects end-users globally, it is hard for government all over to organise themselves and put clear responsibilities and liabilities in place. As the IoT is still relatively new and evolving, it could take some time before governments are able to clean the market from rotten apples and punish the appointed responsible parties if security still lacks. Simple improvements such as labelling the level of security of devices could improve the purchasing environment (Johnson et al., 2020), but even for such small improvements, incentives are lacking.

Therefore, the current mitigation techniques mostly come from Internet Service Providers (ISPs) or informing campaigns to create awareness. These ISPs are seen as the key player in all of the five stages of mitigation (Pijpker & Vranken, 2016). Still, prevention is hard for these parties, as they have to rely on awareness of the end-users. What ISPs can aid in for prevention, is the reinfection rate of previously infected end-users. After such an end-user has gone through the other steps of cleaning, they will be back at the prevention stage. ISPs have an important role in making sure that the end-user has changed their behaviour appropriately during the other stages of mitigation. This will suppress the reinfection rate.

The second stage of mitigation is relatively easier. Because of the obvious patterns that Mirai leaves behind when in practice, detection of the malware on infected networks is no longer a main issue. Owners of effective honeypots, which capture malicious traffic and pinpoint infected networks as they pretend to be a vulnerable target, are currently sharing compromised IP addresses with internet service providers, who are then able to connect the IP addresses to actual consumers. This is where the notification stage can take place.

As users often do not experience any negative impact from their infected Internet of Things (IoT) devices, there's no incentive to clean these devices (Bauer & van Eeten, 2009). Because of this, notifications are needed to activate the end-users in cleaning their infected devices. In general, activating end-users by sending them notifications has a significant impact on the clean-up rate (Durumeric et al., 2014; Li et al., 2016b; Çetin et al., 2019a). However, this remediation rate is far from perfect.

Internet service providers have different techniques in place to notify the users and to support them to act. Because the effectiveness of notifications is key in mitigation, it is important to look into as many aspects of these notifications as possible. Figure 4 displays a conceptual model of how notifications are sent and what dimensions are important to consider.

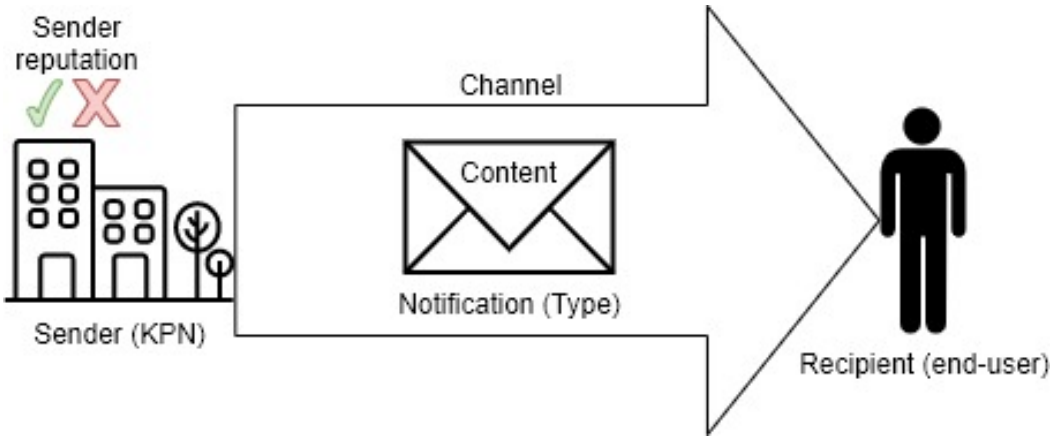


Figure 4. Notification dimensions.

2.4.1 Sender (reputation)

The first dimension of a notification is the sender of the message. As it differs how people react to advice depending on the actor or person that gives the advice, this could also be the case with notifications about infected IoT devices. Redmiles et al. (2016) found that users mostly focus on the trustworthiness of the sender when receiving digital-security advice. In contrast, according to Çetin et al. (2016), the reputation of this sender has no significant impact on how well people clean up their infected devices. Although trustworthiness is not completely equal to reputation, these different findings are hard to grasp. Stock et al. (2018) show that there should be a minimum level of trust by the receiver in the sender to make the notifications effective. Though the research of Stock et al., (2018) is about domain owners and not about regular customers of ISPs, the contradicting findings are notable.

Li et al. (2016a) describes that a personal message can increase the effectiveness of the notification. This could be as simple as ending the message with personal credentials of who should be contacted by end-users when they need help. Independent from the organisation that sends the notification, a personal context should help.

2.4.2 Content

The content of a notification should be understandable and clear for the target group in order to reach the desired goal. Research has found that sharing the steps that should be taken in detail increases the effect of the notification (Çetin et al., 2016; Durumeric et al., 2014; Li et al., 2016a; Vasek & Moore, 2012). On the other hand, Forget et al. (2016) mention that the message should be plain and simple.

There is a dilemma in preparing the content of notifications between explaining the steps to be taken in full detail and keeping the message understandable and usable. If the message contains too much information, users might experience the message as an intrusion into their privacy (Redmiles et al., 2016). If it lacks information though, users are unable to take the correct actions.

Reder et al. (2012) mention the importance of understandable messages when approaching end-users. Story telling about the risks of their infected device can improve the awareness and actionability of the users. Moreover, the study showed that when users can relate to the incidents that take place because of their failing security or the failing security of others, their willingness to act increases.

2.4.3 Channel

The last main dimension of notifications that can be altered is the channel through which the message is forwarded from the sender to the receiver. Two channels have been investigated elaborately in research.

As ISPs often have customers email addresses at their disposal, sending the notification through this email is a logical way of working. Literature is contradicting on the effectiveness of these email-only notifications. According to Stock et al. (2018) and Li et al. (2016a) , sending email-only notifications without additional measures can help in pushing end-users to clean their devices, while Çetin et al. (2019a) and Altena (2018) suggest that the remediation rate does not increase significantly compared to cases where no action is taken at all.

Stock et al. (2018) states that reasons for non-compliance could be found in the rarity with which important notifications are forwarded through email. Users tend to ignore the email message or interpret is as spam. This process strengthens if multiple notifications were to be sent. Although one might think that sending more than one notification might help in activating the users, this only enlarges the chance of them interpreting the messages as spam or phishing (Vasek & Moore, 2012).

The other mitigation technique used by ISPs is to put infected users into a so called “walled garden“. This way of working gives users an actual incentive to act, as their internet access is removed until they can prove to have taken the required actions for remediation. Next to creating awareness of the security issues, walled gardens have the ability to impact the end-users. Literature shows that this internet quarantine technique has the highest impact on the remediation rate by end-users (Çetin et al., 2018, 2019a, 2019b).

However, as only 50 to 75 percent of quarantined users manages to clean their infected device by themselves, the usage of walled gardens invokes a large number of calls to ISPs helpdesks and complaints about the ISPs services. Also, users have a negative experience during the quarantine as it means the ISPs deliver the opposite of their regular product: a lack of internet service (Çetin et al., 2019a). Although the mitigation technique is effective, is has large side-effects, which ISPs cannot manage. Therefore, this way of supporting remediation is not used widely.

2.5 End-users thinking process for IoT security

Even when the best practices for notifying users are put into place by ISPs, the remediation rate does not reach 100%. Although receivers of the notifications have positive feelings about the messages they received (Li et al., 2016b), they often fail to perform the needed steps appropriately.

End-users care more about the ends of their IoT devices than the means that are used by the devices to reach these ends. Ironically, users state to purchase IoT devices to increase their feeling of security at home (Zimmermann et al., 2019). However, for example by bringing in smart security cameras, the security of another kind drops. The perceived risk caused by this drop in security has no impact on purchases (Klobas et al., 2019).

Moreover, users do not want the responsibility of managing the existing risks, as their usage of IoT decreases if they have control over the security settings. This can be explained by the lack of knowledge of end-users about security settings, which is caused by the absence of information sharing by the manufacturers and vendors (Blythe et al., 2019; Furnell, 2007; Gibson et al., 2017). The few end-users that do care about the security also lack in knowledge to perform the right actions due to the heterogeneity in IoT devices that exists (Zimmermann et al., 2019; Forget et al., 2019).

Even if a device starts to malfunction, users often believe that this is caused by the inherent flaws of their relatively cheap IoT product. Rarely, users will link the weird behaviour to malware or hackers (Huijts et al., 2019). Again, this points at the importance of notifications as a wake-up call for end-users.

As has been stated before, in the case where an end-user is notified about their infected IoT device, the message is often regarded as spam or ignored. If the notification is taken seriously, there are still some barriers that could keep users from performing the needed steps.

Noteworthy is a typical protocol that users receive when their IP pops up in the Mirai infected list. Users are asked to take the following 5 steps:

Step 1. Determine which devices are connected to your Internet connection. The Mirai virus mainly infects Internet connected devices such as a digital video recorder (DVR), security camera or printer connected to the Internet (not computers, laptops, tablets or mobile phones).

This first step is present to remind end-users about which devices are more prone to become infected, to point them in the right direction.

Step 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.

As Mirai makes use of brute force attacks with common username/password combinations, changing the password to something less standard will prevent the device from being reinfected shortly after is has been cleaned.

Step 3. Restart the Internet connected devices by turning them off and on again. Hereafter, the Mirai virus has been removed from the memory of the devices.

Because Mirai only exists in the memory of the infected device(s), resetting the device will remove Mirai as this memory is cleaned (Cao et al., 2017).

Step 4. Reset your modem/router to the factory settings. On https://www.kpn.com/service/internet/wifi-en-modems/herstart-reset-experia-box.htm it is described how you can do this for an Experia Box.

Step 5. Change the password of your modem/router. On https://www.kpn.com/service/internet/wifi-en-modems/wijzigen/servicetool-wifi-naam-wachtwoord-wijzigen.htm it is described how you can do this for an Experia Box.

Because there is a significant chance that end-users are not able to perform the necessary steps with regards to the infected device itself, additional steps should be taken regarding the modem/router. The reset of the router to factory settings is a means that should lead to the end that the ports are closed, and there is no space for DMZ (Demilitarized Zone) or UPnP (Universal Plug and Play), which means the devices within the network cannot be approached from the wider internet.

Even in the case where users are aware of the problem and activated to act, there is often uncertainty about what device is infected or about how to take the required steps (Redmiles, 2019; Redmiles et al., 2016). Zeng et al. (2017) add that users rely on older techniques to solve the problems on their current devices, which often is not the correct way for the new types of devices and infections. The article also mentions the interaction between multiple owners of the IoT device as a barrier for correct remediation.

Although Verstegen (2019) had the possibility of asking users about their reasoning, it turned out their stated behaviour did not always match their observed behaviour. Notified users stated to have taken certain actions, which had not been taken. This could be explained by the social desirability bias (Fisher, 1993) that makes people move their behaviour to what is socially acceptable. The discrepancy could also be explained by the lack of knowledge of end users about the effects of their own actions. More thorough data is lacking in this area of research, which could point out the real reason behind users decision to act or not to act properly. Knowledge about the phase of remediation where end-users go wrong is not yet available.

3. Research context

This chapter shows the context in which this research is executed. Section 3.1 identifies the important actors involved, which is put into perspective in section 3.3. Section 3.2 is an introduction to the Abusedesk of KPN, which is the main source of the data used in this research. The processes analysed through the Abusedesk are also displayed in section 3.2. The chapter closes with section 3.4, where the influence of the current Covid-19 crisis is elaborated upon.

3.1 Actor arena

As has been stated in section 1.1.5, the remediation of infected IoT devices should be seen in the context of its complex, multi-actor system. Sheng et al. (2009) mapped the system for phishing crimes, of which the structure has been used for the arena of IoT malware in table 1. Table 1 gives an overview of the involved actors, their relevance and their role within the system. Chapter 2 explained why it is uncertain which of the actors is responsible for the abuse, and why there is no perfect solution to the problem yet.

Actor	Relevance	Role
Infected end-users (households, companies)	These actors can be seen as the army, as they own the infected bots that are used by attackers. They often do not feel the consequences of their bot, which is why they have no real incentive to act.	Victim
DDoS targets	These are the main victims of the bots, as their services become unavailable temporarily. They can only try to prevent the attacks from being successful, or try to minimize the damage done in case of a successful attack.	
Internet Service Providers	These actors provide the end-users with their connection. They have the possibility to give incentives to their users to act.	Infrastructure provider
IoT manufacturers	These actors deliver poorly secured devices to the market, as this saves money and time, which is crucial in today's fast economy.	
IoT vendors	These actors decide to sell the products that manufacturers deliver. They could create standards for their products, but do not have a real incentive to do so.	
Researchers	Researchers can inform other defenders about the right steps to take, but also create awareness amongst end-users. Their role is less direct.	Defender
Non-profit security vendor (Shadowserver, AbuseHub)	These actors are able to detect malicious traffic and pinpoint the IP addresses behind this traffic. They can alert parties who have more power to act.	
Governments	These actors are in favour of general security and can start awareness campaigns or even change the law to force markets into delivering more secure products. These actors are less directly involved.	
Botnet owners	These actors are the suppliers of DDoS attacks as they have control over the infected devices. Although their incentive is often money related, they provide opportunities for DDoS attacks.	Attacker
DDoS customers	These actors are on the demand side of DDoS attacks as they have some incentive to keep a certain service from being available.	

Table 1. Involved actors in the system.

3.2 KPN Abusedesk

As one of the main Internet Service Providers in the Netherlands, KPN has an exemplary role to provide secure internet connections, but also to aid in the ensuring of clean internet traffic. The KPN Chief Information Security Office (CISO) owns this important task as their daily business. Within the office, the Security Operation Centre (SOC) takes care of securing customers systems and making sure its protected against potential DDoS attacks and compliant with the law. Part of the SOC is the KPN Abusedesk, which focuses on the remediation of vulnerabilities and the abuse of network traffic that makes use of KPN services. The Abusedesk is responsible for the mitigation mechanism that is used to help customers remediate their infected devices.

The Abusedesk performs activities in all of the steps of the mitigation cycle in figure 3. For the detection of infected devices, the Abusedesk relies on external parties to provide lists of infected IP addresses. The non-profit organisation Shadowserver tracks a large amount of data, about which they report to governments, ISPs and others who could benefit from it (Shadowserver, 2019). Next to the data that the Abuse team receives from Shadowserver, the Abuse Information Exchange organisation, which represents most Dutch ISPs, collects data through their AbuseHub system (Abuse Information Exchange, 2020). The Abusedesk uses both data sources to get an overview of their infected customers, as they can combine the infected IP addresses from the beforementioned sources with their own data on their customers.

After the data feeds have been processed, the infected users are placed into a walled garden by the Abusedesk. From there, the customer can only perform the most important tasks on the internet (such as sending e-mails and performing financial transactions), including the landing page where they are informed about their infected device(s). Accompanied, users receive an email notification which explains the reason why the users have been placed in quarantine and a protocol that the users should follow in order to get out of the walled garden. Appendix B shows the exact messages that users receive when placed into a walled garden. After performing the needed steps, the users are asked to notify the Abusedesk about their actions, however, mostly users who need help in performing the needed steps actually send contact the Abusedesk.

After this process, customers should have recovered from their infection and should be more aware of the dangers that IoT devices bring along. If customers are infected more than two times, they need assistance from a member of the Abuse team to be released from the walled garden, which is an extra incentive for users not to become infected once again. This can be seen as a prevention measure by the Abusedesk.

3.3 Researched system

Because of the complexity and nested nature of the system at hand, as has been explained in section 1.1.5, it is important to map the system and understand what is in- and excluded from the research. Figure 5 shows an overview of how the actors identified in section 3.1 relate to each other and where the processes explained in section 3.2 take place in this actor arena. Full arrows represent physical actions or streams, while dotted arrows represent informational or nonphysical streams. The red arrows show the unwanted effects of Mirai infected bots, while the green arrow shows part of the solution to the problem. Figure 5 shows what is covered in this research, related to the greater system at hand. What is excluded from the research has become blurry. Next to the focus, also the sub-questions are added to figure 5, which makes clear which parts of the system are investigated in what sub-question. As can be seen, question 1 is about the present devices and the household characteristics, question 2 is about the reaction of end-users to the notification, question 3 is about the specific steps that end-users take and sub-question 4 is about the barriers of the whole system.

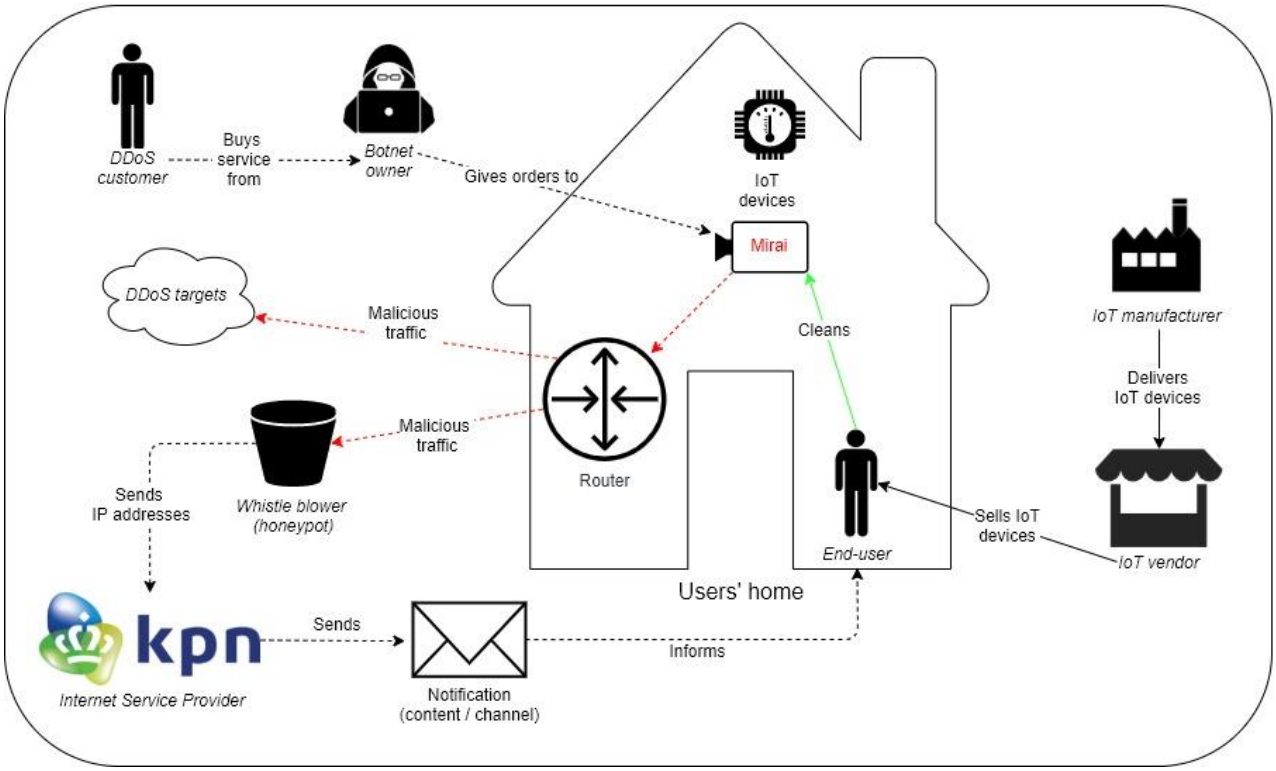


Figure 5. Conceptual model of researched system.

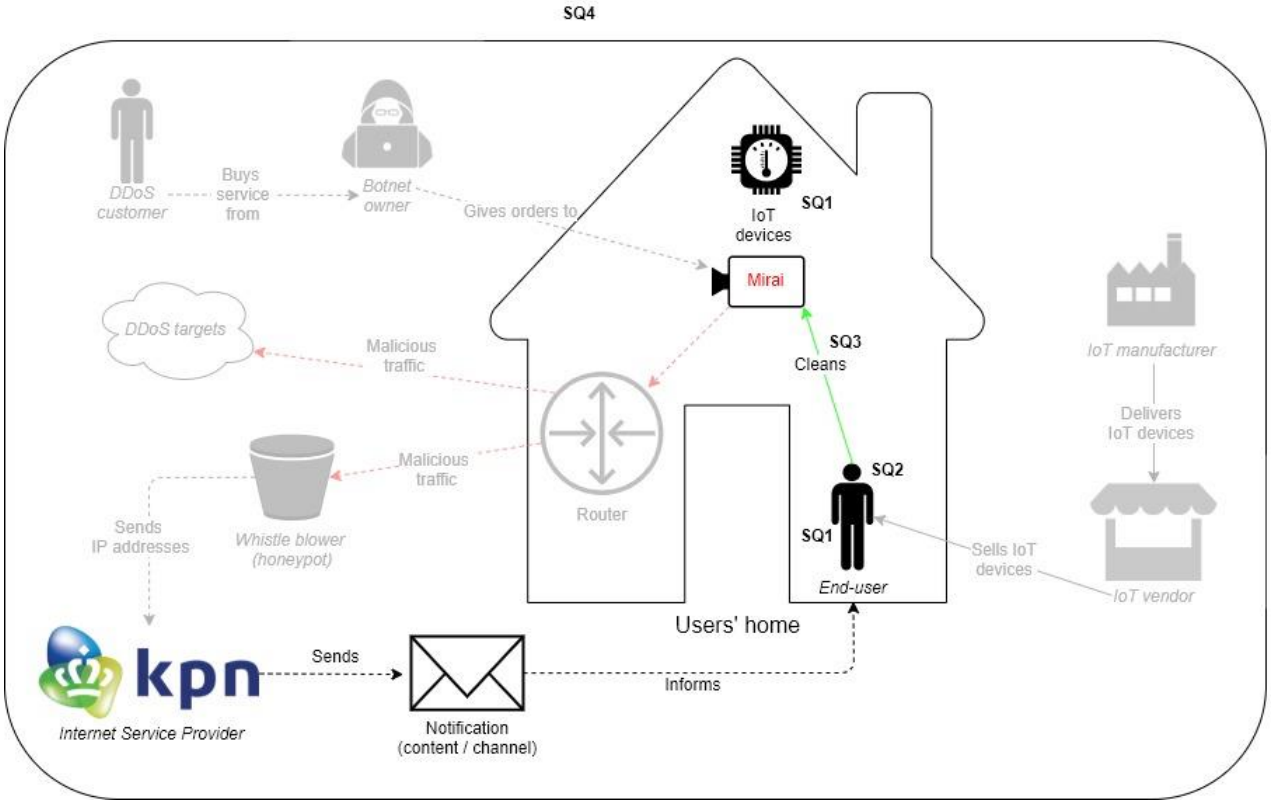


Figure 6. Conceptual model of researched variables.

3.4 Covid-19 circumstances

One might wonder why the visits in this research are organised over the phone, instead of performing actual physical visits to the end-users' home. As the knowledge gaps in the literature state that there is an incomprehensible difference between the stated behaviour of end-users and the perceived behaviour, paying real life visits to end-users might discover the reason for this gap and would help to clarify what processes take place at users' homes. The idea would then be to create a safe environment for the end-users in which can be observed how they approach the notification they received, and which steps they perform as a reaction to this notification. However, due to the countermeasures against the Covid-19 virus, physical visits are out of the picture for this research. The request from the Dutch government to stay inside as much as possible has created a situation that requires desk research for the largest part of the project. Therefore, the decision has been made to move away from physical visits to virtual ones. As many people are stranded in their homes, they will have a better opportunity to participate in the research. Instead of observing the end-users' behaviour, the research now relies on the statements of end-users, while they are performing the steps that they think are correct to remediate their infected devices.

A part of the identified knowledge gaps is about the difference between end-users' actual and stated behaviour. Because this research still relies on end-users' statements, this discrepancy can still take place. However, the difference between this research and previous research is that the end-users perform the steps while they are on the phone. To make sure the end-users feel that they are in a safe environment, the virtual visit protocol is built around the idea that the end-users have done nothing wrong and can state their feelings and actions without any judgement.

The countermeasures against Covid-19 also have a small positive effect on this research, as the availability of many end-users has increased. Now that many users work from home, the chance of them having an opportunity to be part of the experiments has increased significantly. Next to the increased number of opportunities, some people are not able to work from their homes, which increases the likeliness of participation even more.

In general, when analysing the results of this research, one should keep in mind the extreme circumstances of the context of this research, which probably have put several biases on the gathered data and thus the findings.

4. Methodology

This chapter explains how the main research question will be answered. The research is based around a virtual visit, where we observe customers while they perform the requested steps to resolve their Mirai infected device(s). This includes an e-mail notification which will inform the user, a virtual visit where data is gathered about the user and their actions, and the processing, analysing and concluding upon this gathered data. During the call, two intertwined methods are used, as there is no single methodology that can ensure valuable data on all aspects of this research. A part of the call exists of a small survey, in which demographics and information about users' home network environment is gathered. The other part of the call exists of a think aloud protocol, where we virtually observe the users in their actions and thoughts.

Section 4.1 sketches the available data sources and how they relate to each other. In section 4.2 is explained how the available data is used and displays a general overview of the complete experiment. Section 4.3 describes how the virtual calls, which are a main part of the experiment, are prepared. The execution of the virtual visit is discussed in section 4.4, including the think aloud protocol. Section 4.5 discusses how the results of the observations are processed and analysed. The chapter ends with several limitations of the methodology in section 4.6 and an overview of the ethical considerations in section 4.7.

4.1 Available data sources

As this research is positioned mainly around the end-users of infected Internet of Things devices, these end-users themselves are seen as the main data source for the research. Literature states that detailed and fitting information is required to be able to remediate their devices properly. Although ISPs, such as the Dutch KPN, try to implement positive findings of research, remediation results are still far from perfect. The thinking process of end-users and the actions they take are valuable data for ISPs to get insight into the critical points of their notifications and the awareness and understanding skills of their customers. Figure 7 gives an overview of the large cycle through which end-users are processed. The following sections elaborate on the parts in the figure.

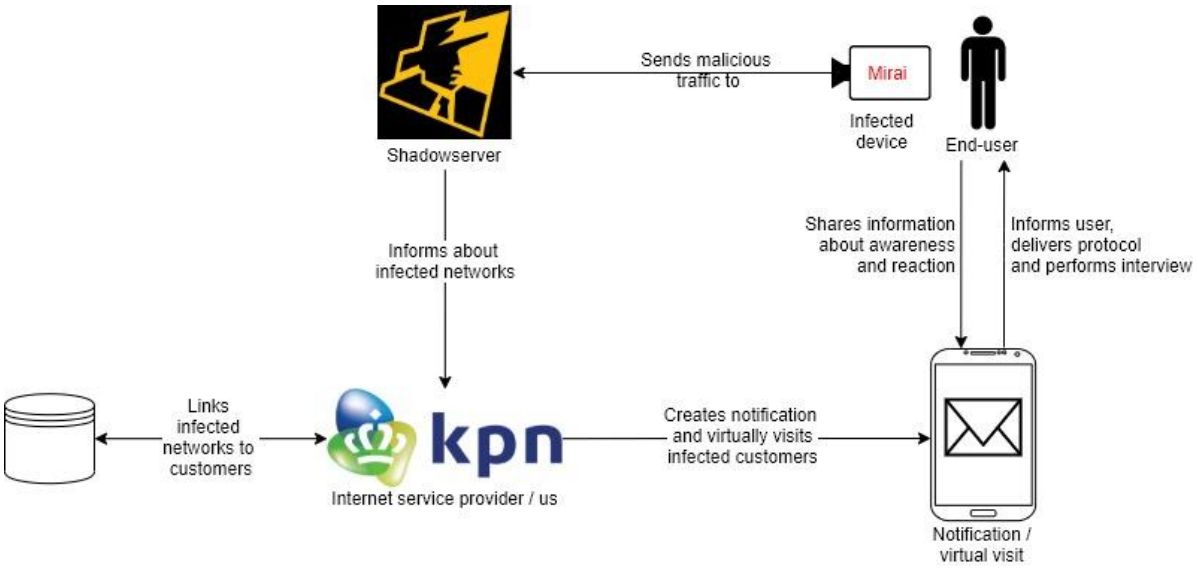


Figure 7. Cycle investigated in research experiments.

4.1.1 Monitoring infected home networks

Shadowserver (2020) and other infection feeds are able to discover Mirai infected home-networks across the internet due to the detectable signature of Mirai malware (Antonakakis et al., 2017; Sinanović & Mrdovic, 2017; Kambourakis et al., 2017; Margolis et al., 2017; De Donne et al., 2018). ISPs can then link these infected networks to their customers, which makes it possible to contact these customers about the infections. This gives us two sets of data: the IP addresses of infected networks over time and some customer information and contact details on customers who own at least one infected device. The monitored infected networks can show whether an infection is remediated after a certain period of time, as it will stop showing up on the logs of Shadowserver and other infection feeds if the infection is gone. This is how previous research has measured clean-up rates and infection time (Çetin et al., 2016; Vasek & Moore, 2012; Altena, 2018; Verstegen, 2019).

4.1.2 Mapping the connected devices

When the infected customer is located, KPN is sometimes able to map the connected devices in the network of this customer. With the consent of the customer, we can look into this map of devices and identify which IoT devices are connected, together with some information about the devices themselves, such as a display name, a model name and sometimes the manufacturer. Because previous research has already pointed out certain devices that commonly become infected (Antonakakis et al., 2017; Çetin et al., 2019b), this data can help in finding the infected device.

Next to the mapping tool, we also used a tool that takes screenshots of Mirai infected networks. By actively scanning IP addresses that match the Mirai fingerprint (Antonakakis et al., 2017) in real time using masscan, zgrab and gowitness, we uncover which ports were open and we collect screenshots (gowitness) and banners (zgrab) of the home pages of these ports. Using the login pages and the banners collected from the open ports, we can label the device that is behind the open port and thus potentially could be the cause of infection. That the ports are open is not actual proof that the labelled device is infected. However, the fact that the device is visible to the open internet shows that that specific device is vulnerable, as attackers use similar approaches to uncover devices they can infect.

Although both data sources could help in identifying the infected device in the network of the end-user, there is no guarantee for this. The tools might be inaccurate or incomplete, which means they should not be seen as ground truths. However, if information is available, it can and should be used to verify users perception of their infected devices. Figure 8 displays the process of collecting the data on devices and comparing this data to the perception of end-users.

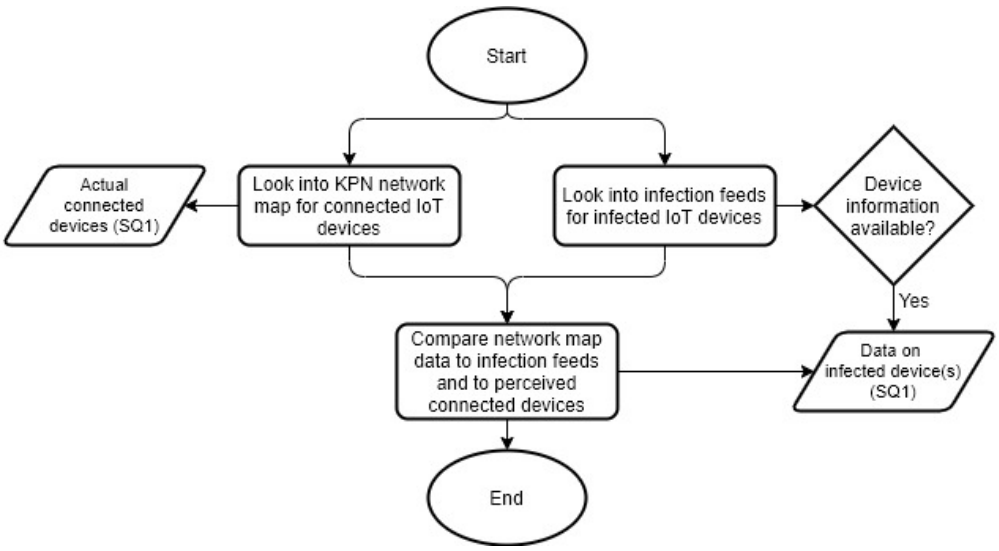


Figure 8. Data sources for connected devices.

4.1.3 Contacting the end-users

KPN can link the infected home networks to customers, which means they possess the contact credentials of the infected network owners. Therefore, we are able to contact these end-users directly, and personally. During these virtual visits, users are asked about their IoT devices, which can be compared to the mapped connected devices. As the first step in the remediation protocol is to find the infected device, the comparison between these two data points can show to what extent users are aware of their home environment and if they are able to find the infected device. Comparing users' perception to the known vulnerable devices can strengthen the extent to which users' are unable to perform the first step of remediation successfully. Also, users are asked about some personal characteristics, such as their age, gender, and the size of their household. Here, it is important that we collect the data on the end-user in the household that is the most likely to perform the remediation steps. To make sure that the right person is on the phone, the protocol includes a check for this.

The other part of the virtual visit is about the steps that end-users take as a reaction to the sent notification. This data has a format that can be linked to the steps in the protocol. One the one hand there is a number of actions that the users perform, such as "pressing an on/off button for five seconds" and on the other hand there is the actions that users took linked to the actions that the users thought they were performing, such as "reset the infected device". These datasets can show what steps are hardest to perform by end-users and how the protocol steps are perceived by the end-users. Next to that, end-users could perform steps that are not listed in the protocol, but that are successful for cleaning the device. This is where a part of the natural remediation could be explained.

Lastly, the willingness to participate in the experiment is a data source of its own. How the notification email is interpreted by end-users and the extent to which they receive and read this notification are two variables that can help explain issues in the notification protocol.

4.2 Experiment setup

To get the most out of the data sources described in section 4.1, a clear setup for the virtual visits is needed. This section elaborates upon the process of discovering a Mirai infected customer to ending the virtual visit with that particular customer to end.

4.2.1 Experiment population

In any study, it is important to map the possible data-subjects in the population that is analysed. The target population exists of *KPN consumers that own at least one Internet of Things device that is infected with Mirai*. Moreover, the experiment focuses *on the household member that is responsible for the clean-up attempt*. As the largest pool of infections can be found in regular consumers' homes, only households are considered within the experiment. Business customers of KPN are excluded as it can be hard to pinpoint the person responsible for clean-up and the setup of the network is significantly different and often bigger than a regular home setup. There can be boundary cases where a consumer runs a small business from their own network. These consumers are included in the research as they are comparable to regular consumers.

Both Verstegen (2019) and Altena (2018) distinguished KPN customers from Telfort customers in their research, due to the different mechanisms that are used within KPN to notify the two markets and the potential differences in personal characteristics of the customers. However, in this research the two populations are seen as one, as the notification mechanism is generalised, and the virtual visit is about end-user behaviour rather than their characteristics. Moreover, as Telfort is currently In a process of changing into KPN, the differences are insignificant. When referring to KPN customers in this research, this includes Telfort customers.

4.2.2 Experiment protocol

During the experimentation period, on all working days, the incoming infection feeds of the day before are analysed for Mirai infected home networks (Figure 9). This leaves out networks that are only detected as infected on Friday and/or Saturday, on which will be elaborated in section 4.5. If an entry has already been notified before the start of the experimentation period, the corresponding consumer is discarded from the experiment. If an entry has been notified within the experimentation period, this means no virtual visit has taken place yet, or the remediation during the virtual visit was unsuccessful. A last reason for reappearance in the feed is the reinfection of the network or the presence of more than one infected device, which was not identified during or before the virtual visit.

All participating customers, which are found by processing the infected IP addresses are notified through an email about the situation and our research (appendix E & F). Directly together with the notification message, an opt-out is sent to the infected customers, which they can choose at any point. If the user does not reply to this notification email, it is assumed that they would like to participate in the research. Still, they can step out at any time, of which they are informed again during the call. Section 4.3 elaborates upon the notification email. In appendix E & F the complete email can be found.

The virtual calls take place one working day after the notification is sent, which means notifications are sent on Sunday to Thursday and virtual visits take place on Monday to Friday. The choice to plan virtual visits only one day after the notification was sent has been made because the study is more valuable if users perform the remediation steps only during the virtual visit. If the virtual visit was planned at a later stage, users may have already acted. Section 4.5 elaborates on the limitations of this choice.

If a phone call is not answered, a voice mail message is left behind stating that we will attempt to reach the end-user later that day. After three attempts to call the end-user, they are discarded from the experiment and will be dealt with using the standard protocol of KPN. That is, the customers are put into quarantine until they have proven that they took the required actions to resolve the problem.

As a day is limited by 24 hours, it could be that the number of infections exceeds the maximum number of virtual visits that can take place on a single day. If this happens, a random set of the total number of infections is chosen to take part in the experiments, whereas the remaining infections are discarded from the experiment and dealt with as usual (section 3.2). The participants are chosen based on the algorithm in appendix C.

The first part of the call is about informed consent. This consent involves both taking part in this research anonymously, as well as the virtual visit taking place and the recording of this virtual visit. Lastly, the consumers' consent is asked to look into the setup of their home network. Users can step out at any time. If they do not wish to participate, they will be processed as usual by KPN (section 3.2), of which they are informed before the call ends.

Because most of the data is gathered during the virtual visits, it is important to fine-tune the used protocol as much as possible. Section 4.3 and 4.4 elaborate on the protocol that is used during the virtual visits and the adjustments that were made. After a virtual visit has taken place, the end-users' network should be remediated from Mirai. To make sure the virtual visit was successful, the infection feeds are scanned for the virtually visited users' IP addresses until the end of the experimentation period. This can also identify possible reinfections of remediated end-users.

Figure 9 shows the outline of the experimentation period and the possible states that can occur within the experiment. As can be seen, consumers that were notified before the experiment period started are not part of the experiment. If a consumer shows up on the feed of infections, but has already been

notified, this indicates that either the virtual visit takes place on that particular day, or that the virtual visit has already taken place, but was unsuccessful in terms of remediation. The whole process regarding one consumer takes up to two days. On the first day the consumer is notified and on the second day the user is called. To make sure that there is equal time between all cases in the experiment, notifications are also sent on Sunday, and not on Friday. The virtual visit process (coloured blue) is elaborated upon in section 4.3.

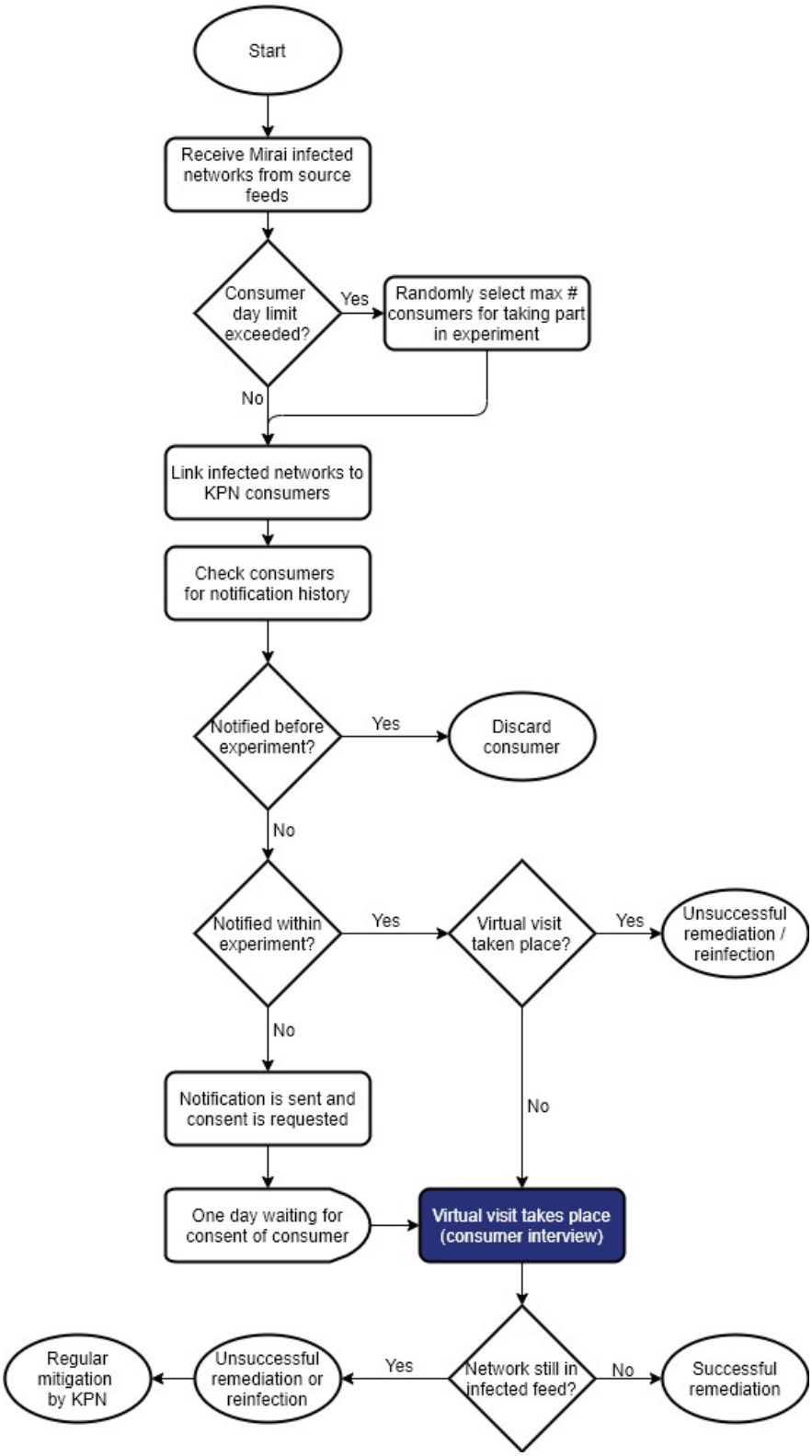


Figure 9. Flowchart of the experimentation processes

4.3 Email notification and virtual visit preparation

As end-users often do not notice the Mirai infection on their device(s), it is important to inform them about the existence of this infection. This is done through an email notification, which can be found in appendix E & F. In this email, it is also explained to end-users why they are not quarantined, which is the usual way to process Mirai infections at KPN (section 3.2). The reason and usefulness of this research is discussed shortly and the processes that are needed for this research are elaborated upon.

Although the email is not part of the data that is collected in this research, it can have a significant impact on the willingness of end-users to participate in the research and their mindset during the experiments. Next to this, the image of KPN should not be harmed and the email should be trustworthy, as there is no real way for the end-users to exclude that the email is not dispatched from KPN. This means it should be seen as crucial for the experiments.

To make sure that the email is read and understood by as many end-users as possible, several communication experts have transcribed the text into B1 level of the CEFR, the Common European Framework of Reference of Languages (European Union & Council of Europe, 2004). Moreover, to ensure as many end-users can participate in the research, the email is written both in English as well as in Dutch.

The email includes the steps of the protocol that end-users are asked to perform only during the call, which means they have access to the needed information to perform the steps on their own before the call happens. As participants should be able to choose not to participate, it is ethically necessary to share the steps with them in another way than during the experiment. However, it could be that part of the participants already acted before the call took place.

During the preparation of the virtual visits, we gained experience in dealing cases where we will not be able to aid consumers with the problems they have. Firstly, we went along with senior KPN mechanic K. Plugge for a day to see and ask what his reaction is in cases of unsuccessful problem solving or unsatisfied consumers. The most important step in such an occurrence is to inform the consumer of the situation and to let them know about the possible ways forward. Even for a mechanic who has been practicing his job for over 30 years, sometimes new, hard to solve problems can occur, which is human. Next to informing the consumer about the possible futures, also a supervisor should be informed about the issue. What is also important, is that extra effort is not needed in cases where remediation is hard. Instead of infinitely trying to help consumers in finding the right steps, a few trials are enough, after which informing is the best we can do.

Secondly, we learned from senior desk member T. Datema, who is experienced in dealing with unexpected problems over the phone and how to react to those. Patience and understanding are key as the beginning of a phone call can set the tone for the remaining of the call. According to Datema, it is sometimes, useable data can only be gathered in a comfortable environment.

Lastly, the experts of the abuse team informed us about their reaction to failed remediation cases. On a daily basis, the abuse team has to deal with end-users that own unsavable products or end-users that lack knowledge to solve the problems at hand. For the end-users not to be constantly quarantined, they must either replace their unsafe device with a safer one, or setup a physical visit with one of the KPN mechanics/experts who can perform the needed steps properly during a physical visit.

4.4 Virtual visit protocol

As has been discussed before, within 24 hours after the notification email has been sent, the customer is contacted through the phone. The virtual visit is an observational study that entails a think aloud protocol (TAP). This type of study can be found under the umbrella term survey.

Observational study

Although the virtual aspect of these virtual visits may disguise it as an interview, the method should be seen as an observational study. In observational studies, there is no control over the independent variables. As this research is mainly about understanding processes at end-users' homes, no altering of the independent variables takes place. Observational studies are useful for providing information on "real world" uses of practices (notification mechanism) for the general population and to help formulate hypotheses for subsequent experiments (Nahin 2012).

Although observational studies cannot make definitive statements about the effectiveness of the practice (notification mechanism), it is seen as the most useful way forward. Research already exists on the effectiveness of notifications (Durumeric et al., 2014; Li et al., 2016b; Çetin et al., 2019a) and users' experiences with notifications (Li et al., 2016b; Redmiles, 2019; Çetin et al., 2019a), but literature lacks information about end-users' thinking and acting processes, which can be analysed best through observation, as this removes most influence of outsiders, leaving the natural situation of end-users. This is considered to be the most effective way forward.

Think aloud protocol

As the visits in this research are virtual, there is no actual direct observation possible. Even though there is a possibility to change setup a video call, which creates an opportunity to observe nonverbal behaviour, the observations are still lacking complete understanding of what is going on in the end-user's home. This is why we make use of a think aloud protocol (TAP).

The think aloud protocol has been applied for about 30 years, which is since Lewis (1982) discovered its usefulness. Research still agrees with the strengths and weaknesses of the method that Lewis noted (Nielsen, 1993; Hornbæk, 2010; Fan et al., 2020). TAP can give detailed insights in the thinking process of the observed subject. The social desirability bias (Fisher, 1993) is limited, as subjects are asked to verbalize their thoughts. This means there is little time for subjects to think over their words and the data is close to the real thoughts of users (van Someren et al., 1994; Rubin and Chisnell, 2008).

Moreover, TAP is a way for us to remain on the side-line, which creates valuable data on end-user processes without many of the biases of regular interviews. Staying on the side-line can be hard as users might expect us to jump in when they are facing difficulties, especially, because they are used to receiving aid actively when dealing with computer related issues (Poole et al., 2009). A common benefit of TAP is the pinpointing of specific issues that subjects encounter during the execution process. These are often issues that are hard to anticipate when designing and even testing a certain system or protocol.

Next to identifying a problem, TAP ensures that the cause of the problem is brought to light. Before this study, ISPs such as KPN only have infection feeds to find out whether clean-up efforts by end-users are effective. In a mail conversation, KPN might find out at which of the steps the end-user misunderstood, but not why exactly. Next to measuring the effectiveness of the notification, TAP can show the attitude of end-users to some extent, based on their comments and tone.

The limitations of the TAP method are also consistent over time (Lewis, 1982; John & Marks, 1997; van den Haak et al., 2003). Even though the researcher is trying to stay at the side-line, their presence influences the participants, who will typically pay more attention and effort to the tasks at hand. Moreover, as words and language are limiting, it is impossible to explain perfectly what mind processes take place. This is why the comments of participants are always incomplete.

In fact, there are four layers between the truth and what we can observe. The first layer is the understanding of the participant of their own thoughts. The second layer is about the translation of this understanding to words. The third layer is about us understanding the words and the fourth layer is about us translating the words back into thoughts. Due to these layers, information loss is inevitable.

A last relevant limitation of TAP is the complicatedness of the information that is transmitted. As it is sometimes unclear what actions end-users have actually taken, the think aloud protocol creates data that can be hard to process. There have to be clear classifications of actions, which can be used to process the stories that end-users tell. Still, it might be hard to recognize these needles in the haystack of participants' thoughts.

Within think aloud protocols, there is often a clear distinction between concurrent and retrospective protocols (Rubin and Chisnell, 2008; van den Haak et al., 2003). The latter means cutting the experiment in half, where participant only explain their thoughts after they have finished the requested tasks.

A benefit of this type of TAP is that it is typically found easier to do by participants, because they can focus on one thing at a time. However, a major drawback of the retrospective TAP is the additional time needed for the experiment which is about twice as long. Additionally, participants tend to rationalize their actions rather than only report their behaviour. This increases the number of layers between the real data and the reported data. This is why we have chosen to use a consecutive think aloud protocol instead.

A major difference between this study and regular TAP studies, is the system that is tested. Due to the heterogeneity of Internet of Things devices on the market (Zimmermann et al., 2019; Forget et al., 2019), different participants will encounter different issues. In regular TAP studies, this system is the same for all participants. This is why the number of subjects studied in this research should be higher than the regular number. This will increase the usability of the recorded data.

According to Nielsen (1994), of all the issues that will be found by participating subjects, the proportion of these total number of discovered issues for increasing numbers of subjects fits a Poisson model with equation $1-(1-\lambda)^n$ where λ varies between 0.12 to 0.48 and n is the number of participants (Nielsen & Landauer, 1993). This means that about 80% of detected problems is often discovered by only five participants and fifteen participants should be able to discover 95% of the detected issues. At some point, the probability of additional issues being found by additional subjects approaches zero, which can be seen as an asymptote. Again, due to the heterogeneity in this research, the number of participants is higher, namely 17. Section 5.2 elaborates upon the reasoning behind this number of participants.

Creating a safe and comfortable environment

For end-users to share their thoughts in as much detail as possible, they should feel safe and comfortable during the experiment. Therefore, it is important to look into the best practices of performing observational studies and surveys. Literature is rich on how to perform such experiments as effective as possible.

Almost a century ago, Bingham and Moore (1931) mentioned the importance of keeping the interest of the participant and that preparation is key. To make sure that the virtual visits happen smoothly and without surprises, a clear and thorough protocol is created. Also, the protocol includes some explanation of why the virtual visits are useful for the consumers to keep their interest. In the classification of survey types by Turner (2010), the virtual visits in this research can be seen as "Standardized Open-Ended surveys", as the observations should be identical in the virtual visits to different consumers, but the answers of the consumers to the questions can be largely different.

Although these types of surveys are excellent for comparing and aggregating the findings in multiple experiments, it can be hard to code the data that is gathered, as the answers can be different (Creswell, 2007). It should be clear up front as much as possible how the varying answers are classified and interpreted. To increase the likelihood of understanding what the participants are trying to communicate, external researchers should look into the answers of the participants and review the evaluations of these answers.

Moreover, standardized open-ended surveys are more likely to be successful if the purpose is explained (1), the terms of confidentiality are addressed (2), the format (3) and duration (4) of the experiment are mentioned and if it is made clear how the participants can make contact at a later stage (5). Lastly, the experiment should be recorded to save as much data as possible (6) and the participant should be asked clearly if they have any questions before the actual experiment can start (7) (McNamara, 2009). If something still seems unclear during the experiment, either for the researcher or the participant, asking follow-up questions can be helpful in reaching clarity. Although the abovementioned best practices increase the usability of the experiments, it cannot be known where the critical points of potential problems are without performing pilot tests (Kvale, 2007). Important in these pilot tests is that the setting is the same or at least similar to the setting in the real experiments. Section 4.4.4 elaborates upon the pilot tests that have been executed in this research.

Castillo-Montoya (2016) introduced a framework that can increase the quality of a survey protocol by shaping it through four suggested phases. In line with the findings above, this framework suggests the usage of pilots and adjusting the protocol based on feedback. What has not been discussed yet, is the alignment of the parts of the survey with the research questions. The alignment increases the usability of the experiments and ensures that the right data is obtained. To make sure that this alignment is present, the parts of the virtual visit are linked directly to one or more of the research questions. Seidman (2013) stresses that the right data in surveys is about the stories of participants that matter. The experiments should above all support the participants to put their experiences into words. It is important to create a safe environment for the end-users and to make sure that they are willing to and feeling satisfied with sharing the truth. This also holds for the TAP part of the call. For example, to create an environment that is as "safe" as possible, the consumers are asked during the call whether they wish to switch to a video call. Having a face next to the voice might help in the level of comfort that the consumers feel. It can also help to observe the nonverbal language of participants better. Lastly, it is important to remain on the side-line as much as possible as researcher. It is key to have a little influence as possible.

Keeping in mind the best practices in the abovementioned literature, a virtual visit protocol is created which is divided into three parts. Each of the three parts is discussed in a corresponding section. In Appendix B, also the complete set of questions can be found. This protocol is the result of many iterations of adjusting and testing through pilot phone calls. Section 4.4.4 elaborates upon the major changes that the protocol went through, through iterations and due to unwanted phenomena during the pilot test-calls. The protocol that was used in the pilots can be found in appendix C.

The following sections explain what processes happen in the different phases of the virtual visit protocol and how they relate to each other. Each section is supported by a figure of the corresponding part of the virtual visit. To make sure the research questions are covered well within the experiment, each of the figures shows how the processes and obtained data relates to this research (Seidman, 2013). Notably, sub-question 4 is missing from the graphs, as this sub-question will be answered by the process as a whole, instead of particular parts of the virtual visit. Although sub-question 2 can be found in figure 10, also this question is answered by the complete process and not only by the reaction to the email notification, labelled in figure 10.

4.4.1 Introduction, consent, and person check (Virtual visit part 1)

The first part of the virtual visit is about engaging the consumer (Bingham & Moore, 1931) and ensuring they give consent for the recording and usage of the call for this research. The participants are informed again that they are allowed to quit the experiment at any time. If for any reason a participant does not wish to be recorded during the call, their thoughts are written down in keywords for which a scheme is prepared that exists of the needed data points that should the collected during the call. For example, when a participant explains about their connected devices, a list of devices is written down live and when a participant thinks aloud during the resetting of their infected device, their actions are written down in keywords. This way of collecting data is less rich than when the call can be recorded, because there is no way of revisiting the thoughts of the participant by other researchers in a later stage to ensure the statements of the participant are interpreted correctly.

Moreover, the first part of the virtual visit is about checking whether the right person is on the phone, which should be the person in the household that is responsible for the remediation process. If the right person is on the phone, they are shortly informed about the purpose, the format, and the duration of the call (McNamara, 2009). Then end-users are asked whether they are interested in switching to a video call setting, which could increase their trust in the experiment and create a safe environment for them. The wearing of KPN clothing during the calls will increase the level of trust even more. As has been mentioned before, the video calls can also increase the usefulness of the calls, as it adds nonverbal language.

If the user wishes to change to a video call, they are invited to make use of a Skype video meeting. Skype is widely used within KPN and has been pen tested thoroughly. Although there is a promising platform within TU Delft for video calls as well, which has been created to deal with the constrictions of the countermeasures against Covid-19, trust is key in having successful virtual visits. This is why the more commonly known Skype is selected over the Big Blue Button platform from Delft (Fiebig, 2020). Because the virtual visit itself might already feel intrusive into people's private home environment, it is important to relieve the protocol of disturbance wherever we can.

Before the actual data gathering starts, there is some time for remaining questions the customer might have (Turner, 2010). The first part of the experiment ends with asking the consumer whether the notification has arrived properly. If this is not the case, another notification is sent real time and the consumer is asked to read it.

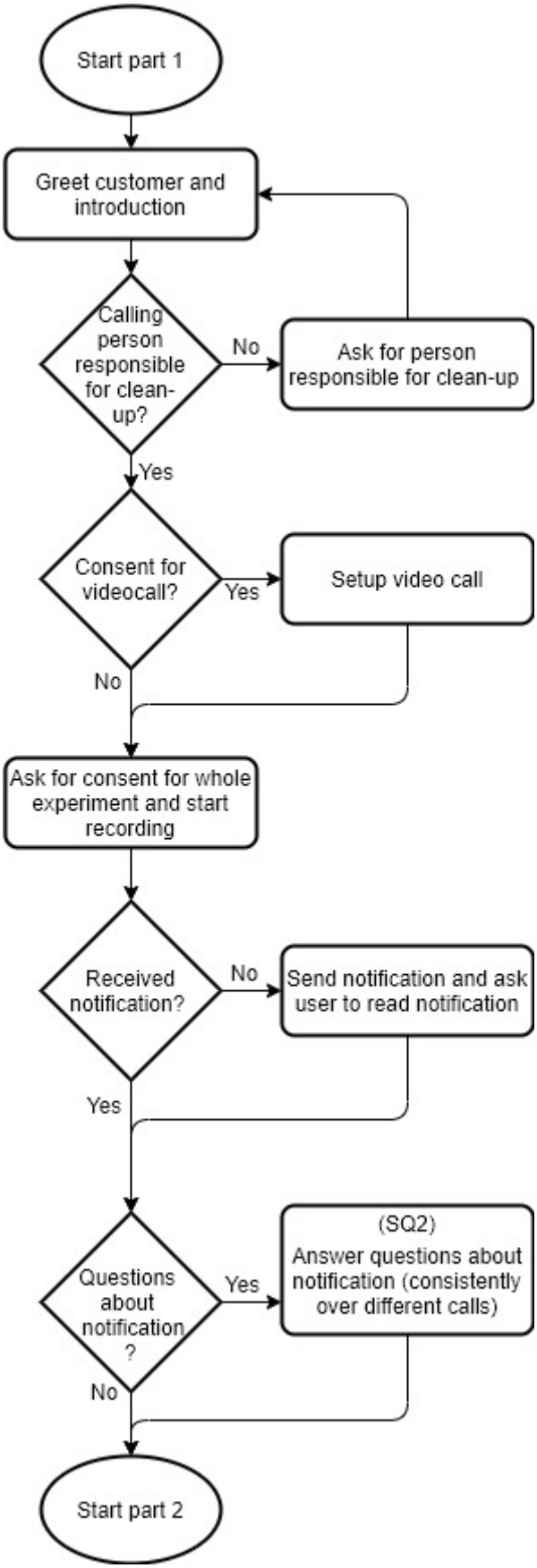


Figure 10. Flowchart of part 1 of the virtual visit protocol.

4.4.2 Think aloud protocol (Virtual visit part 2)

The second and largest part of the virtual visit is the real-life performance of the notification protocol that users have received. This part exists completely of the think aloud protocol. As the virtual visit takes place only one day after the users have been notified, in most cases the user will not have acted yet. During the call, the end-user is asked to look for the notification they received and to perform the steps in the order of the protocol in the notification. For each of the steps, users are first asked to perform the steps on their own, without any interference.

Even if users fail to perform a step correctly, they are asked to continue the remaining steps. Only at the end of the protocol, users are offered help to still perform the needed steps properly. This choice to help the end-users only at the end of the protocol ensures the consumers are not disturbed or influenced during their performance. Section 4.6 elaborates further on this choice. While the users perform the steps, they are asked to think aloud during the complete process. Here it is important to make sure the consumer feels comfortable sharing their actions, which is achieved by constantly reminding the user that they are helping the process which is needed due to its flaws and that there are no right and wrong answers. Their mistakes are not their fault.

The **first step** of the protocol is to identify the devices connected to the network environment. In this phase of the call, users are supposed to pinpoint one or multiple devices that are infected according to them out of the list of connected devices. They should think aloud about which device they identify as infected and also the reason why they believe that this choice is correct. If users do not mention any of the known vulnerable devices, such as printers, security cameras or DVRs, this is a strong indication that their remediation efforts will not be successful. In other words, the mental representation of all the devices is a crucial part of being able to identify potentially infected devices.

According to literature, this step is hard for consumers, due to their belief that their devices inherently malfunction and their lack of knowledge that some devices are connected to the internet in the first place (Huijts et al., 2019; Klobas et al., 2019). If users are not able to identify a potentially infected IoT device for any reason, they are offered help through the phone, but only at the end of the call. Using the heuristics of known vulnerable devices, users are sent into the right direction. Still, there is no guarantee that the correct device is identified during the call.

The **second step** of the protocol is to reset the device identified as infected. During this step, users think aloud, while trying to reset the device. This step typically includes actions such as pulling out the plug or trying to press available buttons for several seconds. It is important that the users explain their actions more in detail than stating that they are “resetting the device”. Therefore, we try to hint at the users to think aloud in more detail. For the research it is key that the actual actions of end-users are mapped, instead of their perceived actions. Again, if users fail to perform the steps, they are aided in doing so after all steps have been performed. Like with the first step, there is no guarantee that the reset actions are successful. Whether this is the case can be seen in the infection feeds of the following days.

The **third step** of the protocol is to change the password of the device identified as infected. According to literature, this third step is likely to be hard for users, as there is a large heterogeneity amongst IoT devices (Zimmermann et al., 2019; Forget et al., 2019) and they often miss a usable user interface (Kolias, 2017; Anthi et al., 2018). Due to the same reasons, this step will still be hard to execute properly with aid as well.

The **fourth step** of the protocol is to reset the router to which the infected device is connected. As these routers are KPN routers and there is plenty of information on how to adjust the settings of these routers, this step should be relatively straightforward compared to the previous steps. Still, users should still think aloud in detail.

The **fifth step** of the protocol and also the final step that should be executed, is to change the password of the router to which the infected device is connected. Again, as consumers make use of KPN routers, there should be a possibility to log in to the router and easily change this password. Not surprisingly, users are asked to think aloud while they perform the needed actions.

After the last step has been performed, users are explained to which of the steps they misunderstood or have to redo in order to reach the wanted effect, if any. The home network environment should be remediated successfully. This intervention takes only place at the end of the virtual visit because the influence of the caller should be minimized in this observational setting. As the steps in the protocol are not dependent on each other and can be executed separately, no intervention is needed in between.

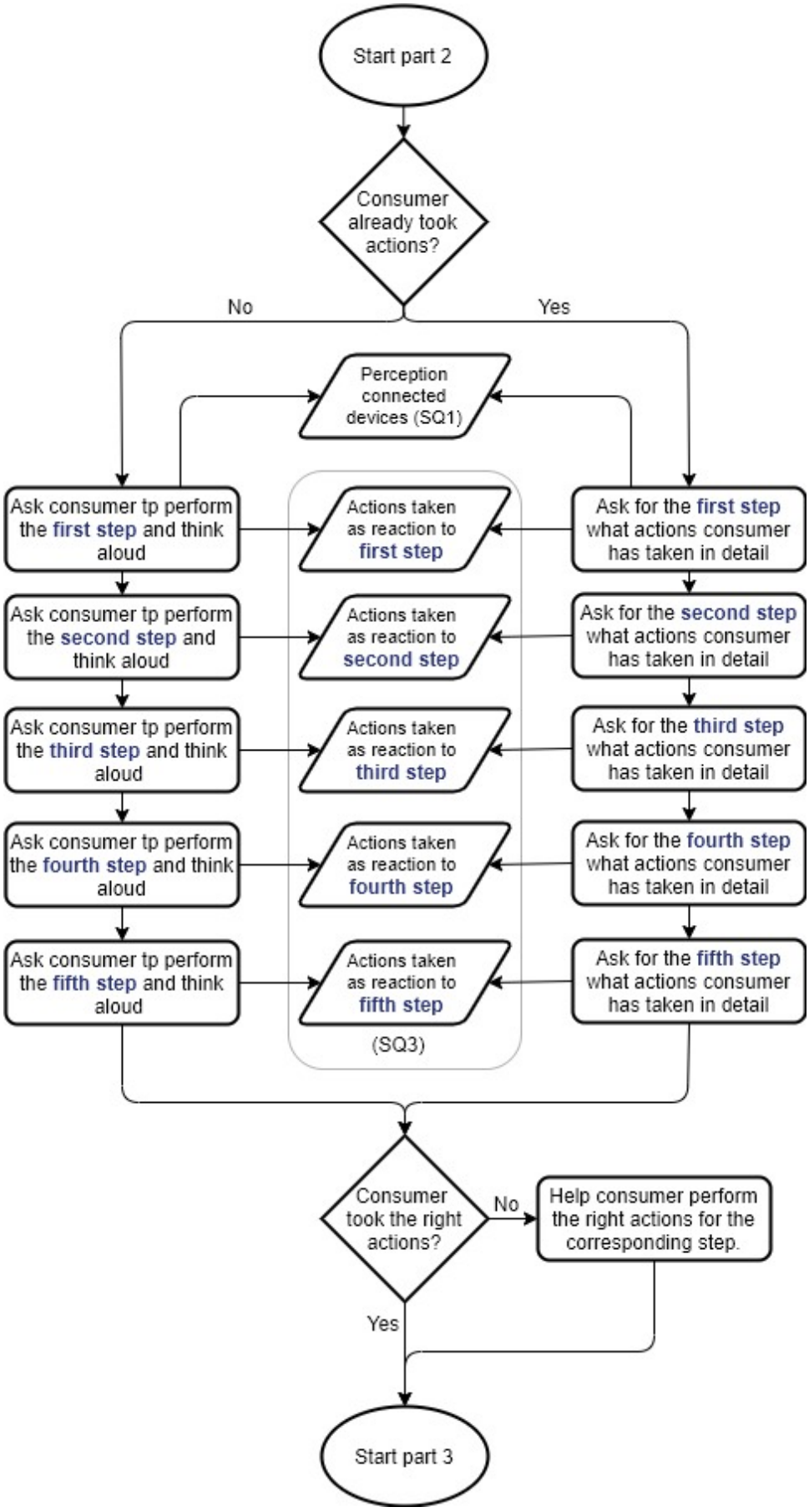


Figure 11. Flowchart of part 2 of the virtual visit protocol.

4.4.3 Demographics survey and ending (Virtual visit part 3)

The fourth and final part of the protocol is a small survey to capture some demographics of the end-users, which can be useful in explaining the captured data and finding differences over several end-user groups. As we aware are that the virtual visit is a relatively intrusive and exhausting experiment for end-users, the demographics part of the call is short and only at the end of the call. We capture the age and the household size of the end-user, as well as the type, brand, and vendor of the infected device. Age has been pointed out by Verstegen (2019) as an influencing factor for infections, which can be verified in this research. Verstegen (2019) also was not able to include the household size and the number of devices as indicators for who gets infected. The protocol ends with asking the end-users permission for looking into the KPN network mapping tool. If this tool shows any potentially infected devices which have not been cleaned yet, the users are informed of this the next day if their network still is listed as infected.

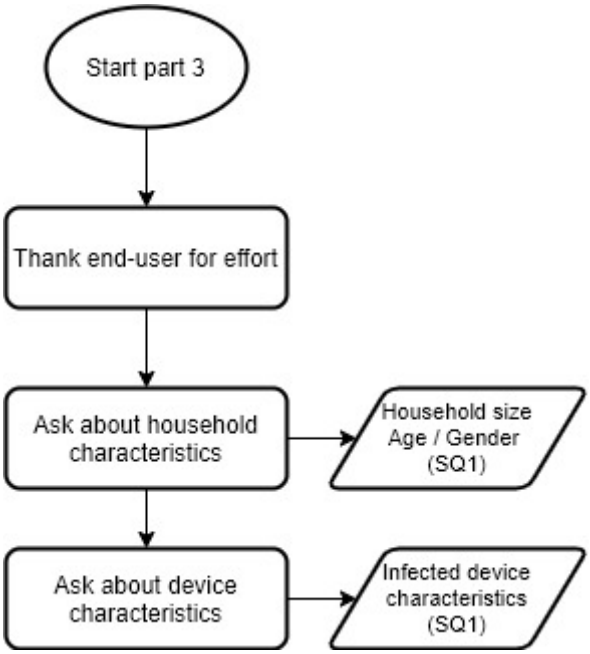


Figure 12. Flowchart of part 3 of the virtual visit protocol.

Data is now captured about end-users household settings, their perceived and real set of connected devices, the actions they take to accomplish what is asked in the steps of the protocol and the actions they take per step of the protocol, which indicates the difference between their perceived actions and their real actions. The virtual visit closes with sharing the details on how the end-users can keep in contact (McNamara, 2009). Users are also thanked for their time and effort and are informed about what would happen in the case where their network has not been cleaned from Mirai properly.

The day after the virtual visit, the household should no longer show up on the infection feeds. If they indeed do not show up again, the virtual visit was successful. However, if a supposedly cleaned home shows up on the infected feed after a virtual visit, this is an indication that the clean-up efforts were not successful. Reasons for unsuccessful clean-up could be the existence of another infected device in the household that was not identified as infected or the uncareful execution of the required steps by the end-user.

In most cases, unsuccessful remediation will be noticeable during the virtual visit, as both the end-user and the caller are not able to find a way to perform one of the needed steps properly. This could be caused by a lack of interface, which is common in IoT devices (Kolias, 2017; Anthi et al., 2018), or by a misunderstanding with an end-user.

4.4.4 Feedback and virtual visit pilots

As mentioned before, the best way to find out whether an experiment protocol works properly is to test it in a setting that is the same as, or at least similar to the real setting in which it will be used (Creswell, 2007; Kvale, 2007; Castillo-Montoya, 2016). This section elaborates on the pilot calls that have taken place and what has been changed to the protocol as a result of the pilots.

During the pilot call phase, 12 owners of at least one infected IoT device have been approached. Out of these 12 customers, we were able to reach ten through the phone of which a subset of seven customers took part in the complete process of the experiment. The results of the pilots are varying, and many potential pitfalls have been discovered.

The experiment protocol

First of all, the execution of the experiment protocol should be consistent and clear. Mistakes have been made with executing this protocol properly, such as sending the wrong notification email or putting the end-users in internet quarantine next to sending the email, which is not part of this experiment. Naturally, the end-users who were affected by these mistakes acted before the virtual call took place. To resolve these potential mistakes, the process of sending the notification email has been standardized and there is a check for whether the right email is loaded. Moreover, technical issues can arise during the experiments such as a failing microphone or recording tools. It is added to the protocol to test both of these tools before a call is made.

On the end-user side of the experiment, some unexpected phenomena happened as well. In two cases, the person responsible for security issues was not part of the household. An example is the family son who comes to take care of such issues. However, it can be hard to reach this family son. In cases where the responsible person is not part of the household, still the household size of the household that owns the infected device is considered in the analyses.

Another common issue is that customers make use of different email addresses in their KPN account than the commonly used addresses, which also holds for their telephone number. Out of ten customers that we could reach through the phone, two customers did not recognize the email address that the notification was sent to and four customers saw the email but did not read it. Naturally, the customers that did not recall receiving an email felt relatively cautious about the phone call. Both customers did not want to take part in the research and hung up shortly the start of the call. In one case, the answer to a telephone call was by a completely different person than the targeted customer, due to a wrongly inputted telephone number.

As there is no way to prove the authenticity of either the email or the call, five out of ten called customers mentioned their caution out loud and tried to receive evidence of the authenticity of the research. An extended explanation of this research and the use of the calls as well as mentioning the opt-out possibility took away the worries of four out of five customers. One customer contacted the service desk of KPN to verify the authenticity of the email and call, who took part in the research one day after this verification.

The virtual visit protocol

The pilot calls have shown that it can be hard to keep a structured conversation according to the virtual visit protocol. Participants tend to share all their thoughts and feelings as soon as possible. Although this was unexpected, it should not be seen as harmful, as long as the necessary conversation parts take place. To be sure that consent for being part of the experiment and recording the call are still present at the start of the call, the protocol has been changed slightly, to having a larger block of information sharing before any questions are asked (appendix B & C).

A second identified flaw of the draft protocol is part two (appendix C), where a separate part is dedicated to the present devices in the end-users' homes. As the information regarding the present devices is already being processed during step one of the remediation steps, this is redundant and unnecessary. Instead of dedicating a part of the virtual visit to the connected devices, the screenshot tool is utilised right before the call, so the user can be helped with identifying the infected device at the end of the call if needed. The KPN network mapping tool is only used after the call, to increase the focus on the think aloud protocol part of the virtual visit.

Another improvement is the timing of asking consent for insight into the network mapping tool of KPN. In the draft protocol this happened during the section on present devices, which felt unnatural, as the participant would expect us to help in finding the infected device if we are able to see all the connected devices. In the final protocol, users are asked consent for this tool at the end of the call. If a participant had a hard time figuring out what device is infected, it is natural to look into this tool after the call. If something peculiar turns up, the user can be informed about this if they are still infected the next day. However, as the tool turned out to be unstable and often lacking in detailed information, this change to the protocol is only useful in a handful of cases.

Lastly, the pilot calls took approximately 10 to 15 minutes, which is included in the informing start of the virtual visits, as it can help create a trusted environment for participants (McNamara, 2009).

Pilot experiment findings

Out of the seven customers who took part in the experiment, all seven were remediated on the same day as the virtual visit. Two out of five customers who did not participate were still infected after three days. Two participants already acted before the call took place, where one customer took all the necessary actions and the other customer only did not perform the last step of the protocol, which is to reset the password of the router.

First remediation step

Six participants identified a plausible device as infected, which was a security camera in five cases and a printer in one case. Two customers could not figure out the brand of the camera, who had all bought it online somewhere. The other three cameras where a Velleman CAMCOLD 26, Astak Wireless Security Camera (CM-811T) and an Avtech camera. The printer as a HP photosmart 55-20. The remaining participant was sure that their laptop was the cause of the problem. Afterwards it turned out the participant also had a security camera which probably was the issue, as the infection was removed after performing the steps on this camera.

Second remediation step

Three out of seven participants were unable to reset the password of their device, as there was no manual or support page they could find. Two participants simply disconnected their camera from their internet connection and kept the camera up for scaring purposes. One participant was able to change the password of the device through a browser and one participant was able to find a manual for resetting the password through a google search.

Third remediation step

Four out of seven participants were able to reset their device by pulling out the plug for approximately 5 to 10 seconds. Two customers already disconnected their device in the previous step, which also holds for this step. One customer was not able to reset their device.

Fourth remediation step

Seven out of seven participants were able to reset their router. Five participants made use of the link that was sent along with the notification email on how to reset the KPN router.

Fifth remediation step

In contrast to the fourth remediation step, only three participants were able to reset the password of their router. One participant gave up after the website where this reset should take place did not respond for over 4 minutes. Three participants had no idea how to reset their router, even though a link is included in the notification email on how to perform this step.

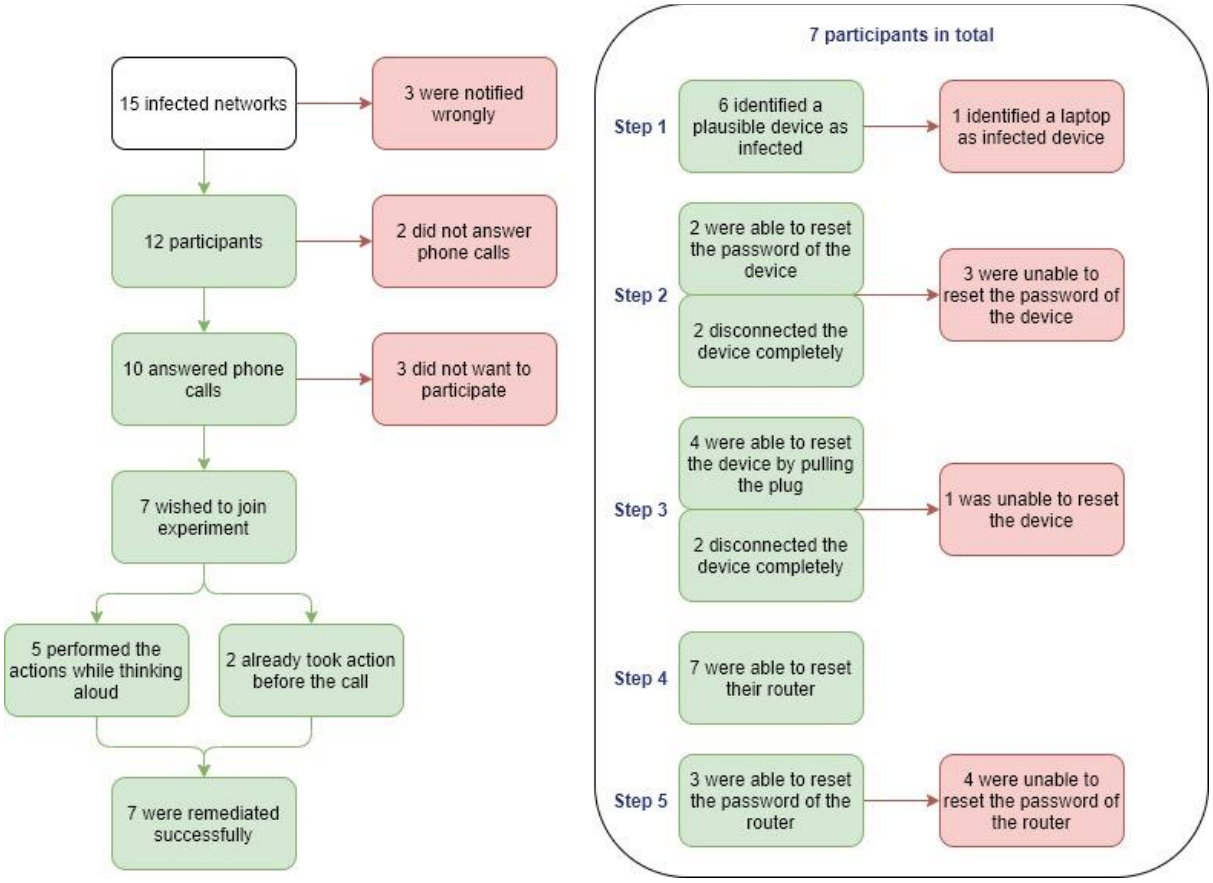


Figure 13. Pilot experiment results.

4.5 Processing of results

This section discusses how the captured data will be processed and conclusions can be created based upon this processed data. For each of the captured data variables, the variable is discussed, after which the methods for processing and analysing this variable are described. For all described variables, they can be separated over or adjusted through the characteristics of the end-user, namely *age, gender,* and *household size*. Before the sub-questions are answered, an overall summary of the experimentation period is displayed and discussed in chapter 5. That chapter discusses general findings such as the part of the approached end-users that wished to take part in the experiment in the first place.

4.5.1 Data on connected devices

The first data points that are of interest exist of the devices that end-users own and their perception of these devices. The following variables are measured:

- *Connected devices and identified infected device(s)*

Within a household, one or more IoT devices are connected to the internet, which end-users are asked to list during the virtual visit. This gives an indication of how many devices are present and what types of devices are often present. It is important to keep in mind that there is a difference between the set of devices that are actually connected to the network and the set of devices state are connected to the network according to the end-users. Some devices might be overseen by end-users and some devices might not be connected to the network, even though the end-users believe so. The Venn diagram in figure 14 clearly shows this discrepancy.

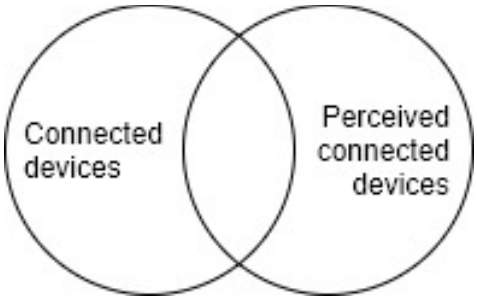


Figure 14. Venn diagram perceived and actual connected devices.

Related to the connected devices, are the devices that end-users identify as infected. The identified device can be compared to the data from Shadowserver and whether the clean-up efforts were successful can show if the correct device was identified or not. Descriptive statistics can show commonly infected devices and to what extent users are able to pinpoint the correct device.

Chapter 6 about end-user homes makes use of this data to sketch the target audience that ISPs are dealing with when they try to aid in remediating infection IoT devices. The data is compared with findings of previous research.

4.5.2 Data on remediation success rate

Although previous research was mostly based around the survival probability of end-users with help of the Kaplan-Meier estimator (Altena, 2018; Verstegen. 2019; Çetin et al., 2019a), this research does not consider the time aspect of infections and their remediation. As the virtual visit is supposed to be successful and it takes place one day after the notification, only an actual clean-up rate can be observed:

- *Clean-up rate of Mirai infected end-user home networks*

The success rates of previous studies can then be compared to the success rate found in this research, which points out to what extent current notification mechanisms have reached their potential. Here, it is assumed that the clean-up rate is no longer bounded by end-users' willingness to act and skill to perform the requested steps right. As we are present during the clean-up efforts, while making sure that the steps are performed correctly by the end of the virtual visit, only three barriers can keep the network from being cleaned successfully. The first explanation for this is that the end-user might not be able to perform the steps, even though we clearly explain to them in detail what to do. The second reason for failed clean-up efforts is that the infected device or present router/modem does not allow for the steps to be executed, for example due to the impossibility to change its password.

The third and last possible reason for failed remediation is that we ourselves lack in skill to understand the steps that should be performed. As there is a large heterogeneity in IoT devices, it is not unlikely that some for certain devices, it is hard to perform the requested steps. The clean-up rate is therefore deemed to be at its potential as far as end-users are concerned.

The clean-up rate in this research can show the effect of mobile support for end-users while they are performing the required steps. Moreover, it can show how much there is still left to gain in fine-tuning the notification mechanism. Chapter 8 makes use of this data to explain the overall performance of end-users on the remediation steps.

4.5.3 Data on performed actions

As users are asked to explain in detail what actions they perform to comply with the requested steps in the notification, the following dataset can be created from the answers of end-users:

- *Performed actions per requested step in the notification*

This list of actions shows the common actions that users take and for which of the required steps they spend the most effort. This will be elaborated upon in chapter 8. Moreover, this list can be used to look into the mental models of end-users considering their interpretation of the requested steps in the notification. The comparison between the notification protocol and the performed actions shows to what extent users are able to execute the protocol successfully, per step of the protocol. This is analysed and reflected upon in chapter 8, in which each of the cleaning steps are analyzed separately.

4.5.4 Exploration of the virtual visits

During the calls, end-users express their feelings and thoughts about the protocol, the requested steps in the protocol and the process as a whole. To make sense out of their stories, a thematic content analysis (TCA) is performed (Joffe & Yardley, 2004; Smith, 2000; Anderson, 2007), which results in several *themes* that can occur during the remediation process. TCA is a descriptive presentation of qualitative data, where common themes in the data subjects' stories are recognized and highlighted. Important is the minimal influence of the researcher in terms of interpretation (Joffe & Yardley, 2004). TCA is about the stated thoughts by the end-users, not about the researchers' feelings.

The method has an inductive approach, which means that there is no coding to categorize the expressions of the end-users up front (Smith, 2000). Instead, the virtual calls are transcribed and labelled afterwards. As there is no real expectation of what end-users could do or say during the calls, this method fits better than a deductive approach.

For performing the TCA, the stepwise approach listed by Anderson (2007) is used. After all virtual visits have taken place, multiple copies of the transcripts are created. Then, all relevant parts of the texts are highlighted, leaving out small talk. Parts are considered to be relevant if they in any way represent the end-users opinion about or experience with the experiment as a whole, and the steps in the protocol that they need to perform. An example would be the following statement by an end-user: "So I believe it has to be the NAS, as we just bought and installed it".

Within the highlighted areas, all separate statements are distinguished from each other, leaving several distinct expressions. These sections can differ largely in size. The next step is to put similar expressions together and to label these groups of expressions. The example above would receive the label: "Identifying infected device".

This process of separating and labelling should be done for several iterations. Some labels will be merged, and some labels will be split into multiple ones. In the example, the label might be made more specific to: “Time related identification of infected device”. Also, some expressions might switch labels during the process. Then, the work is left alone for a few days, after which the original transcripts are looked into again. The multiple revisions help in keeping the overview of the texts and labels. After several iterations, the number of changes to the labels start to decrease, which is a sign that the process is completed.

This process results in a set of common and less common themes, that can be discussed and related to existing literature to make sense out of the initial thoughts and statements by the end-users. To minimize the influence of our own perception and interpretation, the statements by end-users are copied literally, without translation or adjustments. As this type of analysis relies greatly on the independence of the researcher, minimizing abstraction is key. However, as it is also important to verify the researchers interpretation, the eventual labelled statements are translated for a wider reach.

4.5.5 Reactions to the notification mechanism

Next to the stated thoughts of participants of the experiment, which are analyzed through a thematic content analysis, several other occurrences take place during the experimentation period. Also, the potential participants who did not become part of the experiment have a story that is interesting to tell. These stories will be told with some descriptive statistics about the participation rate, the reasons behind participating or choosing not to and important stories about the notification process as a whole. The distinction between the analysis described in section 4.5.4 and this section reflect the figure 9 in section 4.2.2, where the blue part of the figure is analyzed through the TCA and the findings of the whole protocol are analyzed through descriptive techniques.

4.6 Limitations of the method

Although this chapter has explained the reason why the used method is fitting to answer the research questions, there are several limitations to it. This section elaborates upon the most important and impactful limitations of the method that is used.

4.6.1 Monitoring infections

One of the main limitations of the method is the certainty of the measured infections. Although Mirai has a known fingerprint and multiple honeypots are used to detect and identify Mirai infected network environments, the only data that can be measured is active scanning by the infected devices. If for some reason, a device only scans the internet for new bots for a short period of time, the likelihood of it reaching out to the honey pot is smaller. If a consumer network does not show up on any infection feed after the virtual visit has taken place, there is never a 100% certainty that the infection is completely gone.

On the other end of the spectrum, after the virtual visit, the consumer network may still show up on the infection feeds. The most obvious explanation for this is that the virtual visit did not lead to a successful remediation, and one of the steps was not executed successfully. As we trust on consumers ability to translate their actions into words for us, it could be the case that we believe the remediation should be successful, but the consumer did not act as was stated.

Another possibility for the presence of a consumers network environment is that the consumer owned multiple infected devices but only identified and cleaned one of them. As only the router can be seen in the infection feed, there is no way to exclude this from occurring.

Thirdly, even though Mirai has been cleaned properly, another malware could have taken over from it in the meantime. As there are many different variations of Mirai and other malware and malware is known to compete for control over IoT devices, this is not unlikely to happen. Because the signature for Mirai is out for the world to see, other malware might have copied this signature, which will then be identified as Mirai. However, if it is indeed a different type of malware, the cleaning steps might not be successful. If a network shows up on the radar significantly later than the time of the virtual visit, a device could be re-infected by Mirai, which means the clean-up was successful, but the device is still left vulnerable.

The last possible scenario that is considered for this research, is the presence of multiple routers in a consumers home. This could mean that we are measuring connected devices to router A, and the consumer is performing clean-up steps on router B. In many cases, none of the above possibilities actually occurs, but it is important to keep them in mind. Note that it is also possible that multiple of the phenomena above could happen with a single customer.

4.6.2 A need for trust

In a time where phishing mails are common, and supposed Microsoft help desk members call people all around the globe, it is hard to gain the online trust of an end-user. It is critical in this research that end-users have a minimum level of trust in the authenticity of the research. Therefore, in every choice that has been made, this level of trust has been decisive, such as the structure of the email notification, choosing the video platform Skype over the more obvious BigBlueButton, the choice to wear KPN clothing during the video calls or the inclusion of comforting words in the virtual visit protocol. As has been discovered in the pilot experiment discussed in section 4.4.5, even with these choices, trust is sometimes weak and fragile.

The need for continuous trust also has its dark sides. As end-users may at no time during the virtual visit feel like they have done something wrong, the calls require patience which may take a lot of time. As we do not have received professional training in the art of performing such sensitive experiments, the virtual visits will probably not take place in the most effective way. Although the pilot calls take away part of this inexperience, we are no experts. Moreover, it may be hard to get the answers that we are looking for from the end-users, as pushing more and more to get them to explain their actions in the level of detail we are looking for, is not worth the cost of losing them completely.

4.6.3 Observing from a distance

Even though the virtual visit protocol is designed in a way where consumers are in a safe environment, there is no way of removing the barriers of communication during the virtual visits. First of all, the consumer has to understand which actions they are taking, then they have to translate these actions into words, then we have to understand these words and translate them back into actions that the consumer took. Moreover, it could be the case that consumers do not want to share their real actions because they feel embarrassed or uncomfortable. In cases where consumers already took actions before the virtual visit, it can be the case that they forgot exactly what actions they took, although the protocol minimizes this by planning the virtual visits only one day after sending the notification. Lastly, even the proposed videocalls might not be able to create a safe environment for consumers, who might still not be willing to share all detailed information that is requested from them.

4.6.4 Impact of virtual visits

The virtual visit protocol is designed to keep on the side as much as possible. Still, consumers are inherently influenced by the virtual visit taking place in the actions they take. Because this research focuses on the whole protocol, it is important that consumer perform all steps of this protocol. Previous research, however, has found that most users do not perform all requested actions and also

not in the requested order (Verstegen, 2019; Altena, 2018). During the virtual visits, consumers are asked to perform the steps one by one, which is a different setting than what would happen without the call.

Moreover, the presence of a KPN employee and TU Delft student can create more willingness to act right with the users. This leads to an overestimation of the actions consumers take and the effort they put into these actions. This is important to keep in mind when answering the research questions, which are about the awareness and skills of consumers rather than their willingness to act.

4.6.5 Timeliness of method

The chance that an infected device only scans the internet for vulnerable targets during the weekends is not that big. Still, it is important to keep in mind that these devices will not be considered by the experiments in this research. Moreover, if devices start to scan during the weekends, their users will only be notified three days (at most) after the infection started. This means that the probability of users having acted before the virtual visit is highest during the calls on Monday, as they were notified on Sunday.

Another time related limitation is that the calls take place only one day after a notification is sent. There is a significant chance that consumers have not noticed their notification before the virtual visit, which means they will have to read it during the call. On the one hand, this could mean that users have less time to read the notification carefully, which means the actions they take are less precise than they would be in the case where they had more time to read the notification. On the other hand, consumers might put more effort into reading the notification carefully, as they might feel pressure from the KPN employee / TU Delft student that is present.

Lastly, the large impact of the countermeasures against Covid-19 should be stressed in this research, as they enlarge the probability that consumers are willing to take part in the research and take the requested actions as they have more spare time. Although the research questions are formulated in a way where these countermeasures are not present, their impact should be kept in mind during all stages of the research.

4.7 Ethical considerations

In this research, we are working with real people and their capabilities. To make sure no legal or ethical boundaries are crossed in the research, this section elaborates upon several ethical choices that were made.

4.7.1 Lack of intervention

During the virtual visits, consumers' actions are only intervened at the end of the call. This means that they might be struggling with performing the right steps in the right ways. Still, to make sure the data is usable, they will not be aided during this process. If consumers start to feel uncomfortable, we can intervene shortly and let them continue with the successive steps.

4.7.2 Unsuccessful remediation

Although the consumers are aided at the end of the call if this is deemed necessary for the clean-up to be successful, there is no guarantee that the remediation efforts will be successful. Due to the heterogeneity amongst IoT devices (Zimmermann et al., 2019; Forget et al., 2019) and their lack of user interface (Kolias, 2017; Anthi et al., 2018)., it remains hard to properly execute every step of the protocol. Thus, it could be that a consumer still owns an infected device when the virtual visit is finished. Although these cases are minimized by preparation and the cases are dealt with in an ethically acceptable way, it should be noted that unsuccessful remediation even with the effort of a KPN employee is an unwanted situation to say the least.

4.7.3 Data management

As we are dealing with personal data in this research, the data processes are in line with the General Data Protection Regulation (GDPR). For taking part in the research, all consumers are asked for their consent. For the sensitive parts of the experiments, such as the overview of their connected devices and the recording of the virtual visit, consumers give their consent separately. Contact details are only looked up right before they are needed, either for sending the notification or for the virtual visit. As this data can be found in the KPN subscription accounts, this means that the data about their contact credentials is not collected in the research. All data that is collected during the virtual visit, including the recordings, is stored on a KPN laptop within the KPN network and is used only for the purpose of this research. In processing the data, individual persons are not identifiable.

5. Experiment results

This chapter outlines the proceedings of the experiment period. Although this chapter does not answer one of the sub-questions directly, it gives an overview of what the experiment entailed, and how the data was processed. Section 5.1 is a short summary of the experimentation period that shows the different flows in the process that end-users could follow. Section 5.2 explains how the Thematic content analysis was performed. It shows how the obtained data was cleaned and how the themes were created. Section 5.3 is a summary of the flow of the infected networks and a list of the notable findings of the infection feeds during the experimentation period. The chapter ends with section 5.4 which concludes on the experiment process.

5.1 Course of experiment

Over the course of 10 weeks, lasting from May 24th until August 2nd, both the Shadowserver and Darknet feeds have been monitored for networks that show Mirai-like behaviour and are part of the big network of KPN. 59 unique IP addresses that belong to regular customers of KPN have been spotted on the radar, which shows an average of 0.84 unique IP addresses per day. Out of these 59 networks, 37 fell within the experiment period of which 12 showed up only during the weekends, and were thus not included in the experiment. Section 5.3 elaborates on these weekend-only networks. Due to a failure in the automatic emailing tool of KPN, 3 end-users could not be informed timely, and were thus not visited virtually. Chapter 9 elaborates upon this failure and the possibility of similar failures, which can shut down the anti-botnet cycle at the very start. As the tool of KPN does not only process the email notification, but also more critical actions, it took relatively long before the tool could be used again, as management had to be sure of the trustworthiness of the tool. Because Telfort customers were informed through a different tool, part of the infection feed could still be processed.

The period of 10 weeks can be split up into 3 periods, that can be seen in each of the figures in the following chapters. The first week existed of pilot calls (May 24th – June 1st), the seven weeks thereafter were the actual experiment period (June 2nd – July 17th) and in the two final weeks of monitoring (July 18th – August 2nd) there were no virtual visits due to an in-between holiday, and the movement of KPN to a new system which caused a temporary inability to send email notifications. This period is still considered in the research, as end-users were scanned after the virtual visits to monitor the cleaning rate and possible reinfections.

Total number of detected unique IP addresses in the monitoring period.	**59**
Total number of unique IP addresses in the pilot period.	15
Total number of unique IP addresses in the experimentation period.	37
Total number of unique IP addresses in the post experimentation period.	7

Table 2. Course of end-users over different periods.

Table 2 shows how the total number of unique IPs are divided over the three periods within the monitoring period. As we performed 17 virtual visits, 20 were lost during the experimentation period. This means that the experiment had a response rate of 46%.

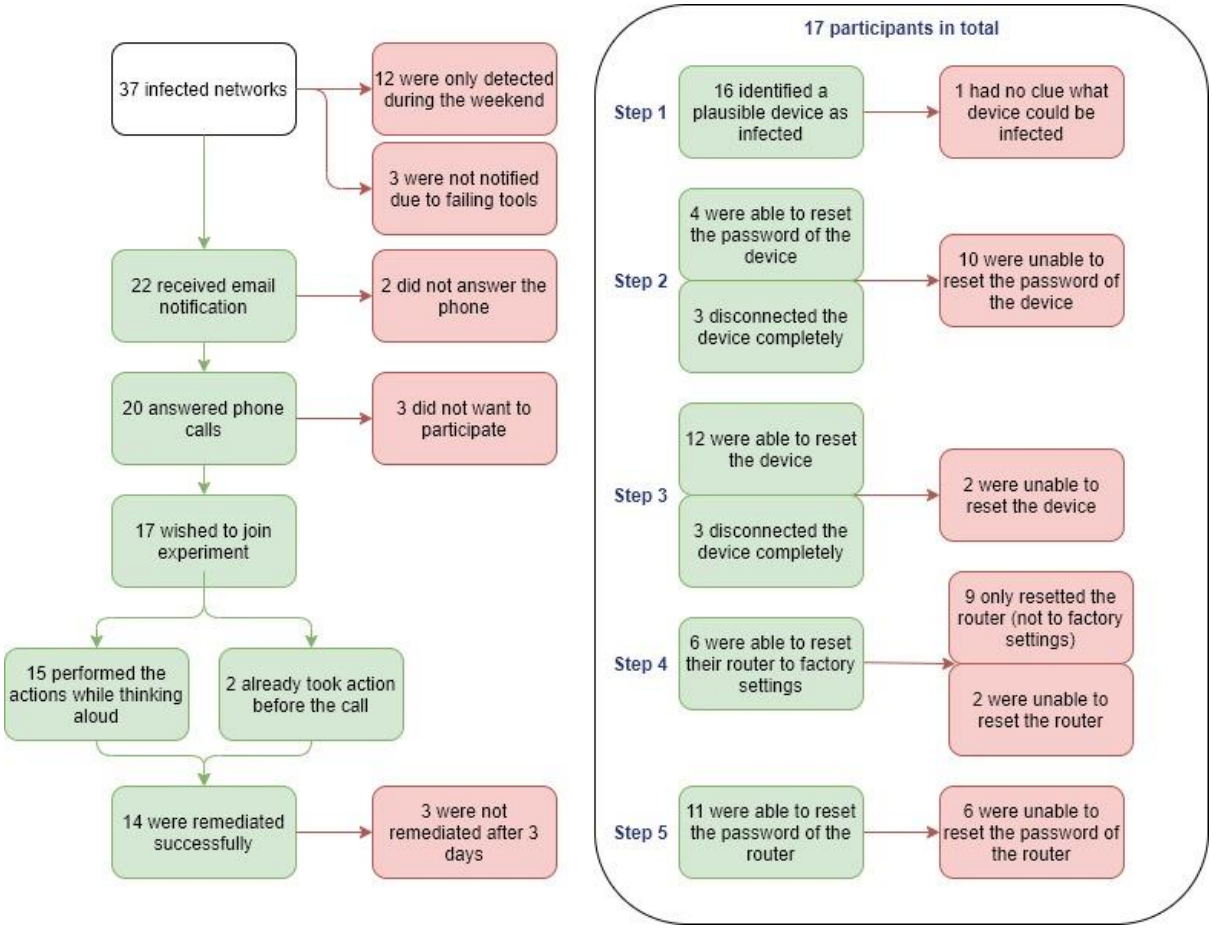


Figure 15. Results of the experiment.

Figure 15 shows the different states that end-users could have during the experimentation period and why the 20 end-users were not part of the experiment. It has the same structure as figure 13, which shows the same states for the end-users in the pilot calls.

As figure 15 shows, 12 networks were only detected during the weekends, which means they could not be part of the research. Another 3 network owners could not be informed due to the technical issues at KPN. The end-users that own the 22 networks that remain were all informed about the presence of Mirai in their network, by using the email in appendix E & F. Of these end-users, 20 also answered the phone one day later and 17 wanted to participate in the research, which is a participation rate for any experiment. This rate shows that participation was indeed attractive as we had hoped, and end-users prefer our presence while they perform the steps. This comes down to approximately 0.5 participants per day, as the experimentation period lasted from June 2nd to July 17th, 5 days a week. None of the 17 participants had an issue with the recording of the call. Chapter 7 elaborates further upon the reasons why 20 end-users were not included in the experiment. The next section is about the way the recorded calls have been processed.

5.2 Data processing

As has been described in section 4.5.4, the lion's share of the virtual visit is analyzed through a thematic content analysis, based on the stepwise approach by (Anderson, 2007). Because we are interested in both the cleaning steps protocol as well as the notification mechanism as a whole, the first round of labelling separated the protocol steps from any other comments and thoughts by the

participants. This meant the labels “step1” to “step5” existed, which were specified in a later iteration to different ways in which participants dealt with performing the corresponding step of the protocol.

The first round of iteration had a set of only 8 labels, consisting of the 5 steps, a lack of cooperation by the end-users, the idea that regular protection should suffice against Mirai, and a lack of support. These 8 labels were the base for the 25 different labels that we added to the end-users statements in the last iteration (the disconnection of a device has been put under 3 different subjects as well as the usage of the email notification by the end-user). The eventual labels that were used to describe, summarize, and explain the thoughts of end-users after several iterations can be found table 3. As can be seen, the 8 labels have been changed into subjects that are now summarizing themes for more specific phenomena.

Subject	Theme	Number of participants
Step 1. Identify the infected device	Taking stock of IoT devices	16
	Device malfunctions	8
	Process of elimination	12
	Time related reasoning	8
	Make use of the email notification	8
	Malfunctioning not due to Mirai	3
	No clue which device is infected	1
Step 2. Reset the password of the device	Successful in resetting the device’s password	4
	There is no password setting	4
	Stop using the device	3
	No help from brand	5
Step 3. Reset the device	Successful in resetting the device	13
	Unsure to what extent memory is cleaned	1
	Stop using the device	3
Step 4. Reset the router	Successful in resetting the router	6
	Too much effort to change the settings	3
	Factory settings are the same as just resetting	9
	Make use of the email notification	2
Step 5. Reset the password of the router	Successful in resetting the password of router	11
	The software does not work properly	2
	Make use of the email notification	6
	Unable to change the password	6
Reason for (non)compliance	Too much effort	6
Reason for (non)compliance	Indication of distrust	3
Reason for (non)compliance	Lacking email communication	3
Reason for (non)compliance	Anonymous calling is an issue	5
Lack of support	Lacking support by device brand	4
Alternate solutions	Regular protection is enough	2
Alternate solutions	Disconnection as a solution	3

Table 3. Themes found in cleaning steps protocol and notification mechanism experiences.

After three weeks of virtual visits, 8 end-users took part in the full virtual visit, for which labels were created in an iterative process. The additional 9 participants only added one more label to the existing set, namely an end-user without a clue which device could be infected. In TCA this is known as saturation principle, which is reason to assume the majority of possible outcomes has been captured (Wolff et al., 2010, Mason, 2010). Together with the time boundaries and the change in software at KPN to send notifications, this was reason to close the experimentation period.

Figure 16 shows the saturation of themes for additional end-users that participated. Note that this figure shows the saturation in more detail than we used during the experiment, as the labels were only created for each participant from the 4th week. The first 8 participants where labelled at the same time.

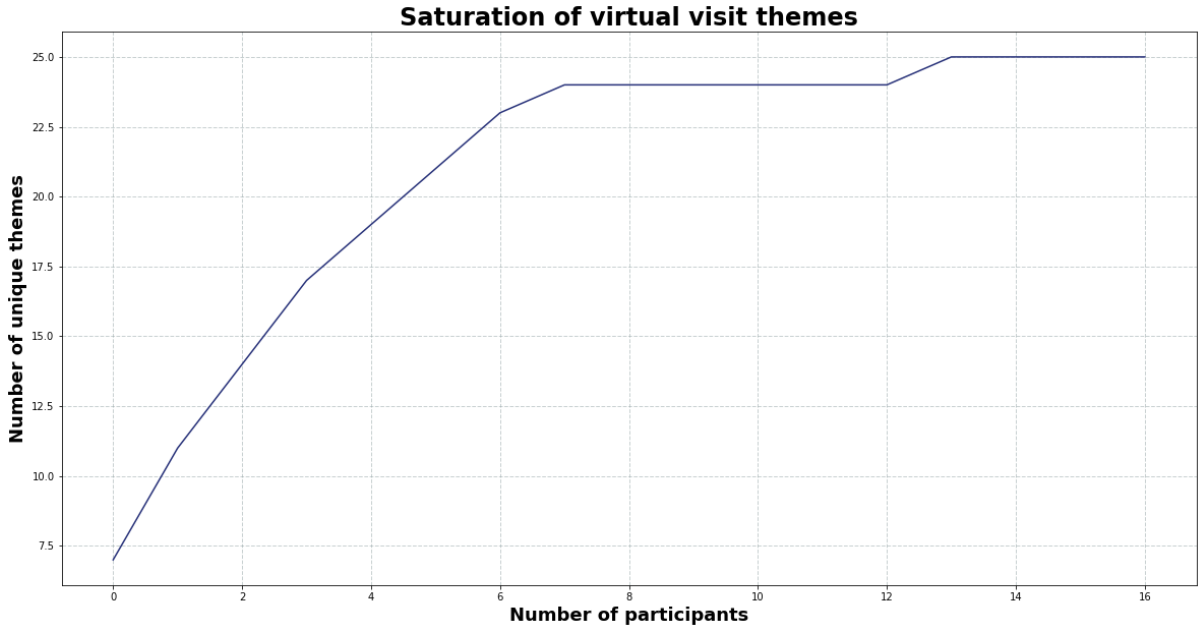


Figure 16. Level of saturation of virtual visits over time.

The following chapters will make use of the themes that were a result of the thematic content analysis, together with findings of the customers that did not want to be part of the experiment, for which the TCA is not possible, as there are no transcripts. However, notes were kept of the reasons why a full participation did not take place for these customers, which is useful information for answering some of the sub-questions and therefore, the main research question.

5.3 Tracking results

Figure 17 shows the cumulative number of unique IP addresses that have been detected over the whole monitoring period. Note that the pilot phase, experimentation period and the post experimentation period can are visualized within the figure and forthcoming figures. Table 4 shows the min, median, mean, and max number of old, new and total IP addresses that were detected each day. Note that the median of new IP addresses is 0.5 as the median was exactly in-between 0 and 1.

	Old IPs per day	New IPs per day	All IPs per day
Min	0	0	0
Median	3	0.5	4
Mean	3.87	0.84	4.71
Max	10	7	11
Total	271	59	330

Table 4. Infection feed of Mirai infected IP addresses over time, divided over old and new detections.

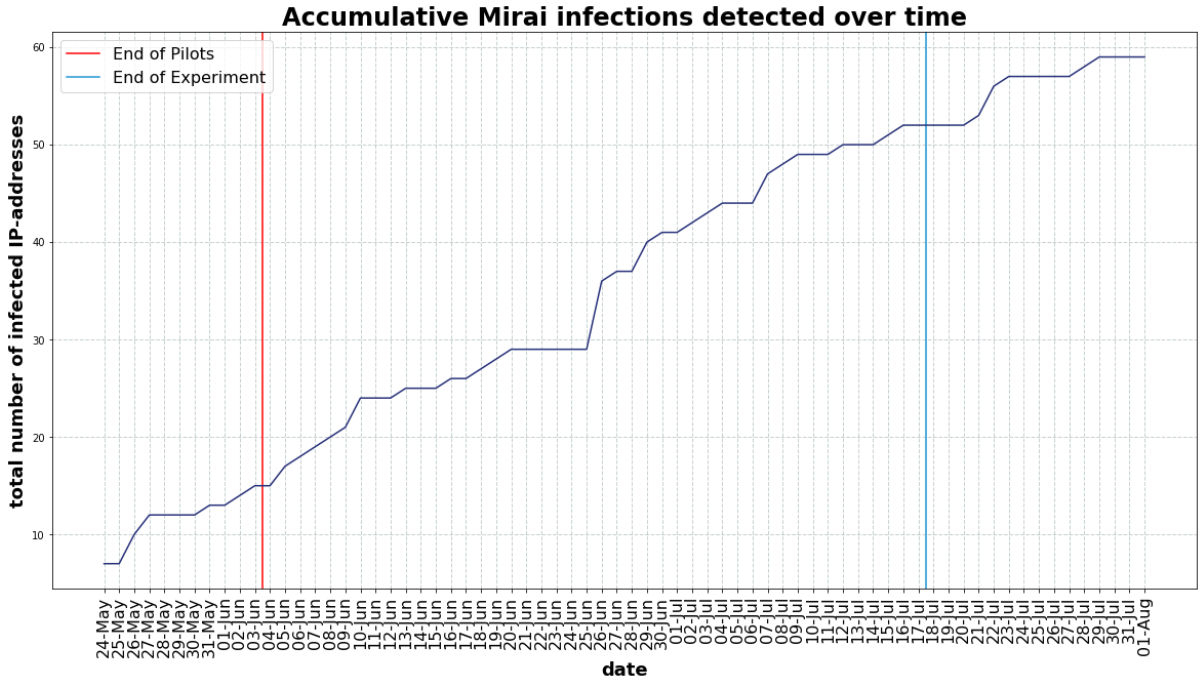


Figure 17. Cumulative Mirai infections detected over time.

Figure 18 shows how many infections showed up on the radar during the experimentation period, distinguished over old and new detections. Old detections are IP addresses that have been detected before, within the experimentation period. This is the reason why the first day (24th of May) only has newly discovered infection, as no infections were mapped before then in this research.

Note that we did not receive an infection feed from Darknet about the infections on the 18th, 19th and 20th of July. As Darknet has constantly and consistently contained all networks that Shadowserver reported (and often more), the data of Shadowserver has not been used on these dates either, as this would be showing only partial data. Moreover, Shadowserver did not show any new information on these days, so only a part of the old IP addresses is missing from the graph.

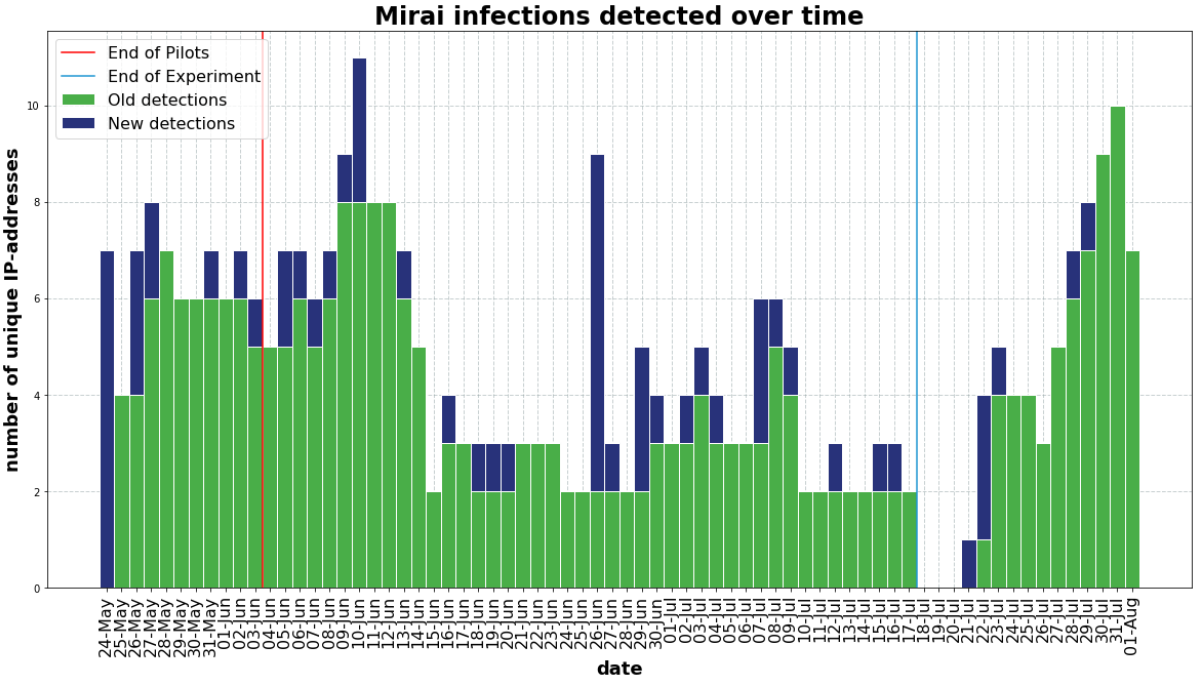


Figure 18. Mirai infections detected over time.

Note that figure 18 shows data on every single day of the monitoring period, including weekend days. As there is no activity at the abuse desk during the weekends, IP addresses that only showed up on either Friday, Saturday, or both, could not be part of the experiment. As the IP detections are checked the day after the detections take place, Sunday detections are processed on Monday, but Friday detections cannot be detected on Saturday, as the abuse desk is closed. We will still refer to these days as weekend days.

It is notable that 12 networks out of a total of 59 networks fit into this category. Figure 19 shows the new infections over time, using a different colouring scheme for workdays and weekends, to display the significant role of weekends in the detected infections of Mirai.

A speculative reason for these weekend-only infections could be that people often purchase new products on Friday evening or during the weekends, which would explain the peaks during the weekends. However, as these detections disappear again after the weekend, something else is happening too. A reason for the disappearance could be that the end-users have a quick trial with their new device to see if it works properly, and only then install it in the right way, by setting a non-

trivial password and such. This could result in a situation where their newly obtained products are infected for a very short period of time. However, only 1 of the participants in the experiment stated that their infection was a result of their delayed installing of the product. There is no theory that could explain why end-users would have more secure behaviour in weekends than during workdays.

Another speculative explanation is that some products are only used during the weekends. If a certain device is only switched on during the weekend, the malicious traffic is only detected in that period of time as well, when it is turned off again, the infection is not detected. As we were not able to speak to the weekend only infected end-users, there is no data on what devices these end-users own. The screenshot tool did not show any notable devices for the weekend-only IP addresses (4 IP-cameras, 4 unknown devices and 4 IPs were not in the system), which means it is relatively unclear what type of devices are in this category.

A last, highly speculative possibility is that attackers have adjusted to the behaviour of the defenders. As ISPs such as KPN only operate during workdays, it could be that attackers have changed the orders to their army of bots to operate differently during the weekends. Attackers have an incentive to do so, as their bots are removed less often during the weekends. This does not explain, however, why the same bots do not turn up every weekend, but often only once. In any case, it could be interesting to look into the weekend only infections to find out what the reason is for their existence in future research.

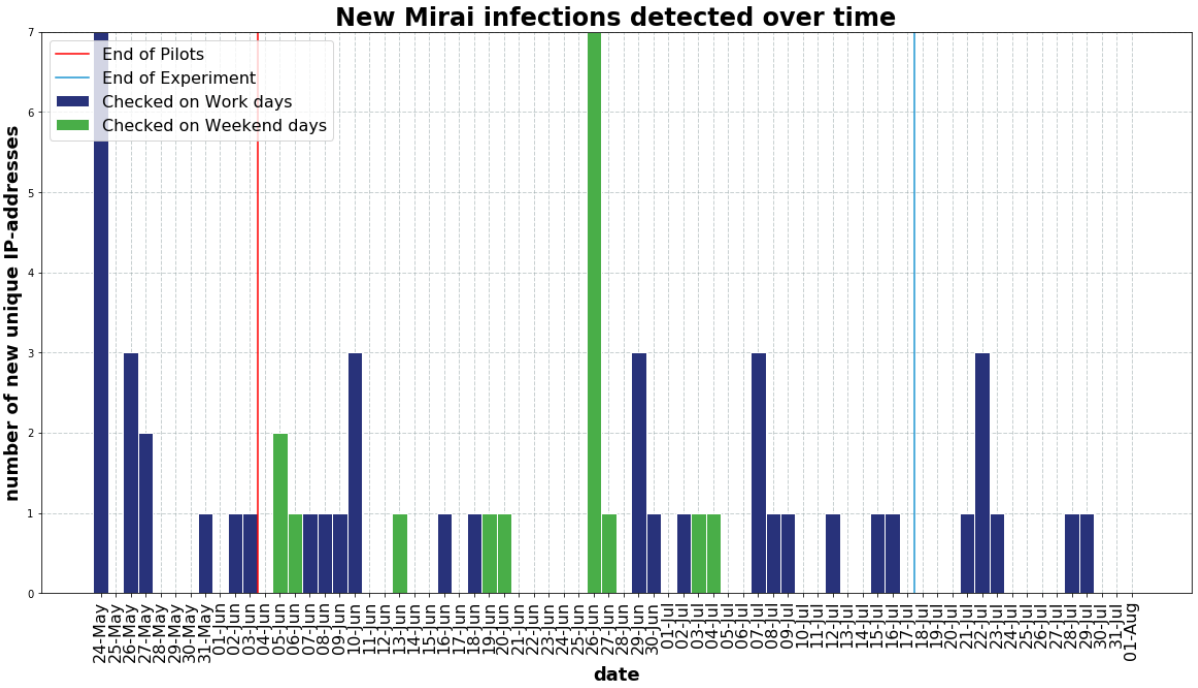


Figure 19. New Mirai infections detected over time, distinguished over workdays and weekends.

6. Crime scene investigation

This chapter is about the end-user network settings and elaborates upon the target audience for notification mechanisms. The chapter answers the first sub-question: *(SQ1) What does the crime scene of Mirai infected Internet of Things devices look like?* Section 6.1 is about the characteristics of the infected end-users, and how this relates to findings in earlier research. Section 6.2 specifically zooms in on the devices that end-users state they own, and section 6.3 is about the supposedly infected devices. The chapter ends with a conclusion in section 6.4 that answers the first sub question.

6.1 End-users characteristics

Although the focus in this research is mainly on an in-depth understanding of the processes that take place at end-users homes and thus does not include enough participants to create statistical findings on demographics of end-users, such as their age, gender and household size, this data has been measured and can be compared to findings in previous studies. It is important to keep in mind what the expected target audience looks like, when designing and executing protocols such as the notification mechanism for Mirai infected Internet of Things devices.

Small businesses

First of all, although the research supposedly only included commercial customers of KPN and Telfort, and not business customers and such, the virtual visits dug up that there is a relatively large set of small businesses running on regular customer connections. This regular connection is less costly and thus attractive to small businesses that do not need a large connected infrastructure within their business. Moreover, these businesses are extra prone to become infected, as they often have an incentive to make use of devices that are known to be vulnerable to Mirai, such as security cameras and NAS devices. 3 out of 17 networks were located within a business location, which were a storage location, shop, and a restaurant. The household size of these participants is coded as NULL, to be sure that they are not considered. Notable, is the participation rate of businesses in both the pilot experiment as the real experiment of this research, which is 100%. Although we cannot be completely sure that all of the non-participating end-users were actual households, their email-addresses and the short conversations over the phone suggested that these were regular household networks.

A high level of cooperation by small businesses can be expected, as businesses have a financial incentive to keep their network clean and working properly. On the other hand, small businesses have more reason to be careful with unexpected calls and might have a higher level of distrust.

Portrait of end-users

Verstegen (2019) has pointed out the overrepresentation of male end-users in the set of Mirai infected KPN customers, which can also be seen in this research, as only 1 out of 17 participants was female (figure 20). However, the household size data shows that these men often are not alone within a household. As only 2 out of 14 household owners lived alone (there were 3 small businesses), it can be stated that men are overrepresented in being the responsible person for online security, but not in being the victim of Mirai infections. However, it remains the case that males are overrepresented in the set of participants.

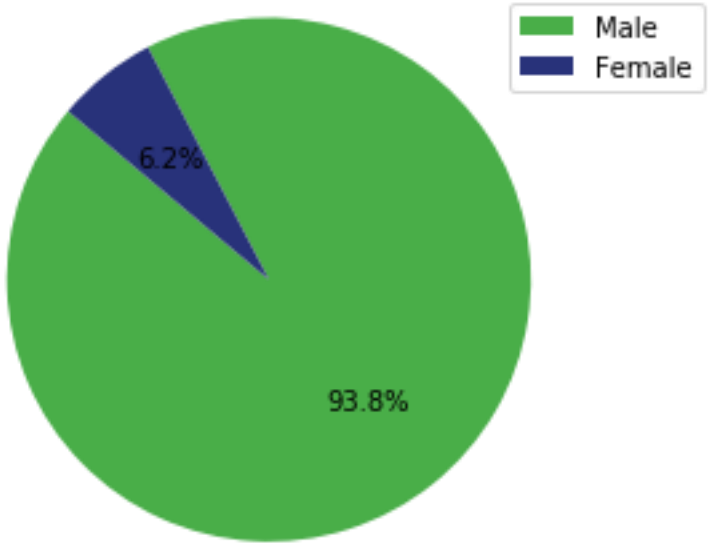


Figure 20. Pie chart showing gender distribution of end-users.

Considering age, Verstegen (2019) has found that Mirai victims are relatively young, but that end-users that own an infected device have wide-spreading ages. This is not different in this research. Table 5 shows the descriptive statistics on the age of the participants and figure 21 shows a distribution of age across the participants. Like in Verstegen (2019), the age ranges from 21 to a high 80, which is unexpected as Mirai focuses merely on IoT devices, which often have a target audience that is much younger.

The relatively young age of Mirai infected end-users has not been explained yet. As we have two additional variables in this research compared to Verstegen (2019), we can speculate on the reason for the existence of the young age of Mirai infected end-users. A possible explanation for the overrepresentation of certain groups in the set of infected end-users could be that these specific groups are more interested in IoT devices, and thus own more of these devices (this is elaborated upon in section 6.2) Another explanation could be that these specific groups of end-users live in a household that is relatively big (distribution of household size can be seen in figure 22), and therefore own relatively many devices. Because the data gathered does not meet the minimum conditions that are needed to execute a Pearson correlation test (n >30, normal distributed data), we have executed a Spearman correlation test as a base for our speculations of household size and the number of owned devices as explanations for overrepresentation of certain age groups in the set of infected end-users.

No correlation is found between age and household size (r-value: -0.314, p-value: 0.275), as well as age and number of owned devices (r-value: 0.194, p-value: 0.455). More data could aid in creating stronger findings, but the first impressions point towards a lack of explanation.

	Age
Min	21
Median	49
Mean	50
Max	80

Table 5. Descriptive statistics on the age of infected end-users.

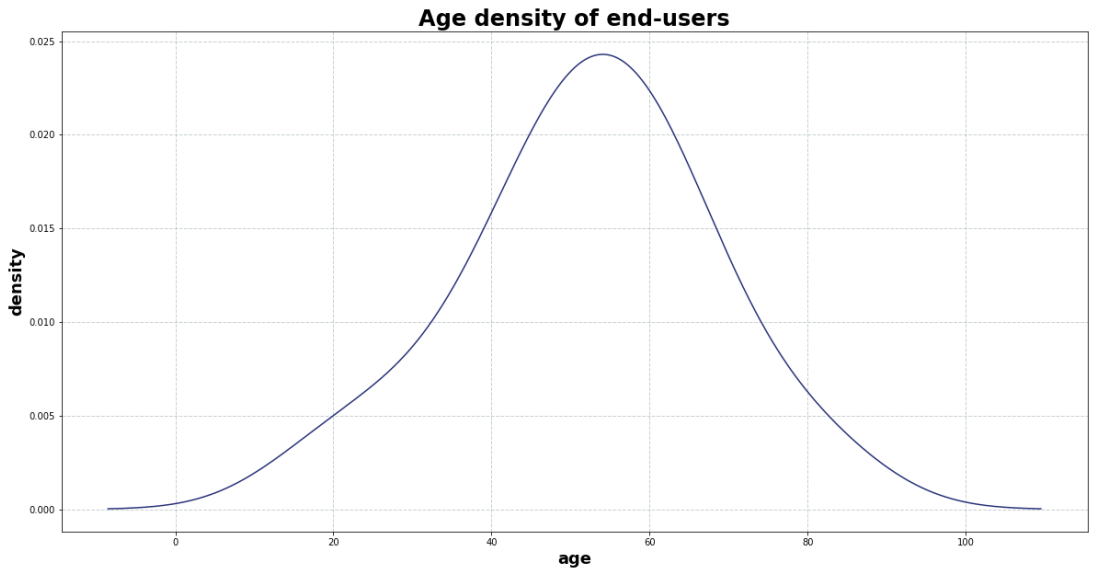


Figure 21. Distribution of the age of infected end-users.

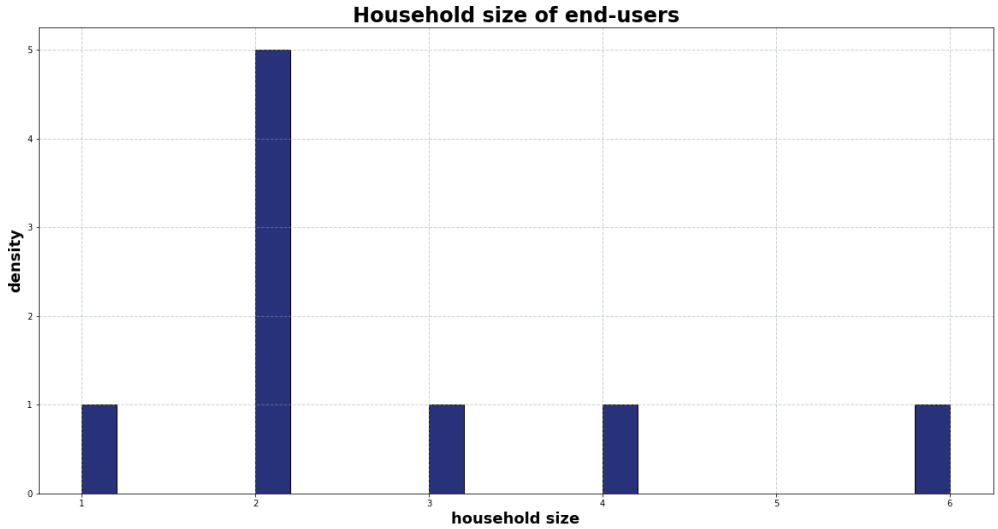


Figure 22. Distribution of the household size of infected end-users.
(the 3 small businesses have been removed)

In future research, it is important to incorporate both household size and the number of IoT devices that end-users own into the analyses. These two variables should be kept in mind when researching end-users, their properties, and their behaviour. Although we have not shown the exact impact of these variables in the remediation processes, it is likely that they influence the thinking processes of end-users during their performance.

6.2 Present devices

As a beginning of understanding the processes that take place in end-users' homes, it is important to look into the devices that are present at these end-users' homes. Although research is exists on analyses into infected devices (Antonakakis et al., 2017; Çetin et al., 2019b), it is important to keep monitoring the types of devices that become infected as time passes, to make sure the advices are up to date. To be able to pinpoint the infected device, end-users must first have knowledge on all their connected IoT devices, which can cause problems.

First of all, if a significant part of the end-users owns so many devices that it can be hard to pinpoint which one out of those is infected, this can be a real barrier for cleaning an infected device. Secondly, owning many IoT devices can increase the probability that one or more of these devices gets infected. If this is the case, the end-users that have a hard time in pinpointing the infected device could be overrepresented in the set of infected end-users, which means we underestimate the difficulties that end-users face while performing the cleaning steps.

Another reason why it could be important to map and understand what IoT devices end-users own, is to get an understanding of the extent to which end-users care about security in their IoT products, and which properties of a device are most important to end-users. Future research could perform a stated preference method map the importance of these properties.

This research created a start in looking into what devices are present at end-users homes, by asking end-users to list the connected devices in their network, while hinting at IoT devices by stating "Mirai often infects devices such as a DVR, a security camera or a printer". Surprisingly, 7 end-users could only recall owning a single IoT device, which was reason to believe that this device should be the one that got infected. Figure 23 shows the number of IoT devices that end-user stated they owned.

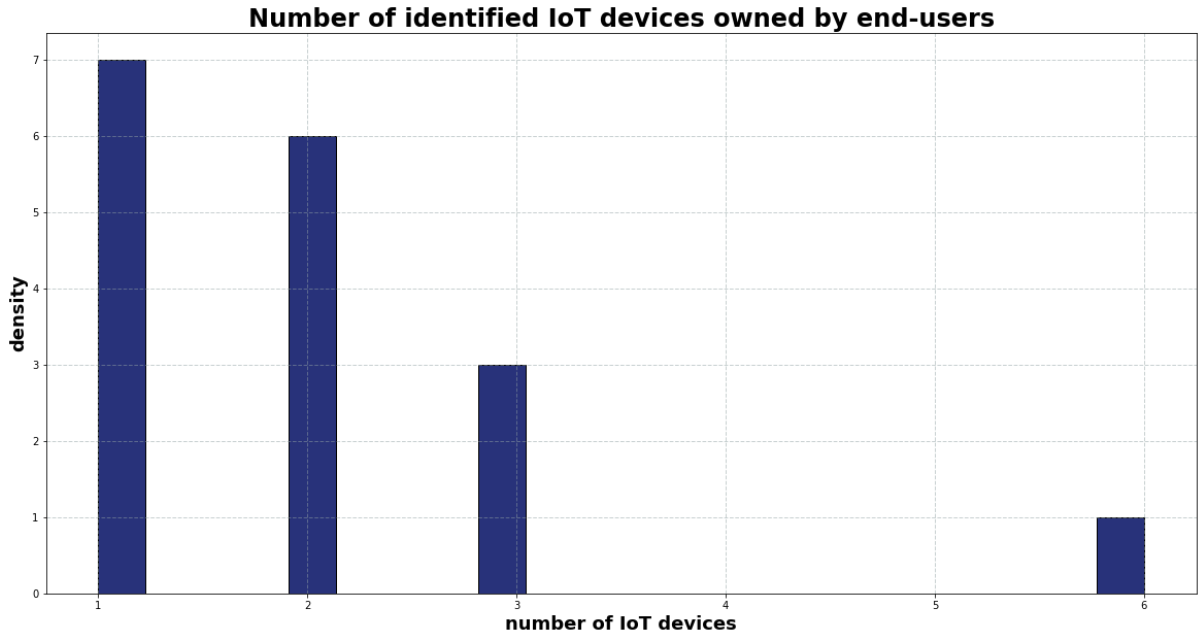


Figure 23. Number of owned devices by end-users.

Most end-users stating they own only one or two devices can be both a blessing as well as a threat. On the one hand, owning only one device takes away the difficulty of finding the origin of the Mirai infection, which should result is more successful cleaning attempts. On the other hand, end-users might overlook certain other devices that could be the cause. As end-users immediately set their mind

on a certain device, it can be hard to stay open-minded about other possible IoT devices that are present in their home, of which end-users might not even know that these devices are connected to the internet. The hinting in the email notification towards cameras, printers and DVRs can create a bias in end-users mind to easily oversee other IoT devices. Table 6 shows all the stated devices by the end-users, and how many end-users owned this device. The IP cameras are probably overrepresented in this set, as they were often the cause of infection (7 out of 10 cases).

Device	**Number of occurrences**
IP camera	10
FritzBox	4
Smart printer	8
Smart TV	3
Solar panels	3
Smart speaker	1
Media player device	1
NAS	2
Doorbell	1
Smart sprinkler	1
Total	34

Table 6. Present devices and number of occurrences.

To verify the list of IoT devices that end-users gave us during the virtual visits, we tried to make use of the network mapping tool that KPN owns. This tool supposedly identifies all connected devices, accomplished by some information on the type of device, and sometimes the manufacturer and other details. However, this tool can only be used for KPN customers (excluding Telfort customers), and turned out not work be suited for our purposes. Only in one case, the tool was successful in identifying two devices that were connected to the network of an end-user. However, even in this case, the tool did not identify any usable information except for the Experia box V10 router.

This means that there was no way to check whether the end-users actually had a correct perception of what is connected to their network and we had to trust on end-users' thoughts. During the virtual visits, end-users often admitted themselves that they forgot one or more devices while they were taking stock of their devices. In a later stage during the virtual visit, 4 out of 17 end-users remembered another device, either caused by chit chat or their movement through the house for example because of moving to the router. In 3 out of these 4 cases, the device identified in a later stage was labelled as infected by the end-user as well as the screenshot tool.

This shows the possibility of incompleteness in listing present IoT devices within households. Still, in most cases (16 out of 17), the infected device was identified with success. Section 6.3 elaborates on the infected devices that were present in end-users networks. Section 8.1 elaborates upon the different tactics that end-users used when identifying the infected device.

6.3 Infected devices

We can hardly ever be 100% sure about the specific device that got infected with Mirai. Only when the actual packages can be analyzed, the infections can be linked to a device. Still, it is important to look into supposedly vulnerable devices and to keep updating this list of vulnerable devices. Although KPNs network mapping tool did not suffice in pointing us in the right direction when the infected device was pinpointed, the screenshot tool used for infected networks jumped in as a navigation tool.

The screenshot tool (explained in section 4.1.2) is sometimes able to label an infected network in the feed as a certain device, which means it can be used independently of KPN or the end-users. Therefore, the supposedly infected devices can be checked for each IP address that enters the infection feed. Figure 24 shows the distribution of types of devices that were the cause of the infection according to the screenshot tool.

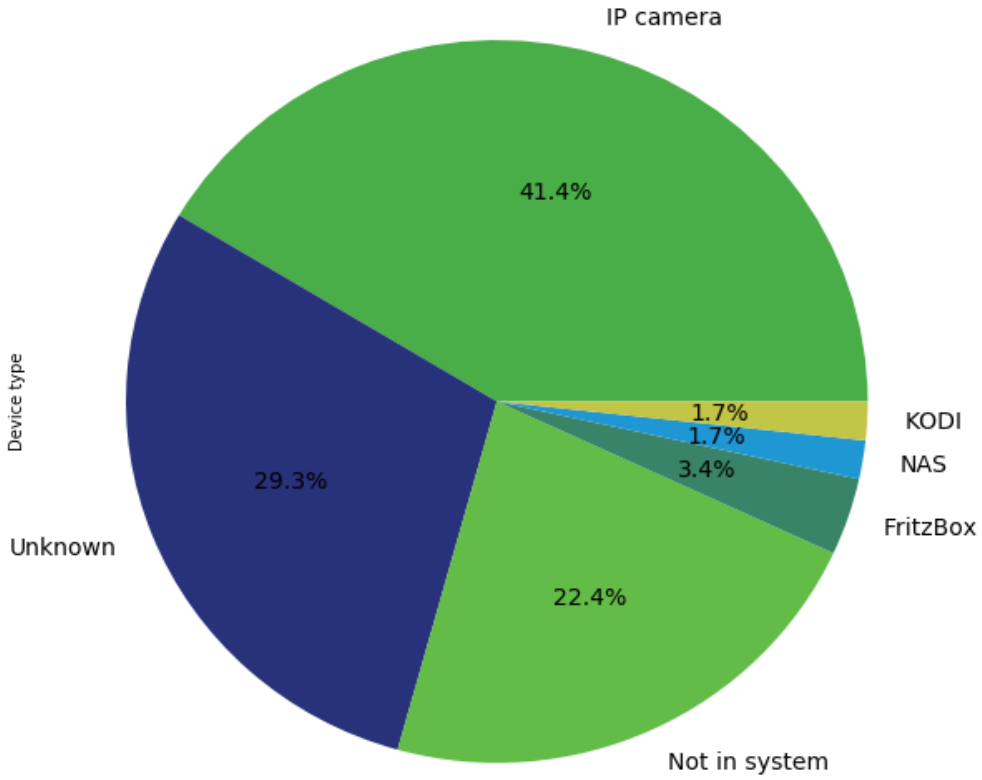


Figure 24. Distribution of infected devices as stated by screenshot tool.

We have good reason to believe that the tool is accurate when it is able to put a label on a device, as it was able to pinpoint the same device that the end-user stated as infected in all 12 cases that it was able to label the device in the first place. It seems that when the screenshot tool is able to put a label on a network, it is consistent with the end-users thoughts. Future use of the tool could ensure the accuracy.

As has been mentioned before, there is no real way to ensure that the “correct” device is identified, but in this research 12 end-users consistently named the same device as the screenshot tool independent of each other. Of the 17 participants in this research, the infected devices were distributed as shown in figure 25.

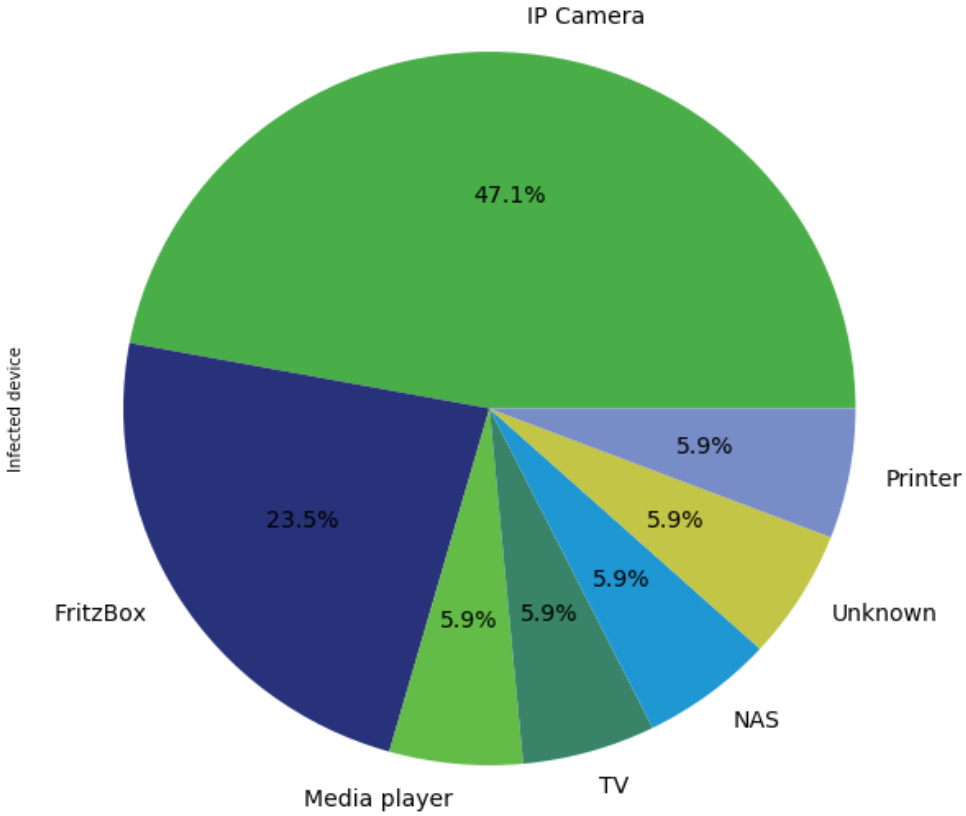


Figure 25. Distribution of pinpointed infected devices.

These findings are also in line with the studies of Antonakakis et al. (2017) and Çetin et al. (2019b), where routers, DVRs and IP cameras were the main source of infection. Except for a Mede8ter and an Epson printer, no new or rare devices were identified as infected.

However, we have supposedly found infections in a type of router that has never showed up before considering Mirai infections. The FritzBox router is known to be one of the safer devices and does not make use of simple standard passwords, which is why its four occurrences cause confusion. Although most known types of Mirai are not able to infect such a device, a possibility exists that a recent version of Mirai can exploits other weaknesses that could be present in a FritzBox.

Another potential reason for the appearances of the FritzBox in both the end-users thoughts as well as the screenshot tool is that there is actually an insecure device that is connected to the FritzBox, but we can only see the router in the screenshots for some reason. This does not explain, however, why end-users would mention the box as the cause.

In all four cases (and one case during the pilots), the end-users network did not show up from the day after the virtual visit, while the cleaning steps were only performed on the FritzBox router. This is why we have reason to believe that there was no other device present that was infected instead of the FritzBox. Future research should be able to find out how and why a FritzBox can be vulnerable to Mirai, as there is no proven explanation at this point. Because the FritzBox is the only rare device that showed up multiple times on the radar, future research could be specific to this device.

7. Reactions by end-users

Section 5.3 has described the way in which the virtual visits have been processed into themed expressions. In this chapter, the experiences with the full notification mechanism are described. To do so, the identified themes that do not correspond to a specific step of the cleaning protocol are discussed here. This includes the experiences during the experimentation of end-user reactions from the moment they enter the infection feed, until the end of the virtual visit. Note that this chapter does not describe the performance of end-users at taking the necessary steps in the cleaning protocol. This is discussed in chapter 8.

This chapter answers the second sub-question: *(SQ2) How do end-users react to an email notification about their infected Internet of Things device(s)?* Section 7.1 is specifically about the way the email notification is received and perceived by end-users. Section 7.2 is about the way end-users reacted to being called after receiving the notification and section 7.3 concludes by answering the second sub question in a structured way.

7.1 Reaction to the email notification

End-users are firstly made aware in the remediation process after they are forwarded an email notification about their Mirai infected Internet of Things device. Their reaction is partly measured in the day between the forwarding of the email and the virtual visit that takes place the day after.

Another part of users reaction to the email can be seen during the virtual visit, in which multiple users tend to start with expressing their thoughts on the email notification, if they have seen and read it, that is. This section elaborates upon the expressed thoughts by end-users considering the email notification process and the notification itself. The themes that arose from the thematic content analysis are used as a base for this section. All themes can be found in table 3 in section 5.2.

7.1.1 A lack of trust

For any theory or hypothesis to be useful, it should be falsifiable (Popper, 1934). It is impossible to verify something for 100%, as there is always a chance that some evidence piece is missing from the puzzle. This can also be seen in end-users behaviour when receiving an email notification about their infected network. Although end-users could call different departments within the sender of the notification, and they could rely on certificates and stamps to decide whether the notification is real or not, there is no way of excluding the possibility that the email is not what it seems to be.

Even though both the email notification as well as the virtual visit have been created around the idea of trust (section 4.6.2) and had a personal note which should have increased trust (Li et al., 2016a), multiple end-users have doubted the authenticity of the notification mechanism. Stock et al. (2018) warned for this to happen, as email is not the regular channel used for important notifications such as these. Another explanation could be the lack of trust that end-users have in KPN as a business (Redmiles et al., 2016). Even though we tried to create a combined email and call and stating our name, trust was still a common issue.

During the pilot phase, one end-user contacted KPNs helpdesk to verify the authenticity and 5 end-users spoke aloud about their doubts about the email notification and call during the pilot phase. In the real experiment phase, this was not different. The theme of trust returned in 3 out of 17 calls, but also in the 3 cases where the end-user decided not to take part in the research.

Again, as people often use multiple email accounts and receive loads of emails, it can be hard for them to find and open the correct email notification. If this process seems to be unprofessional, the trust level sinks. As one end-user stated: “I am not sure about this, things do not seem the run flawlessly”. Even though we tried to improve the protocol after the pilot calls, variables remain that could cause trouble for the process to run smoothly.

One participant explicitly decided not to be part of the research and hang up soon after the start of the virtual visit took place. Even though we do not ask any sensitive information from the participants, they can be afraid that they are asked more sensitive things stepwise, until they do not recognise that they are actually sharing sensitive information. As one end-user stated: “I know how this can work, you ask more and more of my data until it is too late”.

Although it was not part of the protocol, we have elaborated upon the research in 4 out of 17 calls to ensure the participant that the call was no danger in any sense. If we got the opportunity to do so, the participant was satisfied, but in 3 cases (not part of the experiment), the customer hung up before we had the chance to elaborate. None of the participants had an idea of what was needed to increase their trust level, as trust is “a feeling”, and “something you cannot explain easily”.

In many ways, the distrust of end-users is logical and useful for their well-being. As there are many attackers who wish to make use of end-users trust, it is important to remain careful. However, for authentic actors, it only gets harder and harder to remain trustworthy. This problem only gets worse, because attackers have an incentive to become more trustworthy and to imitate real companies as good as possible. For end-users it then gets harder to distinguish the real from the fake, which leads to a limited trust for all parties.

7.1.2 Disconnection as a solution

A returning theme at the start of a call was that end-users already disconnected one of their IoT devices, as they were shocked by the email notification. Both in the pilot calls (2 out of 7) and in the experiment calls (3 out of 17), end-users decided to disconnect the device they identified as infected. These end-users had in common that they felt relieved when they picked up the phone so they could explain that they took effort to solve the problem, as they took the issue very seriously. As one end-user stated: “I immediately threw the camera out, I mean, a virus is just terrible”. Another reason why these end-users decided to disconnect their device fully, was the malfunctioning of the device. One end-user stated: “That thing just stopped working as it should, so the decision was easily made”.

Although disconnection is not part of the cleaning protocol, it is a successful replacement for both resetting the password of the device and resetting the device itself. In all cases, the identified device was an IP camera, and the network was not seen again in the infection feeds. In literature, this type of behaviour has been seen before (Verstegen, 2019) and it could be argued that it could become part of the possible remediation steps certainly after multiple reinfections.

7.1.3 Lacking communication

Related to the trust issues, a failing email communication can cause serious problems for successful remediation. In one case, an end-user had received multiple emails, each stating a different issue and different solutions to that issue. The end-user identified this second email as the correct one and started reading out loud the steps that were stated in this email, which were not the correct ones. Luckily, the right email was forwarded as well, so the end-user could still move towards the correct steps, but without our intervention this would not be the case.

Another technical issue that caused problems for the remediation process, was a migration of the customer databases at KPN. This technical change could have caused the wrong customer to receive an email, or even being put into quarantine, which is not desirable. That is why we decided to hold the experimentation period during this technical failure (section 5.2), which lost us 3 potential participants.

7.2 Reaction to the call

Next to feelings about the email notification process and the way the email is structured, end-users also shared their thoughts on the virtual visit itself and how it took place. This section discusses end-users' thoughts on the call by making use of the themes identified during the thematic content analysis (Table 3 in section 5.2). Some of the sections are about themes related to the cleaning steps, but not to a single step, which is why they are discussed in this chapter. Chapter 8 discusses themes that are directly related to one of the protocol steps.

7.2.1 Too much effort

A relatively common theme in the virtual visits is the expression by end-users that the protocol is too much effort to perform. This is often not linked to one particular step but to the protocol as a whole. 6 out of 17 participants lost interest and the will to act at some point during the virtual visit. Statements like "You ask a lot from me.", "If I do that, I have to change all the passwords again. That would mean taking al devices of the wall and back again…" or "Then I would have to call the ICT guy again, which is pretty costly.", showed the irritation amongst the end-users, even though the virtual visit often took no longer than 10 to 15 minutes of their time. Redmiles (2019) and Redmiles et al. (2016) have mentioned this possible behaviour due to the lack of knowledge by end-users. The tasks they had to complete could have been too overwhelming, which is why some end-users did not want to execute the steps.

In 3 out of these 6 cases, the participant actually used the argument that they did not notice anything due to the infection as a reason not to act. Also 3 out of 6 participants ended up choosing not to perform the steps fully, but skip some steps, or perform the steps in a way that took less effort, for example by resetting the router instead of restoring the factory settings.

Even though a large part of the population was working from home during the experiment due to the Covid-19 countermeasures, these participants stated they were fine with the risk of temporarily losing their internet connection as a result of KPNs quarantine system. Notably, none of these participants were seen on the infection feed again during the experimentation period.

7.2.2 Anonymity as a problem

Related to the trust issue, which is discussed in section 7.2.1, is the expressed issue of participants of being called by an anonymous telephone number. 5 out of 17 participants said they normally would not even answer such a call, but due to the conditions of working at home a lot (section 3.4), they did answer the phone. One participant actually answered the call as he was expecting a call from KPN as it was stated in the email that we would give a call, but also stated that they disliked the anonymity.

2 end-users did not answer any of the 3 consecutive phone calls, along with another 2 during the pilot phase, which could have been the result of the anonymity of the call. According to the voicemail messages that these participants had, we did find the right telephone numbers, which means the end-users either chose not to answer the calls, or did not notice the calls. Again, this issue is related to the trustworthiness of the call. Although the email included the name of the person who would call the next day, this anonymous call counters that personal vibe, which decreases trust (Li et al., 2016).

7.2.3 Regular protection is enough

As laptops and smartphones have been around longer than IoT devices, and customers are willing to pay significantly more for the security on these devices, countermeasures against viruses on these devices are generally known. This fame does induce a bias amongst computer users, which creates an illusion that the countermeasures for laptops, smartphones and similar devices is successful for defending against any type of threat. Zeng et al. (2017) have shown this bias which suggests that end-users believe that protection against regular computer viruses can surely deal with Mirai and other Internet of Things related malware as well.

This bias can also be seen in the think aloud protocol that is used in this research. 2 out of 17 end-users have mentioned some sense of disbelieve about their infected network, as they have multiple protection measures in place. One end-user stated: “I have put anti-virus software on 5 of our devices, including my wife’s iPad”. Another end-user said: “I always use a VPN on my laptops and phones, how can something like this enter then?”.

In both cases, the disbelieve also caused an irritation and lack of motivation to perform the corresponding cleaning steps: “I don’t buy it, this has never happened to me and I have wonderful protection measures. My network cannot have an infection so why would I need to clean it?”. Both end-users did perform the necessary cleaning steps, after they were informed further about the way Mirai works and what devices are prone to become infected.

Although the email notification clearly states what devices are generally vulnerable to Mirai, end-users tend to believe their internet connection is completely safe due to countermeasures they take on the relatively accepted and grown devices on the market. Although these countermeasures are useful against many types of malware, the Internet of Things “lives” next to these secure devices and its weaknesses are independent of their secure neighbours. It remains a barrier for end-users to understand the difference between regular internet connected devices and Internet of Things devices, which causes possible delay or stalemates in the cleaning of Mirai infections.

7.2.4 Lacking support by brand

As has been stated in section 2.4, Blythe et al. (2019) and Gibson et al. (2017), researched to what extent the new generation of internet connected devices, IoT devices, is supported by the brands of these devices. The research dug up the reluctance of these manufacturers to put effort into support pages and/or manuals, as the devices are created to be as simple as possible. Similar to the elaborate security measures that IoT devices could have had, manufacturers and vendors do not have an incentive to invest in these properties of their products.

This lack of support has been noticed by several end-users during the experimentation period. 4 out of 17 end-users mentioned that they tried to find help in some way from the manufacturer or vendor of the device, but in all 4 cases, this did not succeed. A manual or support page seems to be a logical place for end-users to look for help in performing the cleaning steps. As is also mentioned in the email notification, these support pages and manuals could explain clearly how to perform the requested steps. If this were the case, end-users would have had a linear process through the steps, leading to successful remediation. However, as the end-users stated: “there is nothing useful in this manual, only things to keep them from being sued.” And “this website is of no help, I already know I that the device is a camera, is there anything useful here?”, the lack of help can be irritating and a significant barrier in the remediation process.

8. Performing the right steps

In this chapter the cleaning protocol for Mirai infected Internet of Things devices is put to the test, as it describes the performance of end-users in each of the steps. This chapter answers the third sub-question: *(SQ3) To what extent are end-users able to perform the required actions to remediate their Mirai infected device(s)?* Sections 8.1 to 8.5 each discuss one of the protocol steps. For this, the labels created through the TCA are used (Table 3 in section 5.2). Section 8.6 summarized the overall performance by the end-users and the chapter closes with section 8.7, in which end-users behaviour is related to existing literature and our lessons learned are described.

8.1 Finding the cause

As detections of Mirai infections only appear at the scale of a network, in shape of the IP address of a router, the first step in remediating Mirai from that network is to discover what device within the network is causing the detection to take place. The email notification (appendix E & F) states that the cause of the infection can often be found through IP cameras, DVRs or printers that are connected to the internet, but this is only a first push in the right direction.

Through the virtual visits we discovered that end-users tend to have one or more out of three ways of reasoning to decide what device is probably infected. 16 out of 17 participants were able to find a device that potentially got infected, which contradicts Redmiles (2019), who stated that end-users are often not able to find the cause of infection. The remaining end-user seemed to try to perform all three of the ways of reasoning (explained below) but none of those was successful. None of the 17 participants thought of the possibility that they had multiple infected devices.

Device malfunctions

The first way of reasoning that end-users use to discover which device is infected, is to look at, what they think are side effects of the Mirai infection. 8 out of 17 end-users stated that one of their IoT devices was malfunctioning to some extent, which is why that device "must be the cause". One end-user stated: "The display of the camera was not working properly, so I was already worried that some hacker took control of it".

3 out of the 8 consumers that used this way of reasoning were actually surprised that Mirai could cause the malfunctioning, as they thought their device just was not working properly as it was "a cheap thing from China" or "There is not even a brand or name on it". Without the virtual visit, it is unlikely that these end-users would have acted. Huijts et al. (2019) already stated this thinking process of end-users. As the vulnerable devices are often cheaper, there exists a relation between the devices that do not work properly and the devices that tend to become infected with Mirai. However, this correlation is not always interpreted as a causation.

Time related reasoning

The second strategy of pinpointing the infected device is to make links between actions the end-user took and the timeliness of the email notification/virtual visit. 8 out of 17 end-users went through their actions with regards to their IoT devices to figure out why the infection turned up at the time it did. One end-user stated "I actually connected it the day before the notification arrived", which was reason to believe the infection must have come from that device. Although the specific action that end-users took right before the notification differ, the way of reasoning is overlapping in each case. One end-user just shifted from Telfort to KPN, which is why they reinstalled their FritzBox and two end-users just installed a new device which the identified as infected. In any case, timeliness is used widely.

Process of elimination

The third and last strategy to identify the infected device in the network, is to use the so-called process of elimination. When end-users make use of this strategy, they tend to take stock of all their IoT devices and start reasoning why N – 1 out of the N devices cannot be the cause of the infection. This leaves them with a single device that is then stated to be the most likely cause. If end-users can only think of 1 IoT device they own, which happened in 7 out of 17 cases, this process is easily performed. Again, especially with this strategy, there is a chance that the end-user overlooks one or more of their devices that could very well be the cause of the infection.

Sometimes end-users turn to the email notification to aid them in finding the infected device, or rather to aid them in excluding their other devices. 8 out of 17 end-users remembered or read on the spot that Mirai often targets IP cameras, DVRs, and printers, which was reason to exclude all other devices from being potentially infected. On the one hand this push in the right direction can thus be helpful to save time in looking into regular internet connected devices, but on the other hand end-users might develop a tunnel vision because of this information. Although the email states that Mirai **most likely** infected one of the three generally vulnerable devices, end-users tend to ignore the bolded part of the message.

An example can be found in the infection of routers. Even though literature has pointed out that routers are relatively high on the list of infected devices (Antonakakis et al., 2017; Çetin et al., 2019b), the email does not specifically mention these devices, which can cause end-users to overlook that device as a potential cause.

8.2 Changing the password of the device

For the second step of the cleaning protocol, end-users tend to have two issues, which retain them from successfully performing the steps. The first problem that end-users can face is that the identified infected device does not have a password option in their perception, which makes it impossible to change it. This was the case for 4 out of 17 end-users.

The second issue that occurs is the lack of help from the brand or manufacturer of the device. 5 out of 17 end-users had the idea to look for help from a device manual or a support page on the internet. In all 5 cases the help source failed to provide the end-users with the needed information, which is in line with the findings in literature on brand support (Blythe et al., 2019; Furnell, 2007; Gibson et al., 2017). The manuals usually do not seem to inform the end-users about anything else than the mechanics of the device, and how to connect the device. There is a dangerous causal loop here, which can lead to end-users distrusting manuals due to their bad experiences. This would then lead to end-users not looking into manuals at a certain point, even if they were to improve.

As has been discussed before, 3 end-users disconnected their device as a solution instead of changing the password and resetting it. Although, technically this does not count as successfully performing this second step of the protocol, the action is successful in removing Mirai from the network, assuming the infected device was removed from the network instead of a clean one. Therefore it is interpreted as a successful action. Only 4 out of 17 end-users were able to perform this step properly, which is relatively low. Together with the 3 end-users that disconnected their device, that comes down to 7 end-users successfully performing the protocol up to the second step. Even if the other end-users were to perform the remaining steps, there is a real danger of their network getting reinfected after some time, depending on the extent to which Mirai is scanning actively across their IPs. Section 8.6 elaborates upon these reinfections.

8.3 Resetting the device

In contrast to changing the password of the device, end-users are more successful in resetting their supposedly infected device itself. 13 out of 17 end-users was able to reset their device, of which 11 end-users pulled the plug for 5 to 10 seconds and one end-user made use of the reset button of the device. The 3 end-users who decided to disconnect their device completely, technically also performed a reset on their device. The remaining end-user was not able to pinpoint the infected device in the first place, which is why they were not able to reset it either.

1 end-user was not sure to what extent the reset actually cleaned the memory of their device, which would mean that Mirai could still be on their device. The network of this end-user was remediated the day after the virtual visit took place.

Again, it is important to note that even though most end-users were able to reset their device, Mirai could come back shortly after reconnecting, if the same standard password remains on the device. Remediation would be successful, but the desired outcome, namely permanent protection against Mirai, would not be reached.

8.4 Resetting the router

After resetting the infected device, users are asked to also reset their router, as there is a possibility that the device did not allow for the previous steps to be executed properly. As we have discussed in section 8.2, multiple users were not able to change the password of the device, for different reasons. In this case, reinfection is likely to happen, as the credentials are still easily invadable.

To make sure the network is closed off from access from the wider internet, a factory reset is needed. Otherwise, the options for open ports, UPnP and DMZ will stay and the devices behind these accessible doors will still be vulnerable. However, as this step comes directly after the reset of the infected device, end-users often interpret this step as only resetting the router by pulling the plug for a while or by pressing the reset button. Even though the step states to make sure that the router returns to its factory settings, and the corresponding link also explains how to reset the router to its factory settings, 9 out of 17 participants only performed a regular reset, which is relatively less effective for the goal of remediating Mirai, as the virus could easily return.

3 out of these 9 end-users actually understood that their actions would not lead to the desired situation, but indicated that it took too much effort to connect each of their devices to the router again. One end-user stated: “But then I would have to arrange all those port forwards again, I do not want to do that”. Even after we informed these end-users of the risks and possible consequences (temporary internet quarantine), they decided to take the risk.

As has been stated before, the email notification includes a URL to the manual for resetting the router to factory settings. However, only 2 participants made use of this link to help them in performing the step. Both were successful in doing so, next to another 4 end-users who were able to reset their router to factory settings.

8.5 Changing the password of the router

As with the steps regarding the infected device, there is a noticeable difference in performance of end-users in performing the two steps. Although only 6 out of 17 participants were able to reset their router to the factory settings, 11 succeeded in changing the password of the router.

Even more surprisingly, 6 out of these 11 also made use of the URL in the email notification to do so, which means 4 end-users decided to look into the notification only after resetting their router. A

possible explanation for this behaviour is that these end-users thought to be capable of resetting the router without any help, but did not feel the same capability for resetting its password. All 6 end-users were successful in changing the password of the router, but the 4 end-users who only consulted the notification at this point did not perform a full factory setting reset before this.

Two end-users stated they had issues with the software that is used to reset the password for the Experia Box V10, which is the standard router for KPN and Telfort customers. After several trials they did not manage to change the password of their router, as the software did not load the page where this action can be performed. Both end-users decided to give up.

8.6 Overall performance

To get an overview of the overall performance of end-uses considering the cleaning steps, this section summarizes the actions across all steps and elaborates upon the remediation rate of the experiment. Figure 26 shows the relative number of participants that were able to successfully execute the cleaning step, for all five steps. As can be seen, step 4, to reset the router to factory settings has proven to be the hardest for end-users, or took too much effort. Closely behind is step 2, to reset the password of the infected device. Interestingly, only 1 end-user was not able to identify the infected device in their network. Because the only dependence in the protocol is that the second and third step can only be executed if the first step has been performed successfully, this finding gives hope for a successful remediation in many cases.

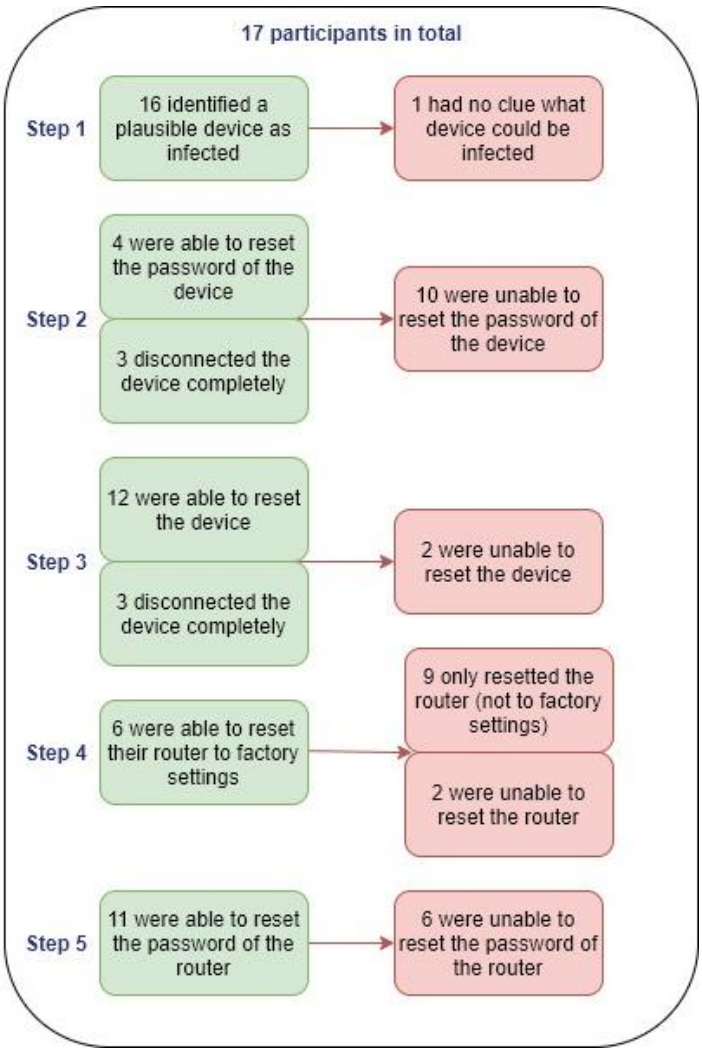


Figure 26. Performance on the cleaning steps.

Note that the failing end-users for one step do not have to be the same end-users that fail to perform another step. In total, only 6 end-users were able to perform each of the steps successfully. These are the 6 end-users that were able to reset the router to factory settings, which seems to be a critical point in the protocol. In cases where the identified infected device was actually the FritzBox router (4 out of 17), the end-users steps are included in both the steps for the device and the router. Also disconnecting the device is counted as successful for the corresponding steps. This means that if the infected device is a router, which the end-users decides to disconnect, all steps are interpreted as being executed with success.

As we tried to help the end-users after the call to help them perform the steps that they could not perform correctly by themselves, we expected the remediation rate of the participants' networks to be 100%. However, often we were not able to be of any help, as the participant either supposedly performed all steps correctly or their perception of their network made it impossible for us to aid. 3 end-users were still infected two days after the virtual visit, which means the remediation was not successful in these cases.

Figure 27 shows how often the IP addresses in the of the participants of the experiment were seen in the detection feed of Mirai infection during the experimentation period. We excluded weekend days from this graph to remove the weekend only detections from the data, as they were not part of the experiment. This means the maximum number of appearances would be 35. It also means, however, that there is a small underestimation of the number of appearances of IPs, as they could have still been infected during the weekend after they have been called. Manual inspection showed that this was only the case for IP addresses that were also still infected after the weekend, which means the remediation for these IPs was not successful. In a best-case scenario, an IP address should appear not more often than 3 times on the detection feed. IPs that showed up 3 times or less are interpreted as remediated successfully (one day before the email, the day the email is forwarded and the day of the virtual visit). A last note that probably caused an overestimation of the appearances of IP addresses, is the shortened time in which customers were quarantined due to the Covid-19 countermeasures. KPN policy changed during this period to a system that manually removed customers shortly after their network was quarantined, which means the incentive of getting back ones internet was missing.

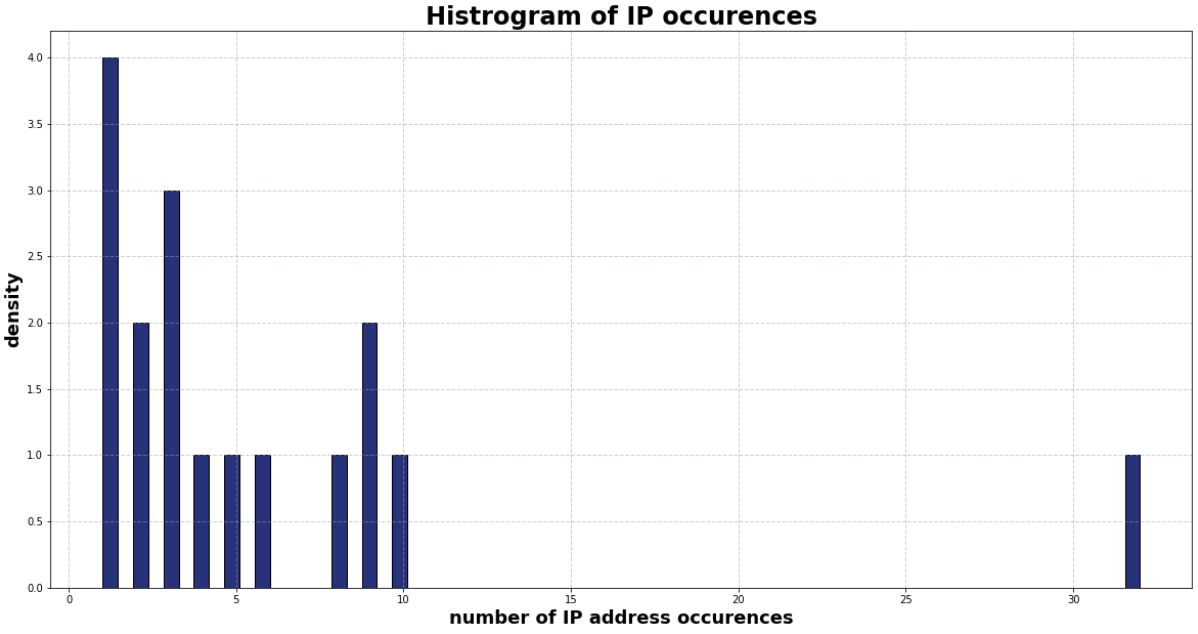


Figure 27. The number of detections during the experimentation period per unique IP address.

In one of the cases, the participant was not able to change the password of the FritzBox or reset the FritzBox to factory settings. After 5 days, their network was still remediated, which was in line with the proposed action by the end-user to ask for help from another family member. Another case of failed remediation was due to an IP camera that supposedly did not have a password setting. The end-user was able to reset the device, which probably got reinfected shortly after. In the last case of failed remediation, the end-user had no clue which device could be infected. As this customer was a Telfort customer, KPN systems were not able to take additional measures temporarily. Therefore this customer was detected as infected for over a month (32 workdays in a row).

Reinfections of networks

Next to the 3 participants whose network was not remediated successfully, another 5 participants' network was detected more than 3 times during the experimentation period. As these detections were not consecutive, this is an indication of reinfection of the network. In cases of reinfection, the experiment protocol states that the customer is dealt with in the regular way, namely the usage of temporary internet quarantine. The finding that 10 participants were not able to change the password of their device is a straightforward explanation for these reinfections/failed remediations. All participants who got reinfected were not able to change their password during the virtual visit. Of the 5 detected reinfections, 1 participant got reinfected twice, as there were 3 separate periods in which their network IP showed up on the radar.

8.7 The thinking process behind the performances

Although particular decisions and actions of end-users during the clean-up efforts have been described and related to literature, it is important to conclude on the virtual visit as a whole. This section discusses to what extent users performances can be related to literature and what we have learned about end-user behaviour in how they engage the problem at hand.

In general, the process of performing the remediation steps does not run smoothly for end-users of Mirai infected IoT devices. The struggles that end-users faced during their performances can largely be explained by literature on IoT devices, notification mechanisms and end-users mental models.

Mohamed et al. (2017) and Redmiles et al. (2016) have warned for the importance of a balance between usability of a notification and the level of security that it ensures when executed successfully. Although we have tried to create an email notification that is usable for most end-users and at the same time upholds the desirable level of security, some end-users were overwhelmed by the tasks they had to perform and ended up in not performing one or more of these tasks. This shows the importance of a clear and easily understood protocol.

Moreover, there exists a difference between mental models for security issues by experts and non-experts. Asgharpour et al. (2007) indicated a large difference in how notification are perceived by the two distinct groups of people, which becomes apparent in this research as the notification and remediation protocol have been designed by experts, but non-experts should execute the requested steps. End-users can have a different mental model than expected of what a Mirai infection entails. This became visible in several virtual visits, in which end-users were sure their regular protection measures should suffice against Mirai (Bravo-Lillo, 2011; Zeng et al., 2017) or where end-users did not know that Mirai could have an impact on their network or devices, which can be found in van Eeten & Bauer (2008) and Huijts et al. (2019) as well.

If the issues above are overcome, the IoT environment causes more, due to the heterogeneity of devices (Zimmermann et al., 2019; Forget et al., 2019) and a lack of interface (Kolias, 2017; Anthi et al., 2018) or aid by the brand and vendor (Blythe et al., 2019; Furnell, 2007; Gibson et al., 2017).

9. Barriers & improvements

This chapter gives an answer to the fourth and last sub question: *(SQ4) What are possible improvements for the remediation process of Mirai infected Internet of Things devices?* Although many issues have been discussed up to this point, it is important to put the problems that the remediation process of Mirai infected Internet of Things devices entails into perspective. This chapter looks at the overall remediation process, by discussing the issues that were found during this research and relating them to possible improvements and corresponding actors. Each of the 5 following section in this paragraph discusses the barriers that could occur for a specific part of the anti-botnet cycle (Online Trust Alliance, 2013).

Although the anti-botnet cycle (Figure 3) might seem straightforward, many things can go wrong in each stage of the cycle. Figure 28 shows the anti-botnet cycle (Online Trust Alliance, 2013), including possible barriers that can cause the step to fail or solutions that can remove certain barriers. For each barrier/solution a potentially responsible actor has been linked, that has the means to lower the barrier or remove it altogether. Statements that start with a + represent possible improvements and statements that start with a – represent present issues. Moreover, blue statements are direct findings of this research and green statements are statements that were derived from the findings in this research. The following sections elaborate upon each of the identified barriers and solutions.

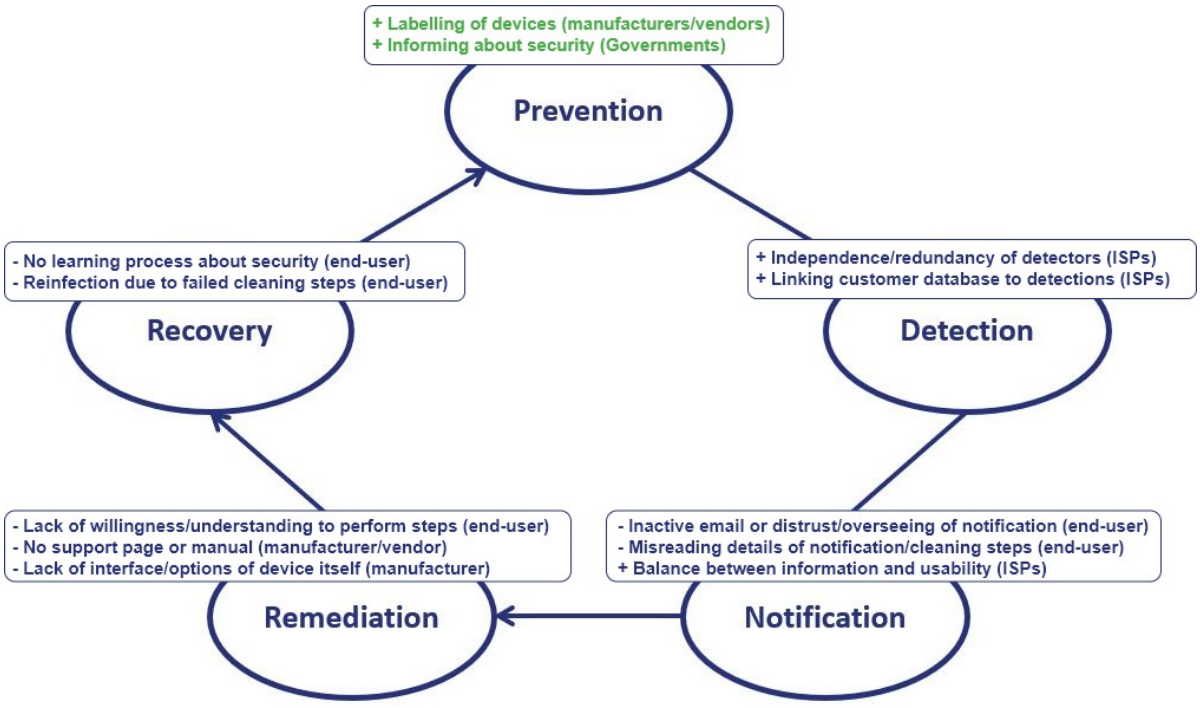


Figure 28. Anti-botnet cycle including barriers and possible solutions.

9.1 Prevention

In chapter 2, literature has pointed out that manufacturers and vendors do not have the necessary incentives to aid in the prevention of Mirai infections taking place. If anything, these actors only encourage consumers to purchase vulnerable devices, due to their low price. Because consumers also do not notice the consequences of Mirai in most cases, this attractiveness of vulnerable devices increases even more. Despite the effort of Government campaigns (Ministerie van Economische Zaken, 2019; Autoriteit Persoonsgegevens, 2019), Mirai remains an existing threat, due to the vulnerable devices on the market. ISPs could aid in these campaigns, but no more than that.

Vendors could play a role in a movement towards an IoT market that has more secure devices, which is also mentioned widely in literature (Kolias, 2018; Abu Waraga et al., 2020; Kambourakis et al., 2017; Mohsin et al., 2017). Johnson et al. (2020) pointed out that labelling the devices based on their level of security can nudge consumers into purchasing better secured products. Blythe & Johnson (2018) add to this that consumers do care about their security, as long as they understand the extent to which it is present.

9.2 Detection

If Mirai has infected a certain device, it is important to inform the right actors about this infection. As consumers often do not know about their infection themselves, another party has to recognize its existence. KPN relies on Shadowserver to detect any Mirai infection of consumers on their network, which is relatively easy due to the obvious signature that Mirai uses when active.

However, the communication between the detector and the ISP can fail from time to time, which means it is not possible to inform the victims timely about the infection, as these victims are not known by the actor that can notify. As Mirai goes hand in hand with open ports, and vulnerable devices, ISPs might be able to setup their own tool to detect Mirai infections on their network, for example by making use of Darknet. This would decrease the number of dependencies within the cycle. In the experiment, we were prepared for possible failure of informing by the detectors, which is why we made use of 2 independent ways of detecting. During the experiment, Shadowserver failed to inform KPN on 4 different days. The Darknet failed to deliver on 3 days, directly after the experimentation period had ended. Redundancy could be key in making sure that the detections are continuously working properly.

In many organisations, activity drops during the weekends, which is no different at KPN. This research has shown that many Mirai detections only occur during the weekends, in which no notifications can be send. Although some Mirai infections last until after the weekends, which means the owners of the infected devices can still be informed, the infections generally last longer because of the inactivity during the weekends.

An automatic notification tool could be linked to the detection mechanism, to be able to also inform end-users about their infection during the weekends. However, as KPN usually makes use of quarantine methods for Mirai infected consumers, this might cause problems during the weekends, as there would be no point of help for consumers to turn to. Especially in times were connectivity is key, this is undesirable. It should not be overseen that ISPs generally want their networks to be clean, but they also want their customers to be satisfied with their services. In cases of Mirai infections, these two objectives are often 180 degrees across from each other.

To be able to inform the right end-users, ISPs must link the detected IP address to a customer, which can then be informed. This means KPN databases have to be up to date and available to the employee that processes the detections. During the experimentation period, 3 end-users could not be informed timely due to a lack of availability of this data, as there was a significant risk that the systems would link a different customer to the detected IP address. Although this issue is not common, it should be seen as a barrier in the Anti-botnet cycle.

9.3 Notification

After an infection has been detected, and the systems are able to link an end-user to the infection, the notification part is key to activate the end-user, as again, the user often does not notice the infection without any intervention (Van Eeten & Bauer 2008). In this research, only email notification were used, without the usage of quarantining end-users, which means the email should be clear and trustworthy to make sure end-users are willing to act upon it. Section 7.2.1 already explained the need for trust, as there is no way to prove the authenticity of the notification by end-users other than verifying with other sections within KPN. Moreover, this research has shown the importance of the personal vibe of the notification also found in Redmiles et al. (2016) and Li et al. (2016a), as end-users often had issues with the anonymity of the calls. Common issues that can arise during the notification phase of the cycle are that the automatic emailing tool that is needed for layout can malfunction, which means no notifications can be send.

The other possible barriers can be found at the end-user side. The email notification could be overseen by end-user, because they do not actively use their KPN email or the notification could be perceived as spam and therefore get deleted. Moreover, as we have found during the pilot calls, a non-existing email can be used for the KPN account, which means there is no way to contact the end-user.

Lastly, end-users might not read the email carefully and perform the requested steps in a slightly different way. Section 8.4 has shown an example of this that happened in this research. Namely, 9 end-users only performed a “soft” reset on their router instead of the factory settings reset that is mentioned in the notification.

Altena (2018) and Verstegen (2019) have found a significant improvement of the remediation rate when internet quarantine is used additional to the email notification, which could be explained by the barriers for the effectiveness of the email notification stated above.

Considering the content of the notification, findings in this research are contradictory. For some end-users, the email notification seemed to have overwhelmed them by the many and difficult steps that it requests. On the other hand, multiple end-users would have like additional information of the issue of Mirai, botnets and DDoS as the knowledge of the issues would improve their motivation to act. This can also be seen in Reder et al. (2012).

9.4 Remediation

The is a big set of barriers that can cause the cycle to fail during the remediation phase. Chapter 8 explains these possible issues in detail, which is why we will only give a smaller overview in this section. The first issue that can arise when end-users should perform the cleaning steps is their willingness to do so. This research has found a lack of this motivation in multiple cases, even in our presence. It can be too much effort for end-users, which can logically be explained by the finding that they do not feel the consequences of their infection. This means their benefit is perceived as nihil, which means any effort is too costly.

If end-users are willing to act, the cleaning steps can still suffer from the end-users lack of understanding. In most cases the end-user is able to pinpoint the infected device, but does not known how to change its password, either due to lacking support by the brand and vendor, or by the lack of interface and options of the devices themselves. Even with our presence and help, only 6 end-users were able to perform all the steps correctly.

9.5 Recovery

After an end-user has gone through the complete anti-botnet cycle, they should be remediated fully and have a better understanding of the dangers of poorly secured Internet of Things devices. However, even in this stage, barriers exist that can cause end-users to become infected again.

If the cleaning steps were not performed perfectly, there is a real chance of reinfection, which happened in 5 out of 17 cases in this research. This finding indicates that end-users might have purchased a new vulnerable device or that the cleaning steps were not executed properly, which caused the weakness in their network to continue existing. Even with our help, in some cases it turned out to be close to impossible to successfully perform the steps, for example because there was no lead to a potentially infected device or because the identified device had no leads to change its password. Lastly, in some cases where we were able to find how to perform the steps, end-users were not able to follow our suggestions.

Informing by trusted parties, such as governments could be key in making sure that recovered but also new potential targets of Mirai understand the threats when they purchase poorly secured IoT devices. This should be relatively successful, as end-users are willing to invest in security, as long as they understand its importance and effects (Blythe et al., 2020; Nguyen et al., 2017; Rowe & Wood, 2013).

10. Conclusion and discussion

The main objective of this research was to get a better understanding of the processes that take place during the remediation of Mirai infected Internet of Things devices, especially at end-users homes. Aiding in the process to answer the main research question, were 4 sub-questions which have been answered in the previous chapters. Section 10.1 recalls and displays the main findings of this research and section 10.2 discusses what these main findings imply for KPN, ISPs in general and policy for the remediation process of Mirai.

10.1 Main research findings

To be able to give and answer to the main research question: *What do we learn about how and to what extent Internet Service Providers can improve the remediation process of malware infected Internet of Things devices by monitoring end-users while they are cleaning their Mirai infections?*, we followed Mirai infections of KPN customers for 10 weeks in total. In total, 59 unique networks have been detected by the Mirai scanners, of which 24 took part in a virtual visit after being notified of their infection through a notification email. During these virtual visits, a think aloud protocol allowed for the end-users to explain their actions in detail while they were performing the cleaning steps.

The first week of monitoring has been used to perform pilot virtual visits, from which we got experience in performing these calls, but we were also able to tweak the protocol that was used during the calls. 7 out of 24 calls took place during this pilot phase.

During the 7 weeks after the pilot phase we executed the actual experiment, in which 17 end-users were visited virtually. The two last weeks of the experiment, only the detected networks were monitored to measure the remediation and reinfection rate of the participants of the experiment. We performed a thematic content analysis on the complete transcribed virtual visits to pinpoint common issues and themes in the remediation process of Mirai infected IoT devices.

We have seen 16 out of 17 male consumers, which is in line with previous research. However, as we have only spoken with the person responsible for computer security tasks in the household, this does not necessarily mean that men get infected with Mirai more often.

The household size supports the idea that male end-users are not infected more often, but deal with the issue more often. The age of the end-users that perform the clean-up steps looks to be normally distributed between a minimum of 21 years and a maximum of 80 years, but there is no statistical prove that the distribution is really normally distributed.

7 out of 17 end-users were only able to identify 1 Internet of Things device in their network, followed by 6 end-users who only identified 2 and 3 end-users that identified 3 devices. Although this often led to an easy time in pinpointing the infected device, as excluding the other devices these end-users owned was relatively straightforward, the end-users could have forgotten about one or more IoT devices during their stocktaking. As we hint in the email notification at IP cameras, printers and DVRs, end-users might induce a bias based on this hint and overlook other potentially infected devices.

However, in 16 out of 17 cases the end-user potentially pinpointed the correct device. Using the screenshot tool that labels Mirai infected networks, we were able to identify 12 out of 17 infected devices, which was the same (type of) device that the end-user identified in each of the 12 cases. 8 IP cameras showed up, 4 FritzBox routers, 1 NAS, 1 media player, 1 TV and 1 unknown device, as both the screenshot tool and the end-user were not able to identify a potentially infected device.

The most notable device is the FritzBox router, which has not been seen on the radar in any previous research and is known to have a relatively high level of security build into it. Because this device was identified in 5 cases (1 during the pilots), and remediation was successful in each of these 5 cases, it is not likely that another device was actually infected. An explanation could be that attackers have implemented a novel way to infiltrate the secure FritzBox through some weakness.

During the experiment, we found out that the anti-botnet cycle can be disturbed in many different ways, mostly during the remediation efforts. 12 out of 59 IP addresses only showed up as infected during the weekends, which means there was no way to inform the end-user about their Mirai infection, as the Abuse desk is closed during the weekends. These detections are either devices that are only turned on during the weekends, new purchases that are secured shortly after they're installed or an attacker strategy to activate their bots during the weekends, as they have knowledge of the absence of notifications during the weekends.

Also the detection systems and customer databases have to be up and running for email notification to be sent. Within the experimentation period, 3 end-users could not be informed timely due to the failing systems. There were 3 days within the monitoring period where no detection feed arrived, which also holds the remediation process from starting.

6 out of 17 users that were notified and visited virtually mentioned that performing the steps was too much effort. Sometimes this resulted into the end-user only performing some of the steps or performing the steps is a poorer way. Another 3 end-users mentioned that they did not fully trust the notification and virtual visit. Although the email notification has been optimized for understanding and transparency, end-users are often cautious about phone calls. The anonymity of the phone calls strengthened this lack of trust, as 5 end-users mentions they preferred the call not to be anonymous. This anonymity could have very well been the reason why 3 end-users did not answer the 3 consecutive phone calls.

Another barrier that 4 end-users came across, is the lack of support that is given by the manufacturers/vendors. Their identified infected IoT device either did not have a manual or support page, or the needed information was lacking from these sources.

If end-users are to clean Mirai from their network, they must first acknowledge its existence and the steps that are needed to remove the virus from their infected device and network. 2 end-users refused to believe Mirai could have infiltrated their one of their devices as they had anti-virus software and VPN installed on all of their laptops and smartphones. As the Internet of Things is still on the rise, some end-users do not know about its functionalities and weaknesses yet, leading to misunderstanding of the issue and its solution.

Regarding the 5 steps of the cleaning protocol, only 6 out of 17 end-users completed all steps successfully. This typically goes wrong at the second step, which is to change the password of the device, or at the fourth step, which is to reset the router to its factory settings. 16 end-users were able to pinpoint a plausible infected device, using one or more out of three strategies. 7 end-users decided on their infected device by eliminating all other connected devices in their network. 8 end-users stated that one of their devices was malfunctioning in one way or another, which was cause to believe that device had to be infected and another 8 end-users either bought, installed or performed some action regarding one of their devices around the same time the notification email was sent, which was reason to believe their actions caused the infection.

Only 4 end-users were able to change the password of their identified device and 3 end-users decided to disconnect their device as they took the infection very seriously and were not able to perform the actions in another way. This leaves 9 end-users that were not able to reset the password of their device, either due to the indicated lack of a password option or the lack of support from a manual or support page. Resetting the device caused no issues, except for one end-user who was not sure about the extent to which the memory and thus Mirai got cleaned when a reset takes place. The person who was not able to identify any infected device logically did not perform this step either.

For the router, 9 end-users only performed a regular reset, instead of resetting the router to its factory settings, even though it is clearly stated in the notification email. However, as only 6 end-users made use of the email notification while performing the steps, for example by looking into the added links, this low success rate can be explained. In cases where end-users are not able to permanently remove Mirai, the steps regarding the router are most important. If these end-users fail to reset their router to factory settings, the network probably still allows for the larger internet to invade the network.

The daily infection feed showed the effects of the sometimes poorly executed cleaning steps, as 3 end-users were still infected 2 days after the virtual visit took place and 5 end-users got reinfected during the monitoring period. This also shows a lack of a learning process for end-users in purchasing secure products and using their IoT devices in a secure way.

10.2 Implications and recommendations

This research has shown that issues and barriers can be found in each step of the anti-botnet cycle, which means the remediation of Mirai infected Internet of Things devices is far from perfect. From monitoring Mirai infected end-users. we have learned that even in the most perfect way of notifying and informing end-users, the influence of Internet Service providers remains limited. Improvements to the botnet cycle should be achieved by multiple actors.

Literature pointed out that manufacturers and vendors could actually protect end-users from investing in poorly secured devices by a simple labelling technique. Trusted parties like governments could strengthen end-users knowledge about the risks and issues with these devices by general informing campaigns. This would improve the prevention of Mirai infections occurring in the first place. Part of this awareness will probably come through time, as Internet of Things should still be seen as a new kid on the block. However, as online availability become ever more important, waiting for this awareness to grow could be too passive of a strategy.

For the awareness to grow further, a list of known vulnerable devices could aid both ISPs in informing end-users and end-users themselves in not purchasing these devices or pinpointing the infected device after they received a notification. Although KPN hints at generally known vulnerable devices, the specific types and names of the vulnerable devices could be the next step. The screenshot tool used in this research could help in labelling the infected devices, and contact with end-users could add to the list as well. Thirdly, future research might be able to check the vulnerability of potentially weak devices by manually letting Mirai try to penetrate these devices. This could also improve the knowledge of how different types of Mirai malware work and to create a one-size fits all protocol for cleaning Mirai infected IoT devices. A good start would be to look into the FritzBox router which has turned up only in this research and does not fit the profile of a regular victim of known Mirai variants.

On their turn, ISPs, such as KPN could invest to become less dependent on other parties and multiple databases within their systems. As KPN relies on only Shadowserver for informing them about the daily infections, there is a critical dependency at the beginning of the remediation process. Redundancy or independency could both remove this thin line.

Regarding recommendations on the cleaning protocol and the way in which end-users are approached, the requested steps are often not performed with success. Because this research took away the temporary quarantine as an incentive for end-users to act, the trust in the authenticity of the experiment went down and the willingness to put effort into the cleaning steps likewise.

Unfortunately, the many experiences with phishing and such has turned end-users into cautious actors who prefer not to act, which means a display of power might be needed to gain trust of the end-users. We highly recommend removing the anonymity from contacting customers, as this could improve trust and willingness.

As has been mentioned before, more info about the potentially infected devices could give end-users a hand in performing the steps, although most end-users were successful in doing so with the current state of the notification. In special, uncommon devices such as the FritzBox could easily be overlooked by end-users as routers are not mentioned in the email notification, even though literature states that routers are one of the big three victims of Mirai. Next to sharing information about potential victimized devices, KPN could focus on making the end-users more dependent on the email notification, to make sure it is thoroughly read by the end-users, before the start to perform the steps.

The regular way of obligating end-users to inform the abuse desk about their specific steps helps with this, but also adds time consuming contact between the two parties. Moreover, this bureaucratic methodology can cause end-users to focus on the goal rather than the process and the way of thinking that goes alongside this process. End-users do not learn from their poor security purchases and the way they use these bought devices. Only in some cases, do end-users take the issue seriously, which often leads to the disposal of the infected device and a striking change in attitude.

In this research, we have found out the common tactics of end-users in performing the steps which could be added to the notification as well. For example, for the first step of pinpointing the infected device, the notification could be enhanced by adding possible tactics to do so, namely statements such as: “Maybe you have just installed a new device, or one of your devices malfunctions…”. Moreover, the notification could become more useful if suggestions are given in the case where end-users fail to perform the step successfully. For example for the step in which end-users are supposed to change the password of their infected device, the notification could explain what end-users should do if there is no password setting, or if the manual/support page does not aid in changing the password. To the extreme, this idea could be connected with the idea of a detailed database on vulnerable devices. For each device, end-users could then look up the necessary steps, which are more specified for the devices than the regular cleaning protocol. It should be noted that this would be an ongoing expensive investment, as more devices enter the market and attackers swiftly adapt their tactics.

This research has shown that without a real incentive, end-users do not perform well on the cleaning steps, but at the same time tried to optimize the way the end-users were informed. This means it is a hard task to improve the protocol as it is and suggests that a more effective change can probably be found at the beginning of the anti-botnet cycle, in preventing Mirai infections from taking place. If end-users are more aware of the Internet of Things, its benefits, and its risks, the IoT could evolve in a similar fashion as regular internet connected devices, such as laptops and smartphones, namely turn into a well secured, easily usable set of devices. At the same time, ISPs should not stop to inform current infected end-users and incentivize them to act upon their infection.

11. Limitations, validity and future work

This final chapter reflects upon the research as a whole and identifies knowledge gaps that are useful to be filled to add more value to this and previous research. Section 11.1 discusses the limitation of this research and its findings. Section 11.2 and 11.3 elaborate upon the internal and external validity respectively and section 11.4 ends the chapter with a discussion on the next steps after this research.

11.1 Limitations

Although many parts of this research have been thought through, it is important to highlight the limitations to put the findings into perspective. Some of these limitations have already been discussed in section 4.6, which will not be discussed again here.

The first major limitation of this research is the presence of many levels of interpretation between the experiment and the conclusion we state. The conclusions in this research are mainly on the reactions and thus thought processes of end-users during the remediation process. First of all, we asked the end-users to put these thoughts into words, which adds abstraction to the ground truth. We interpreted their statements and processed them to what we believe these end-users meant by their statements, which on our turn, we put into words again. It should not be ignored that many biases exist in end-users while telling their story, but also in us while listening to those stories. We tried to minimize these in-between layers of interpretation by choosing for a simultaneous think aloud protocol in which the notification and our roles made sure that end-users felt comfortable sharing any thought, but the biases still exist.

Another major limitation of this research is the generalisation of existing Mirai variants. As we only filtered networks based on the Mirai signature, there was no way of looking deeper into what variant of Mirai we were dealing with in specific cases. It could be that the cleaning protocol has different levels of success for different types of the Mirai virus, which means the end-users actions would not be comparable. However, as this research is mostly about end-users reactions to a Mirai infection in general, the type of Mirai that was present would not change our findings significantly.

Due to technical changes, we lost 3 potential participants for the experiment. Right after the experimentation phase we did not receive a detection feed for 3 days straight, which kept us from monitoring end-users that were virtually visited at the end of the experimentation period. Although these issues are small and the methodology allowed for such flaws to occur, it should be kept in mind that the anti-botnet cycle can have problems in each of its stages.

The existence of natural remediation as mentioned in many previous research attempts has not been explained or accounted for in this research. The parts in this research about the remediation success rate are likely biased because of the phenomenon of natural remediation. Even though we were present while end-users performed the steps, end-users could have performed additional steps before or after the call, or other household members could have taken actions without knowing their consequences.

A large part of increasing security for internet connected devices can be found in releasing firmware updates for known vulnerabilities. However, the devices do require an architecture that allows for overwriting the core coding and a rollback if the update caused unforeseen consequences (Antonakakis et al., 2017). In cases of less evolved IoT devices, such updates can be executed manually by end-users. However, this possible step to secure a network has not been considered in this research and is also not part of the base set of instructions that KPN sends out to Mirai infected end-users. This limits the conclusions that can be taken from this research to a set of possible countermeasures against Mirai.

A last limitation of this research is the number of participants in the experiment. Although we performed more than enough virtual visits to be able to create conclusion about the think aloud protocol and the thematic content analysis according to the saturation principle, the statistical parts about end-user demographics and their infected devices could have been stronger if the number of participants had been higher.

11.2 Internal validity

As we constantly monitored two sources of Mirai infections, we assume that all KPN and Telfort customers have been considered for this research. Only in a later stage of the experiment, we lost parts of this population. Only during the period of technical issues, we lost 3 Mirai infected end-users. Also, there is no way to prove that all Mirai infected KPN and Telfort customers were detected by the sources. Therefore there is no certainty that the complete population has been considered.

Regarding the validity of the gathered data during the virtual visits, some limitations have already been mentioned. Both end-users and we have biases that can influence our interpretation of end-users thoughts. Moreover, consumers could have chosen to be dishonest during the virtual visit or be unknowingly dishonest as their perception of their actions does not fit their real actions. This could have caused the data to be invalid. However, as we have chosen to make use of a simultaneous think aloud protocol, the space for dishonesty shrinks as there is less time to prepare “fake” answers. The time spent to create a comfortable environment for end-users removed the probability of them being dishonest even more.

As the virtual visits often progressed in a slightly different fashion than was prepared in the protocols, the conversations and order of the data gathering is inconsistent over the virtual visits. In some cases the end-user decided to explain more about their experiences with the infected device without being asked to do so and in some cases the end-user mentioned their own age before it was asked. Although the cleaning steps were always in the same order and these differences are minimal, they could have caused different outcomes and levels of detail within the gathered data.

11.3 External validity

As this research has only focused on customers of one Dutch ISP, it is hard to generalize the findings to customers of other ISPs in the Netherlands or let alone the world. As the public opinion on security is a highly cultural issue, the reactions of end-users to a notification about their Mirai infected IoT device likely differ largely across the world.

In terms of demographics, we have included a large range of end-users considering age and household size. This is reason to believe that the findings can indeed be generalised to a broader population. In comparison to previous research, findings largely overlap. The types of devices that became infected are in line, (except for the FritzBox) and most of the themes that we recognized in end-users behaviour have been identified before.

This research is a useful step in understanding end-users behaviour after receiving a notification about an infection on their network. It is too soon to state that this research can be used for end-users reactions to all types of malware, but it can be stated that end-users reactions probably overlap in malware regarding Internet of Things devices and malware that does not per se affect the owner of the device.

11.4 Future work

This research can be broadened in many different ways that could be useful for the remediation process to flourish. First of all, end-users could be visited physically instead of virtually. This will remove a layer of interpretation and will gather ground truths on end-users behaviour and actions. For example, during a physical visit, the router activity could be measured to see all connected devices. These devices can then be compared to the devices end-users state they own to find out to what extent end-users actually have knowledge of what is connected in their network. Moreover, the actions they take as a reaction to a notification can be followed closely and in addition to end-users thoughts, the researcher can observe the end-users during their performance. This would add non-verbal language and emotions to the data. A last benefit of a physical visit is that the network traffic can be measured real time, which means we could be able to check whether the malicious traffic stops due to the end-users' actions.

Regarding the blind spots of this research, future work can look into the weekend only detections of Mirai infections. We have listed several explanations for infections that only show up during the weekends and it could be interesting to exclude x -1 of these explanations. Another finding that needs more research to understand its presence is the multiple infections that took place on a FritzBox router. As has been mentioned before, these devices are not generally vulnerable, which means something else must be going on.

A phenomenon that still has not been explained yet is the existence of natural remediation. Previous research and this research have seen infections disappear without knowledge of the owners performing any critical steps that could remove Mirai. Future research could fill this knowledge gap as there might be hidden keys to a better remediation process.

Lastly, the recommendations in this research should be researched to calculate their added value. One of the suggestions in this research is to keep track in detail what devices seem to be vulnerable and how to tackle the infection for each of these specific devices. Research should analyse how effective this additional source of support can be for end-user performances. Another idea that should be tested is the most effective way of informing end-users of the risks of poorly secured Internet of Things devices. Research has shown that end-users are sensitive to security if they have knowledge on the different levels of security on different devices, but research is missing on the proper method to transfer this knowledge.

References

Abu Waraga, O., Bettayeb, M., Nasir, Q., & Abu Talib, M. (2020). Design and implementation of automated IoT security testbed. *Computers & Security*, *88*, 101648. https://doi.org/10.1016/j.cose.2019.101648

Abuse Information Exchange. (2020). *Home page.* https://www.abuseinformationexchange.nl/

Akhawe, D., & Felt, A. P. (2013). *Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness*.

Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). Internet of Things security: A survey. *Journal of Network and Computer Applications*, *88*, 10–28. https://doi.org/10.1016/j.jnca.2017.04.002

Alcaide, A., Palomar, E., Montero-Castillo, J., & Ribagorda, A. (2013). Anonymous authentication for privacy-preserving IoT target-driven applications. *Computers & Security*, *37*, 111–123. https://doi.org/10.1016/j.cose.2013.05.007

Ali, B., & Awad, A. (2018). Cyber and Physical Security Vulnerability Assessment for IoT-Based Smart Homes. *Sensors*, *18*(3), 817. https://doi.org/10.3390/s18030817

Allen, D. (1999). *Transaction Costs*. Encyclopedia of law and economics.

Altena, L. (2018). *Exploring Effective Notification Mechanisms For Infected IoT Devices*. Delft, University of Technology.

Anderson, R. (2007). *Thematic Content Analysis (TCA)*. 4.

Anthi, E., Ahmad, S., Rana, O., Theodorakopoulos, G., & Burnap, P. (2018). EclipseIoT: A secure and adaptive hub for the Internet of Things. *Computers & Security*, *78*, 477–490. https://doi.org/10.1016/j.cose.2018.07.016

Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., Durumeric, Z., Halderman, J. A., Invernizzi, L., Kallitsis, M., Kumar, D., Lever, C., Ma, Z., Mason, J., Menscher, D., Seaman, C., Sullivan, N., Thomas, K., & Zhou, Y. (2017). *Understanding the Mirai Botnet*. 1093–1110.

Asgharpour, F., Liu, D., & Camp, L. J. (2007). Mental Models of Computer Security Risks. 9.

Ataç, C., & Akleylek, S. (2019). IoT Çağında Güvenlik Tehditleri ve Çözümleri Üzerine Bir Araştırma. *European Journal of Science and Technology*, 36–42. https://doi.org/10.31590/ejosat.494066

Autoriteit Persoonsgegevens. (2019). *Internet of things en smart home? Bescherm uw privacy!* https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/internet-telefoon-tv-en-post/internet-things

Batalla, J. M., Mastorakis, G., Mavromoustakis, C. X., & Pallis, E. (Eds.). (2017). *Beyond the Internet of Things: Everything Interconnected*. Springer International Publishing. https://doi.org/10.1007/978-3-319-50758-3

Bauer, J. M. (2011). *Introduction to the Economics of Cybersecurity*. *81*.

Bauer, J. M., & van Eeten, M. J. G. (2009). Cybersecurity: Stakeholder incentives, externalities, and policy options. *Telecommunications Policy*, *33*(10–11), 706–719. https://doi.org/10.1016/j.telpol.2009.09.001

Biggs, J. (2016). Hackers release source code for a powerful DDoS app called Mirai. *TechCrunch*. https://techcrunch.com/2016/10/10/hackers-release-source-code-for-a-powerful-ddos-app-called-mirai/

Bingham, W. V., & Moore, B. V. (1931). *How to interview.* Harpers.

Blythe, J., & Johnson, S. (2018). The Consumer Security Index for IoT: A protocol for developing an index to improve consumer decision making and to incentivize greater security provision in IoT devices. *Living in the Internet of Things: Cybersecurity of the IoT*, 7. https://doi.org/10.1049/cp.2018.0004

Blythe, J. M., Johnson, S. D., & Manning, M. (2020). What is security worth to consumers? Investigating willingness to pay for secure Internet of Things devices. *Crime Science*, *9*(1), 1. https://doi.org/10.1186/s40163-019-0110-3

Blythe, J. M., Sombatruang, N., & Johnson, S. D. (2019). What security features and crime prevention advice is communicated in consumer IoT device manuals and support pages? *Journal of Cybersecurity*, *5*(1), tyz005. https://doi.org/10.1093/cybsec/tyz005

Blythe, John. M., & Johnson, Shane. D. (2019). A systematic review of crime facilitated by the consumer Internet of Things. *Secur J*. https://doi.org/10.1057/s41284-019-00211-8

Bravo-Lillo, C., Cranor, L. F., Downs, J., & Komanduri, S. (2011). Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *IEEE Security & Privacy Magazine*, 9(2), 18–26. https://doi.org/10.1109/MSP.2010.198

Cao, C., Guan, L., Liu, P., Gao, N., Lin, J., & Xiang, J. (2017). Hey, you, keep away from my device: Remotely implanting a virus expeller to defeat Mirai on IoT devices. *ArXiv:1706.05779 [Cs]*. http://arxiv.org/abs/1706.05779

Cardoso de Santanna, J. J. (2017). *DDoS-as-a-Service—Investigating Booter Websites* [PhD, University of Twente]. https://doi.org/10.3990/1.9789036544290

Castillo-Montoya, M. (2016). *Preparing for Interview Research: The Interview Protocol Refinement Framework*.

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

Çetin, Orçun, Altena, L., Gañán, C., & van Eeten, M. (2018). *Let Me Out! Evaluating the Effectiveness of Quarantining Compromised Users in Walled Gardens*.

Çetin, Orcun, Ganan, C., Altena, L., Tajalizadehkhoob, S., & van Eeten, M. (2019a). Tell Me You Fixed It: Evaluating Vulnerability Notifications via Quarantine Networks. *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, 326–339. https://doi.org/10.1109/EuroSP.2019.00032

Çetin, Orçun, Hanif Jhaveri, M., Gañán, C., van Eeten, M., & Moore, T. (2016). Understanding the role of sender reputation in abuse reporting and cleanup. *Journal of Cybersecurity*, *2*(1), 83–98. https://doi.org/10.1093/cybsec/tyw005

Chad Perrin. (2020). The CIA Triad. *TechRepublic*. Retrieved 27 May 2020, from https://www.techrepublic.com/blog/it-security/the-cia-triad/

Cheung, Ryan. (2017). *Targeting Financial Organizations With DDoS - A multi-sided perspective*. University of Twente.

Chromik, J. J., Santanna, J. J., Sperotto, A., & Pras, A. (2015). Booter websites characterization: *Proceedings of 33rd Brazilian Symposium on Computer Networks and Distributed Systems, SBRC 2015*, 445–458.

Cid-Fuentes, J. Á., Szabo, C., & Falkner, K. (2018). An adaptive framework for the detection of novel botnets. *Computers & Security*, *79*, 148–161. https://doi.org/10.1016/j.cose.2018.07.019

Coulter, R., & Pan, L. (2018). Intelligent agents defending for an IoT world: A review. *Computers & Security*, *73*, 439–458. https://doi.org/10.1016/j.cose.2017.11.014

Creswell, J. (2007). Qualitative Inquiry and Research Design: Choosing Among Five Approaches. *Health Promotion Practice*, *16*(4), 473–475. https://doi.org/10.1177/1524839915580941

Creswell, J. (2009). *Research Design—Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). SAGE Publications, Inc.

Dange, S., & Chatterjee, M. (2020). IoT Botnet: The Largest Threat to the IoT Network. In L. C. Jain, G. A. Tsihrintzis, V. E. Balas, & D. K. Sharma (Eds.), *Data Communication and Networks* (Vol. 1049, pp. 137–157). Springer Singapore. https://doi.org/10.1007/978-981-15-0132-6_10

De Donno, M., Dragoni, N., Giaretta, A., & Spognardi, A. (2018). DDoS-Capable IoT Malwares: Comparative Analysis and Mirai Investigation. *Security and Communication Networks*, *2018*, 1–30. https://doi.org/10.1155/2018/7178164

Denscombe, M. (2008). Communities of Practice: A Research Paradigm for the Mixed Methods Approach. *Journal of Mixed Methods Research*, *2*(3), 270–283. https://doi.org/10.1177/1558689808316807

Durumeric, Z., Payer, M., Paxson, V., Kasten, J., Adrian, D., Halderman, J. A., Bailey, M., Li, F., Weaver, N., Amann, J., & Beekman, J. (2014). The Matter of Heartbleed. *Proceedings of the 2014 Conference on Internet Measurement Conference - IMC '14*, 475–488. https://doi.org/10.1145/2663716.2663755

Ericsson, A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data. *Cambridge, Mass: MIT Press.*

European Union, & Agency for Network and Information Security. (2017). *Baseline security recommendations for IoT in the context of critical information infrastructures.* http://dx.publications.europa.eu/10.2824/03228

European union, & Council of Europe. (2004). *Common European Framework of Reference of Languages (CEFR)*. https://europass.cedefop.europa.eu/sites/default/files/cefr-en.pdf

Evans, D. (2011). *How the Next Evolution of the Internet Is Changing Everything*.

Fan, M., Shi, S., & Truong, K. N. (2020). *Practices and Challenges of Using Think-Aloud Protocols in Industry: An International Survey*. *15*(2), 18.

Fiebig, T. (2020). *BigBlueButton*. https://bbb.surfcloud.nl/b/

Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, *20*(2), 303. https://doi.org/10.1086/209351

Forget, A., Pearman, S., Thomas, J., Acquisti, A., Christin, N., Cranor, L. F., Egelman, S., Harbach, M., & Telang, R. (2016). *Do or Do Not, There Is No Try: User Engagement May Not Improve Security Outcomes*.

Furnell, S. (2007). Making security usable: Are things improving? *Computers & Security*, *26*(6), 434–443. https://doi.org/10.1016/j.cose.2007.06.003

Gartner. (2017). *Gartner says 8.4 billion connected 'things' will be in use in 2017, up 31 percent from 2016*. https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016

Gibson, G., USENIX Association, USENIX Conference on File and Storage Technologies, & FAST. (2017). *"I feel stupid I can't delete...": A Study of Users' Cloud Deletion Practices and Coping Strategies*. USENIX Association. https://www.usenix.org/legacy/events/fast05/tech/

Gill, K. S., Saxena, S., & Sharma, A. (2020). GTM-CSec: Game theoretic model for cloud security based on IDS and honeypot. *Computers & Security*, *92*, 101732. https://doi.org/10.1016/j.cose.2020.101732

Guo, H., & Heidemann, J. (2019). *USC/ISI Technical Report ISI-TR-726B*. 17.

Heartfield, R., Loukas, G., Budimir, S., Bezemskij, A., Fontaine, J. R. J., Filippoupolitis, A., & Roesch, E. (2018). A taxonomy of cyber-physical threats and impact in the smart home. *Computers & Security*, *78*, 398–428. https://doi.org/10.1016/j.cose.2018.07.011

Holl, P. (2015). *Exploring DDoS Defense Mechanisms*. 8.

Hornbæk, K. (2010). Dogmas in the assessment of usability evaluation methods. *Behaviour & Information Technology*, *29*(1), 97–111. https://doi.org/10.1080/01449290801939400

Huijts, N., Haans, A., Budimir, S., Fontaine, J., Loukas, G., Roesch, E., Bezemski, A., Oostveen, A.-M., & Ijsselsteijn, W. (2019). *Users' perceptions and responses to cyber-physical attacks on IoT devices in the home environment: A naturalistic field experiment*. https://www.eventbrite.co.uk/e/society-for-risk-analysis-benelux-conference-2019-registration-53855205369

Ishtiaq, M. (2019). Book Review Creswell, J. W. (2014). Research Design: Qualitative, Quantitative and Mixed Methods Approaches (4th ed.). Thousand Oaks, CA: Sage. *English Language Teaching*, *12*(5), 40. https://doi.org/10.5539/elt.v12n5p40

Joffe, H., & Yardley, L. (2004). *Research Methods for Clinical and Health Psychology, chapter 4: Content and thematic analysis*. SAGE Publications, Inc.

John, B. E., & Marks, S. J. (1997). Tracking the effectiveness of usability evaluation methods. *Behaviour & Information Technology*, *16*(4–5), 188–202. https://doi.org/10.1080/014492997119789

Johnson, S. D., Blythe, J. M., Manning, M., & Wong, G. T. W. (2020). The impact of IoT security labelling on consumer product choice and willingness to pay. *PLOS ONE*, *15*(1), e0227800. https://doi.org/10.1371/journal.pone.0227800

Kambourakis, G., Kolias, C., & Stavrou, A. (2017). The Mirai botnet and the IoT Zombie Armies. *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, 267–272. https://doi.org/10.1109/MILCOM.2017.8170867

Karami, M., & McCoy, D. (2013). *Understanding the Emerging Threat of DDoS-As-a-Service*.

Klobas, J. E., McGill, T., & Wang, X. (2019). How perceived security risk affects intention to use smart home devices: A reasoned action explanation. *Computers & Security*, *87*, 101571. https://doi.org/10.1016/j.cose.2019.101571

Kolias, C., Kambourakis, G., Stavrou, A., & Voas, J. (2017). DDoS in the IoT: Mirai and Other Botnets. *Computer*, *50*(7), 80–84. https://doi.org/10.1109/MC.2017.201

Kumar, A., & Lim, T. J. (2019). Early Detection Of Mirai-Like IoT Bots In Large-Scale Networks Through Sub-Sampled Packet Traffic Analysis. *ArXiv:1901.04805 [Cs]*. http://arxiv.org/abs/1901.04805

Kvale, S. (2007). *Doing interviews*. Thousand Oaks, CA: Sage.

Lee, Y., Park, Y., & Kim, D. (2015). Security Threats Analysis and Considerations for Internet of Things. *2015 8th International Conference on Security Technology (SecTech)*, 28–30. https://doi.org/10.1109/SecTech.2015.14

Levy, Y., & Ellis, T. (2006). A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. *Informing Science: The International Journal of an Emerging Transdiscipline*, *9*, 181–212. https://doi.org/10.28945/479

Lewis, C. (1982). *Using the 'Thinking-aloud' Method in Cognitive Interface Design*. IBM Thomas J. Watson Research Center.

Li, F., Durumeric, Z., Czyz, J., Karami, M., Bailey, M., McCoy, D., Savage, S., & Paxson, V. (2016b). *You've Got Vulnerability: Exploring Effective Vulnerability Notifications*.

Li, F., Ho, G., Kuan, E., Niu, Y., Ballard, L., Thomas, K., Bursztein, E., & Paxson, V. (2016a). Remedying Web Hijacking: Notification Effectiveness and Webmaster Comprehension. *Proceedings of the 25th International Conference on World Wide Web - WWW '16*, 1009–1019. https://doi.org/10.1145/2872427.2883039

Margolis, J., Oh, T. T., Jadhav, S., Kim, Y. H., & Kim, J. N. (2017). An In-Depth Analysis of the Mirai Botnet. *2017 International Conference on Software Security and Assurance (ICSSA)*, 6–12. https://doi.org/10.1109/ICSSA.2017.12

Mason, M. (2010). Sample Size and Saturation in PhD Studies Using Qualitative Interviews. *Forum: Qualitative Social Research.*, *11*.

McNamara, C. (2009). *General guidelines for conducting interviews.* http://managementhelp.org/evaluatn/intrview.htm

Ministerie van Economische Zaken. (2019). *Kamerbrief Voortgang Roadmap Digitaal Veilige Hard en Software*.

Mohamed, M. A., Chakraborty, J., & Dehlinger, J. (2017). Trading off usability and security in user interface design through mental models. *Behaviour & Information Technology*, 36(5), 493–516. https://doi.org/10.1080/0144929X.2016.1262897

Mohsin, M., Anwar, Z., Zaman, F., & Al-Shaer, E. (2017). IoTChecker: A data-driven framework for security analytics of Internet of Things configurations. *Computers & Security*, *70*, 199–223. https://doi.org/10.1016/j.cose.2017.05.012

Nahin, R. (2012). Observational studies and secondary data analyses to assess outcomes in complementary and integrative health care. *National Center for Complementary and Alternative Medicine.*

Nguyen, K. D., Rosoff, H., & John, R. S. (2017). Valuing information security from a phishing attack. *Journal of Cybersecurity*, *3*(3), 159–171. https://doi.org/10.1093/cybsec/tyx006

Nielsen, J. (1993). Usability Engineering. In *Usability Engineering* (p. iii). Elsevier. https://doi.org/10.1016/B978-0-08-052029-2.50001-2

Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. *Human-Computer Studies*, *41*, 385–397.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93*, 206–213. https://doi.org/10.1145/169059.169166

Noroozian, A., Korczyński, M., Gañan, C. H., Makita, D., Yoshioka, K., & van Eeten, M. (2016). Who Gets the Boot? Analyzing Victimization by DDoS-as-a-Service. In F. Monrose, M. Dacier, G. Blanc, & J. Garcia-Alfaro (Eds.), *Research in Attacks, Intrusions, and Defenses* (Vol. 9854, pp. 368–389). Springer International Publishing. https://doi.org/10.1007/978-3-319-45719-2_17

Online Trust Alliance (OTA). (2013). *Botnet remediation overview &practices*. https://www.internetsociety.org/wp-content/uploads/2017/10/ota_2013_botnet_remediation_best_practices.pdf

Pijpker, J., & Vranken, H. (2016). *The Role of Internet Service Providers in Botnet Mitigation*. 8.

Poole, E. S., Chetty, M., Morgan, T., Grinter, R. E., & Edwards, W. K. (2009). Computer help at home: Methods and motivations for informal technical support. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 739. https://doi.org/10.1145/1518701.1518816

Popper, K. (1934). *Logik der Forschung*. Julius Springer Verlag.

Rader, E., Wash, R., & Brooks, B. (2012). Stories as informal lessons about security. *Proceedings of the Eighth Symposium on Usable Privacy and Security - SOUPS '12*, 1. https://doi.org/10.1145/2335356.2335364

Redmiles, E. M. (2019). 'Should I Worry?' A Cross-Cultural Examination of Account Security Incident Response. *ArXiv:1808.08177 [Cs]*. http://arxiv.org/abs/1808.08177

Redmiles, E. M., Malone, A. R., & Mazurek, M. L. (2016). I Think They're Trying to Tell Me Something: Advice Sources and Selection for Digital Security. *2016 IEEE Symposium on Security and Privacy (SP)*, 272–288. https://doi.org/10.1109/SP.2016.24

Rossow, C. (2014). Amplification Hell: Revisiting Network Protocols for DDoS Abuse. *Proceedings 2014 Network and Distributed System Security Symposium*. Network and Distributed System Security Symposium, San Diego, CA. https://doi.org/10.14722/ndss.2014.23233

Rowe, B., & Wood, D. (2013). Are home internet users willing to pay ISPs for improvements in cyber security? *Economics of Information Security and Privacy*, *3*, 193–212.

Rowley, J., & Slack, F. (2004). Conducting a Literature Review. *Management Research News*, *27*(6), 31–39.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed). Wiley Pub.

Safaei Pour, M., Mangino, A., Friday, K., Rathbun, M., Bou-Harb, E., Iqbal, F., Samtani, S., Crichigno, J., & Ghani, N. (2020). On data-driven curation, learning, and analysis for inferring evolving internet-of-Things (IoT) botnets in the wild. *Computers & Security*, *91*, 101707. https://doi.org/10.1016/j.cose.2019.101707

Sanders, C. (2011). *Practical Packet Analysis, 2nd Edition* (2nd ed.). no starch press.

Scheepers, C. E., Wendel-Vos, G. C. W., den Broeder, J. M., van Kempen, E. E. M. M., van Wesemael, P. J. V., & Schuit, A. J. (2014). Shifting from car to active transport: A systematic review of the effectiveness of interventions. *Transportation Research Part A: Policy and Practice*, *70*, 264–280. https://doi.org/10.1016/j.tra.2014.10.015

Schwartzberg, S., & Couch, A. (2004). *Experience in Implementing an HTTP Service Closure*. 18.

Seidman, I. (2013). *Interviewing as qualitative research: A guide researchers in education and the social sciences* (4th ed.). NY: Teachers College Press.

Shadowserver. (2020). *What we do.* https://www.shadowserver.org/

Sheng, S., Kumaraguru, P., Acquisti, A., Cranor, L., & Hong, J. (2009). Improving phishing countermeasures: An analysis of expert interviews. *2009 ECrime Researchers Summit*, 1–15. https://doi.org/10.1109/ECRIME.2009.5342608

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, *76*, 146–164. https://doi.org/10.1016/j.comnet.2014.11.008

Signes-Pont, M. T., Cortés-Castillo, A., Mora-Mora, H., & Szymanski, J. (2018). Modelling the malware propagation in mobile computer devices. *Computers & Security*, *79*, 80–93. https://doi.org/10.1016/j.cose.2018.08.004

Silverio-Fernández, M., Renukappa, S., & Suresh, S. (2018). What is a smart device? - A conceptualisation within the paradigm of the internet of things. *Visualization in Engineering*, *6*(1), 3. https://doi.org/10.1186/s40327-018-0063-8

Sinanović, H., & Mrdovic, S. (2017). Analysis of Mirai malicious software. *2017 25th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1–5. https://doi.org/10.23919/SOFTCOM.2017.8115504

Smith, C. (2000). Content analysis and narrative analysis. *Handbook of Research Methods in Social and Personality Psychology*, 313–335. https://doi.org/Cambridge University Press.

Statista. (2018). *Internet of things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions).*

Stock, B., Pellegrino, G., Li, F., Backes, M., & Rossow, C. (2018). Didn't You Hear Me? - Towards More Successful Web Vulnerability Notifications. *Proceedings 2018 Network and Distributed System Security Symposium*. Network and Distributed System Security Symposium, San Diego, CA. https://doi.org/10.14722/ndss.2018.23171

Turner, D. W. (2010). *Qualitative Interview Design: A Practical Guide for Novice Investigators*.

United Nations (Ed.). (2020b). *Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*. Springer Publishing Company. https://doi.org/10.1891/9780826190123.0013

United Nations. (2020a). *Sustainable Development Goals*. https://www.un.org/sustainabledevelopment/

van den Haak, M., De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology*, *22*(5), 339–351. https://doi.org/10.1080/0044929031000

van Eeten, M., & Bauer, J. (2008). *Economics of Malware: Security Decisions, Incentives and Externalities* (OECD Science, Technology and Industry Working Papers No. 2008/01). https://doi.org/10.1787/241440230621

van Someren, M., Barnard, Y., & Sandberg, J. (1994). The think aloud method: A practical guide to modelling cognitive processes. *Information Processing & Management*, *31*(6), 906–907. https://doi.org/10.1016/0306-4573(95)90031-4

Vasek, M., & Moore, T. (2012). *Do malware reports expedite cleanup? An experimental study*.

Verstegen, S. (2019). *Understanding the role of IoT end users in Mirai-like bot remediation*. Delft, University of Technology.

Wee, B. V., & Banister, D. (2016). How to Write a Literature Review Paper? *Transport Reviews*, *36*(2), 278–288. https://doi.org/10.1080/01441647.2015.1065456

Wolff, B., Mahoney, F., Lohiniva, A. L., & Corkum, M. (2019). Collecting and Analyzing Qualitative Data. *The CDC Field Epidemiology Manual*, 213–228.

Zeng, E., Mare, S., & Roesner, F. (2017). *End User Security & Privacy Concerns with Smart Homes*.

Zimmermann, V., Gerber, P., Marky, K., Böck, L., & Kirchbuchner, F. (2019). Assessing Users' Privacy and Security Concerns of Smart Home Technologies. *I-Com*, *18*(3), 197–216. https://doi.org/10.1515/icom-2019-0015

Appendices

Appendix A. Literature search and analyses summary

The literature search is performed using the stepwise approach that is suggested by van Wee & Banister (2016) and Levy & Ellis (2006). First an initial search takes place by using keywords. Then the found citations in the found articles will be analysed and lastly, the articles that cite the identified articles will be analysed. The used databases, keywords, time boundaries and languages can be seen in table A1. For each of the consulted databases, the number of hits is shown, as well as the size of the subset of hits that is analysed and the size of the subset that is relevant. Lastly, the date on which the search has been performed is shown. Overlapping results in databases or searches are added to the oldest search cell.

For the literature search, two databases and two journals are considered: IEEE Xplore, Scopus, The Oxford Journal of Cybersecurity and Computers and Security by Elsevier. The journals have been chosen by performing a DuckDuckGo search on "Journal AND (computer security OR cybersecurity)" and selecting the two top results. IEEE Xplore and Scopus will be used for their wide range of literature. Because the word security is part of the name of both journals, the search phrase will be only on title, abstract and authors.

Note that the articles are categorized into one of four research subjects: The rise of the Internet of Things, The future of computer crime: Mirai malware, Existing remediation tactics and evaluation, or End-users behaviour in IoT security. For each of these categories, a different keyword search has been performed.

	Source	**Keywords**	**Hits**	**Analysed hits**	**Relevant hits**	**Date of search**
1.1	Computer & Security (Elsevier)	(IoT OR 'Internet of Things') AND (secure OR security)	57	57	4	22/03/2020
1.2		Mirai OR (botnet AND (IoT OR 'Internet of Things'))	86	50	4	24/03/2020
1.3		(remediation OR notification OR warning) AND (malware OR malicious OR botnet)	471	50	1	26/03/2020
1.4		(behaviour OR perception OR actions OR intention) AND (security OR secure)	2807	50	2	30/03/2020
2.1	Journal of Cybersecurity (Oxford)	(IoT OR 'Internet of Things') AND (secure OR security)	6	6	0	22/03/2020
2.2		Mirai OR (botnet AND (IoT OR 'Internet of Things'))	5	5	0	24/03/2020
2.3		(remediation OR notification OR warning)	27	27	2	26/03/2020

		AND (malware OR malicious OR botnet)				
2.4		(behaviour OR perception OR actions OR intention) AND (security OR secure)	39	39	0	30/03/2020
3.1	Scopus	(IoT OR 'Internet of Things') AND (secure OR security)	16816	50	2	23/03/2020
3.2		Mirai OR (botnet AND (IoT OR 'Internet of Things'))	591	50	5	25/03/2020
3.3		(remediation OR notification OR warning) AND (malware OR malicious OR botnet)	497	50	5	27/03/2020
3.4		(behaviour OR perception OR actions OR intention) AND (security OR secure)	89168	50	3	30/03/2020
4.1	IEEE Xplore	(IoT OR 'Internet of Things') AND (secure OR security)	12174	50	1	23/03/2020
4.2		Mirai OR (botnet AND (IoT OR 'Internet of Things'))	366	50	2	25/03/2020
4.3		(remediation OR notification OR warning) AND (malware OR malicious OR botnet)	208	50	2	27/03/2020
4.4		(behaviour OR perception OR actions OR intention) AND (security OR secure)	21525	50	2	30/03/2020
5	Snowballing	-	-	20	7	31/03/2020

Table A1. Literature search and results

By using forward and backward snowballing on the identified articles, 20 additional potential articles were found, of which 7 are relevant. Table A2 shows the relevant hits of both the keyword search, as well as the articles found during the snowballing. The table shows the title of the relevant hits, but also the most important findings, the knowledge gaps, the number of citations and how it was found. Lastly, the table displays the relevance of the articles for this research.

Note that some of the papers were acquired through a different path than the literature search. This indicates that the article was acquired during discussions, or as initial reading into the subject. If such a case was also found during the literature search, it is still reported as acquired before the search. If an article was found in multiple databases, the first source that was searched has been reported. Lastly, cases that could fit into more than one of the categories are only displayed in their most fitting category.

Reference	Findings	Knowledge gaps / future work	Source	Relevance
The rise of the Internet of Things				
Anonymous authentication for privacy-preserving IoT target-driven applications (Alcaide, 2013)	- IoT devices become a part of our daily lives, which means the security and privacy issues should be taken seriously - There are ways to secure privacy and security issues, by using the proposed design. - The design does not rely on a central node in the network, and data collectors and users can be attuned to one another.	- What actor should implement the network scheme presented? - How can the efficient working of each of the devices be ensured?	Computers & Security (Elsevier)	- Although there are thoughts about how the IoT environments should be setup, it is unclear how this can be reached. Also, as IoT becomes an increasing part of our lives, it is necessary to take the issues that come with it seriously.
A taxonomy of cyber-physical threats and impact in the smart home (Heartfield et al., 2018)	- There are many different types of threats that can hurt a smart home. - Progress can be made within the user security posture to defend better against threats for smart homes, which is called Human-as-a-Security-Sensor (HaaSS)	- How effective will HaaSS be? -How can HaaSS be achieved? - How can the effectiveness of HaaSS be measured?	Computers & Security (Elsevier)	- There is definitely space for end-users to contribute to the defence of IoT environment. To achieve this HaaSS, it is however, necessary that the attitude of end-users changes and awareness is created as well as the right knowledge to become a sensor.
Modelling the malware propagation in mobile computer devices (Signes-Pont et al., 2018)	- Diseases spread fast, which is not different in the case of IoT malware. - User awareness could mean a shift in the tipping point for an exploding number of infections.	- To what extent does the model relate to the real world? - The model has not yet included human interaction with the system, which means it is unclear how adaptation could affect the system.	Computers & Security (Elsevier)	- Although Mirai is known at this point, and also the signature for recognizing it, there is no effective way yet of removing the “disease” from the online world. We still depend on user interaction with their IoT network, which means that is the place to invest.
EclipseIoT: A secure and adaptive hub for the Internet of Things	- The heterogeneity of IoT devices asks for an advanced tool for scanning the devices for security issues, EclipseIoT could be that advanced tool.	- It is not yet implemented how IoT devices can be protected from each other. It is not hard to imagine how one infected device can infect others within the	Computers & Security (Elsevier)	- Once again, the heterogeneity of IoT devices asks for a smart solution. Although the paper addresses a technical solution to this issue, there is no real answer

(Anthi et al., 2018)	- Implementing such a tool is relatively difficult to manage, as it relies on many services (PubNub in this case).	same network. This is a whole different type of challenge.		yet to the social side of the issue and how the solution will cope with this social side. The heterogeneity should be kept in mind.
Design and implementation of automated IoT security testbed (Abu Waraga et al., 2020)	- The number of IoT devices grows faster than the security measures, meaning more and more issues arise. - Devices should be tested before entering the market, but this is too expensive. Therefore an open source database of known IoT device issues should be started. Some white hat hackers can then find issues and add them for all end-users to see and solve. - There is a big lack of standardization in the big market of IoT devices.	- How can the testbed be implemented and become known for all users that have knowledge about lacks in IoT security. - Even if the issues are known, the products are still out on the market, and there is no easy way to fix the issues, as users lack knowledge and vendors lack incentives.	Scopus	- Once more, end-users seem to be key players in solving the issues that IoT devices bring along. Although it would be helpful to start sharing common IoT issues and their solutions openly, it is still important that users are incentivized to find this database and act upon it. This again moves towards the creation of awareness and skills with end-users.
Security, privacy and trust in Internet of Things: The road ahead (Sicari et al., 2015)	- There lacks ideas about how to secure the IoT before it becomes big enough to cause serious problems. - Insurance could be an option to relief the users from their responsibilities.	- How can we make sure that the deployed IoT devices actually follow the security, privacy and trust policies?	Scopus	- This paper sheds light upon the different issues that IoT devices still have and points at a direction where governments setup strict policies for minimal security efforts for IoT devices. However, it is hard for local governments to keep users from importing IoT devices with poor security. Therefore, a hint is given towards awareness amongst end-users, next to governments and other big organisations.
Security Threats Analysis and Considerations for Internet of Things (Lee et al., 2015)	- Everything can be hacked, which is harmful in the case of for example smart cars or smart homes. - IoT mostly does not have the ability of a clear user interface, which makes it hard for users to	- How can simple and weak IoT devices be secured from getting infected? - Which actor is the right one to appoint for creating	IEEE Xplore	- This paper is one of the first to ask some serious attention for the weaknesses of most IoT devices. As more critical parts of our infrastructures are driven by IoT devices, it is important to become

	make changes to the settings of the device. The IoT exists of black boxes.	responsibilities for solving the failing security.		aware of the issues, which cannot only be done at the producers side.
The future of computer crime: Mirai malware				
Intelligent agents defending for an IoT world: A review (Coulter & Pan, 2018)	- Hard to detect all different intrusions - The aspect of IoT that they are constantly interacting creates more windows of opportunity for attacks - Coordination and knowledge sharing are key in defending against all types of malware	- Signature identification for all types of malware is still needed - How can a defence be created that is layer independent - How will future IoT systems evolve and what are their challenges?	Computers & Security (Elsevier)	- As new devices and attackers will constantly have new ways of intruding IoT devices, the solution for cleaner IoT environments should also be looked at with end-users and other actors that come in only after IoT devices reach the market.
GTM-CSec: Game theoretic model for cloud security based on IDS and honeypot (Gill et al., 2020)	- Game theory suggests that we should be able to discover attackers well enough with the use of honeypots, signature-based and anomaly-based detection measures. However, this is only about discovering, not about cleaning.	- The game theory could be experimented with in the real world, which means altering defence mechanisms to minimize costs.	Computers & Security (Elsevier)	- The problem clearly is no longer at the bottleneck of discovering attacks and attackers. What remains a problem is acting upon these discoveries, which sometime has to happen at ground level.
An adaptive framework for the detection of novel botnets (Cid-Fuentes et al., 2018)	- The algorithm in this paper can recognize new arising botnets without using historical data on older botnets such as Mirai. - With the use of adaptive training, the algorithm can recognize unknown botnets relatively accurate (100% if false alarms are tolerated and 48% if no false alarms can be made)	- As with most techniques to defend against attackers, this tool is only usable temporarily. As soon as attackers discover this tool, they will find ways around it to stay up front in the arms race.	Computers & Security (Elsevier)	- It is interesting to see how researchers and companies try to arm themselves against attacking botnets. However, the solution only starts at discovering the botnet. The steps after the discovery are more crucial in the defence strategy, such as informing vendors, and owners of devices. Moreover, it should also become clear how the malware can be removed from botnet soldiers.
On data-driven curation, learning, and analysis for inferring evolving internet-of-Things	- Although IoT covers more and more critical infrastructures, the problem do not fade away. - In 24 hours 400,000 IP addresses were compromised mostly with Mirai, Hide and Seek and Reaper	- Are there more signatures for the malware binaries? - Do differences exist for IoT exclusive malware?	Computers & Security (Elsevier)	- Mirai remains a large problem in 2020. Also, other types of malware seem to be related to Mirai, which makes the combat more important. As IoT covers more and more

(IoT) botnets in the wild (Safaei Pour et al., 2020)				critical infrastructures, the environment should become cleaner.
Understanding the Mirai Botnet (Antonakakis et al., 2017)	- Mirai was able to grow so big because of the lack of security in IoT devices, even though the code is relatively simplistic. - The mostly attacked devices are security cameras and DVRs. - Mirai infected IoT devices have four states: scan, kill, wait, attack. - While scanning, new devices are brute forced for getting access, using a dictionary of standard username, password mechanisms.	- Mirai should be the best wake-up call for governments, ISPs and other stakeholders to start cleaning up the badly secured IoT environment.	Scopus	- As Mirai makes use of the worst weaknesses in the IoT environment, even small steps in security could decrease the size of the problem significantly. Therefore, increased awareness on its own could already have an impact. - As the Mirai source code has been shared on the internet, the signature for its presence is also known, which could be used by anyone to recognize an infection.
Early Detection Of Mirai-Like IoT Bots In Large-Scale Networks Through Sub-Sampled Packet Traffic Analysis (Kumar & Lim, 2019)	- This paper tried to achieve a maximum efficiency on pinpointing Mirai infected IoT devices, by using the signature of Mirai. - More vulnerable devices have a higher probability of being infected than non-vulnerable devices. - Sometimes there can be a large delay between the infection and the detection, which is caused by either the sleeping state of Mirai, or the lack of measuring all traffic of all devices, which would require a large amount of memory data.	- Mirai might have a hard time evading detection software because of its blunt signature, but newer and altered malware might implement tricks to evade the common measurements of network traffic. This arms race should not be surrendered.	Scopus	- When measuring traffic data of supposedly infected IoT devices, it is important to keep in mind the boundaries such as time and data memory, which can delay detections or completely oversee infections. Also, it is important to share knowledge of newly discovered malware which might be able to evade some of the detection tools.
Analysis of Mirai Malicious Software	- Mirai infected IoT devices are recruiters for new soldiers for the botnet. This is why the botnet explodes in size. Through CNC server, attackers can locate their bots and send them orders. The signature that can detect Mirai is the following:		Scopus	- Two papers have established the signature for Mirai, which is the same in both papers. Also the behaviour of infected devices seems to overlap. This ensures the

(Sinanović & Mrdovic, 2017)	TCP sequence number is equal to the destination IP address.			detection of Mirai during virtual visits at consumers homes.
The Mirai Botnet and the IoT Zombie Armies (Kambourakis et al., 2017)	- The paper elaborates upon the processes that Mirai start within a device, coming down to the same steps as other articles describe. - The steps that are prescribed as mitigation techniques align with those of the protocol of KPN, namely, resetting the password for both device and router and rebooting both devices.	- The paper fails to point out how the prescribed measures should be executed. Due to the interference of a modem or router, the actual infected device stays hidden in a users' home. Therefore, users will be the only actor that is able to mitigate the infection. - Also, although the paper states that the vendors should create cleaner and safer products, it fails to state how these vendors can be incentivized to do so.	Scopus	- Another article that repeats the known processes of Mirai, but also entails the needed steps to ensure remediation. This means that the protocol of ISPs is correct. However, it remains a question how to make sure that a detected bot actually is mitigated.
IoT Botnet: The Largest Threat to the IoT Network (Dange & Chatterjee, 2019)	- As the impact of IoT botnets is relatively high compared to the traditional existing malware, it should be priority to create a safer IoT. - Prevention is better than remediation, but both sides should be invested in.		Scopus	- The paper strengthens the idea that only focussing on the prevention of infections will not be sufficient in the short run. To be able to create a cleaner IoT environment, users should be mobilised. Because of the constantly increasing size of the IoT, the issue grows bigger too, which means faster remediation is valuable.
An In-Depth Analysis of the Mirai Botnet (Margolis et al., 2017)	- This paper illustrates the different states of Mirai malware and shows how it can spread fast. - Although the paper suggests that IoT device manufacturers have the key role in improving security, it also states that end-users could help in the remediation.		IEEE Xplore	- The paper aligns with others on the best way forward for improving, namely incentivizing manufacturers and vendors into bringing more secure products into the market. They fail, like the other papers in

				explaining how this should be shaped. Unlike the other papers on this subject, this article does suggest also looking into the user side of the IoT, which supports the relevance of this research.
DDoS in the IoT: Mirai and Other Botnets (Kolias et al., 2017)	- This paper is one of the first to elaborate on the story of Mirai and its way of working. Mirai has four components: the **bot**, which is the malware that infects, the **command & control server** which gives the attacker the possibility to manage their bots, the **loader** which supports the spread of the executables and the **report server,** which holds records on all infected devices. - The paper holds device vendors responsible for not caring more about the security issues of the products on their shelves.		IEEE Xplore	- The paper is a wakeup call that shows the large impact of Mirai and points a finger at device vendors. However, these vendors feel no incentive to sell secure devices at all. It is also hard to provide this incentive for them, which means the solution probably lies with other parties, such as the buyers of the devices, end-users.
A Survey on Security Threats and Solutions in the Age of Iot (Atac & Akleylek, 2018)	- Lots of IoT devices lack in security and should be enhanced to the required level from the design phase. - After devices enter the market and customers, internet traffic should be measured constantly to point out malicious traffic so action can be taken timely.		Snowballing	- The paper illustrates the many issues that a lack in IoT security can cause. In section 3.6, human factors turn out to be a large reason for insecurity and should thus be addressed in the solution space.
DDoS-Capable IoT Malwares: Comparative Analysis and Mirai Investigation (De Donne et al., 2018)	- This paper illustrates the different states of Mirai malware and shows how it can spread fast. - It shows the “skeleton” of Mirai and that it could become as big as it is due to the lack of security in IoT devices. - This lack of security can be linked to the fast economies, which ask for products to be on the market as soon as possible. This is at the cost of side qualities of the product, such as security.		Snowballing	- The paper shows once more, why the IoT is so poorly secured, and why a simple malware such as Mirai could take over. Moreover, the paper illustrates why Mirai remains an eminent threat for IoT devices and why this threat should be eliminated, or decreased to say the least.

Existing remediation tactics and evaluation				
IoTChecker: A data-driven framework for security analytics of Internet of Things configurations (Mohsin et al., 2017)	- The best place to increase security for IoT is in the planning and manufacturing phases - IoT devices that support over-the-air (OTA) updates have better security - IoT is not safe enough to be responsible for the most critical processes. Minimize IoT usage for these processes. - Products from the same vendor have similar issues, which can be fixed in the same way.	- How can IoT devices be supported for constant automatic updates? - There is need for a centralized repository for IoT devices, their weaknesses and the solutions to those weaknesses	Computers & Security (Elsevier)	- Although the main issues lie with the manufacturers and vendors of weakly protected IoT devices, the overall security of the environment can be increased with the awareness of end-users other actors that play a role after consumption.
What security features and crime prevention advice is communicated in consumer IoT device manuals and support pages? (Blythe et al., 2019)	- The information shared in IoT device manuals and support pages is not adequate enough to inform consumers about the security options. - Even those who did disclose to some extent the security possibilities, cyber hygiene advice (do's and don'ts) is the rarest information.	- Some parts of the security information in manuals should be standardized across the market.	Journal of Cybersecurity (Oxford)	- This paper clearly shows that the problem of security issues in IoT is not only the fault of manufacturers or consumers, but also the way in which information is shared between the two parties. As there are no real incentives to make devices as secure as possible, the effort to include well structured, useful information in manuals is too high. Therefore, either incentives should be created, or other parties should take up the task of informing consumers.
Understanding the role of sender reputation in abuse reporting and cleanup (Çetin et al., 2016)	- No evidence is found that the reputation of the sender of notifications influences the cleanup rate of infected devices. - Detailed notices do improve the cleanup rate compared to relatively simple messages.	- The reason for receivers of notifications to act is unexplored territory.	Journal of Cybersecurity (Oxford)	- This research excludes one of the dimensions of a notification, namely the sender reputation. This suggests that the solution space for an optimal notification covers all of the possible senders.

	- Hosting providers have a higher cleanup rate when there is a link to an informative website, but this is not the case for end-users. - Personal help might increase the cleanup rate.			- The research hints at the importance of technical knowledge of the internet environment to be able to perform clean-ups as good as possible.
Let Me Out! Evaluating the Effectiveness of Quarantining Compromised Users in Walled Gardens (Çetin et al., 2018)	- Roughly 50% to 75% of quarantined users manages to clean their infected machine. - The reinfection rate was very low, meaning quarantined users adapted their online behaviour after their experience in quarantine. - Although quarantining seems to be successful, the costs for ISPs are very high due to the many complaints, and calls to the helpdesk.		Scopus	- Walled-gardens seem to be one of the most effective ways in increasing the remediation rate and even have long term effects as they change the online behaviour of users. However, as the costs are also remarkably high for IPSs, both time consuming and a loss in customer satisfaction. The most important finding is that incentives can be created for users to change their behaviour.
The Matter of Heartbleed (Durumeric et al., 2014)	- Although the Heartbleed vulnerability was publicly disclosed, not all vulnerable parties reacted appropriately to remove the risk. (3% still did not after 2 months) - Sending notifications to the remaining parties had a positive impact on the actions of parties in performing the necessary patches.		Scopus	- This paper shows the value of notifying users about their lack in security. Although the incentive for performing the necessary steps are clearer and more direct for users, as they carry the risks themselves, the paper still hints at creating awareness as a useful policy lever.
Remedying Web Hijacking: Notification Effectiveness and Webmaster Comprehension (Li et al., 2016a)	- Direct communication with affected parties will increase the actions by these affected parties. - Precise infection details and providing tools to perform cleaning has a positive impact.		Scopus	- Although the “end-users” in this paper are webmasters, who have a clear incentive to cooperate, and the infections are of a different kind, this paper shows the importance of close contact with the end-users and the provision of usable tools and information.

You've Got Vulnerability: Exploring Effective Vulnerability Notifications (Li et al., 2016b)	- The study found that notifications have a positive impact on patching by network operators. - Direct messages with detailed information work the best. - Although the notifications were received in a positive way, the effects of remediation were disappointing.	- It is important to understand better what parts of notifications are able to trigger receivers to act.	Scopus	- This paper once again shows the importance of sending detailed information with the notification and contacting the end-user in a direct and personal way. It also asks for more research into the specific parts of the notification that "make or break" the effects.
Stories as informal lessons about security (Rader et al., 2012)	- Story telling could be an effective way of increasing security awareness all over. Instead of having a technical protocol on how to perform the needed actions, it could be useful to add several existing stories on what could go wrong to increase the impact of a notification.	- It is quite unsure whether storytelling also leads to better security, next to the changed thoughts and awareness.	Scopus	- This study strengthens the thought that end-users need a detailed and personal approach to achieve the best remediation. It could be useful to look more into the way the content of notifications is presented.
Tell Me You Fixed It: Evaluating Vulnerability Notifications via Quarantine Networks (Çetin et al., 2019a)	- Both email notifications and walled gardens significantly increase the remediation rate of infections. - Some users seem to be unable or unwilling to perform the requested actions. - Although walled gardens are most effective, user experience is negative.	- It is still largely unsure what the optimal content of a notification should be. Tests with different types of content should be performed - A cost-benefit analysis should point out whether the additional clean up with the use of walled gardens can compensate for the pushback of customers and the extra time and effort spent by ISP employees.	IEEE Xplore	- The paper clearly shows need for more research into the notification mechanism and the reactions of end-users.
The role of internet service providers in botnet mitigation	- ISPs in the Netherlands perform relatively good with respect to botnet mitigation. Still, botnets are a significant threat in the Netherlands. - ISPs are key players in all 5 stages of botnet mitigation: Prevention, Detection, Notification, Remediation and Recovery.	- ISPs should improve their information sharing both with their customers as well as their peers.	IEEE Xplore	- This paper explains the hard position that ISPs find themselves in. It also points out that there are steps to be taken by ISPs, in most of the stages of botnet mitigation.

(Pijpker & Vranken, 2016)				
Evidence on ISP and Consumer Efforts to Remove Mirai (Çetin et al., 2019b)	- The usage of walled gardens can increase the remediation rate of infected IoT devices significantly. The usage of email notification, however, cannot. - Walled gardens are too economically expensive, which makes the option not viable for many ISPs.	- Either a cost-benefit analysis is needed to find out whether the high costs of walled gardens are worth the increased effectiveness, or the alternative notification mechanisms should become more successful.	Snowballing	- This study gives a useful oversight of the possible channels for notifications. The dilemma between costs and effectiveness is something that could be researched more thoroughly.
Didn't You Hear Me? — Towards More Successful Web Vulnerability Notifications (Stock et al., 2018)	- The paper sheds light on how vulnerability notifications are received by domain owners, with a data set of 24000 sent notifications. - Email notifications can be stopped by spam filters or a lack of trust by the receiver. As notification emails are rarely send, the receivers might not expect such a message and ignore it. - The is a need for trust by the receiver in the sender for the notification to work. The study suggests that the sender reputation does influence the effect of the notification, which contradicts earlier research.	- Research is needed on the required technical level that is needed in the notifications in order to be successful. As the level of knowledge differs largely among receivers, a minimum of explanation should be found that can give each of the receivers the necessary knowledge and information.	Snowballing	- Although the study is not about regular consumers, but about domain owners, it shows the large variety of options when notifying receivers about vulnerabilities or infections. The study shows the hurdles for email notifications, which should be kept in mind when finding the best notification mechanism. Also, this research shows the importance of activating people to act more than they do at this moment in time.
Do malware reports expedite clean up? An experimental study (Vasek & Moore, 2012)	- This study researched the effectiveness of notifications for remediating malware. 32% of malware distributing websites is cleaned within a day compared to 13% that did not receive a notification. - Notices should be detailed to work properly. - Only one notification should be sent. Returning notifications will enlarge the likelihood of receivers thinking of the notifications as spam or not useful.		Snowballing	- This paper shows the importance of getting the dimensions of a notification mechanism right. One, clearly explained notification has the most wanted effect. Although the setting is slightly different than in informing ISP consumers, it is likely that the findings hold for that setting as well. At least it is valuable to keep in mind the displayed do's and don'ts.

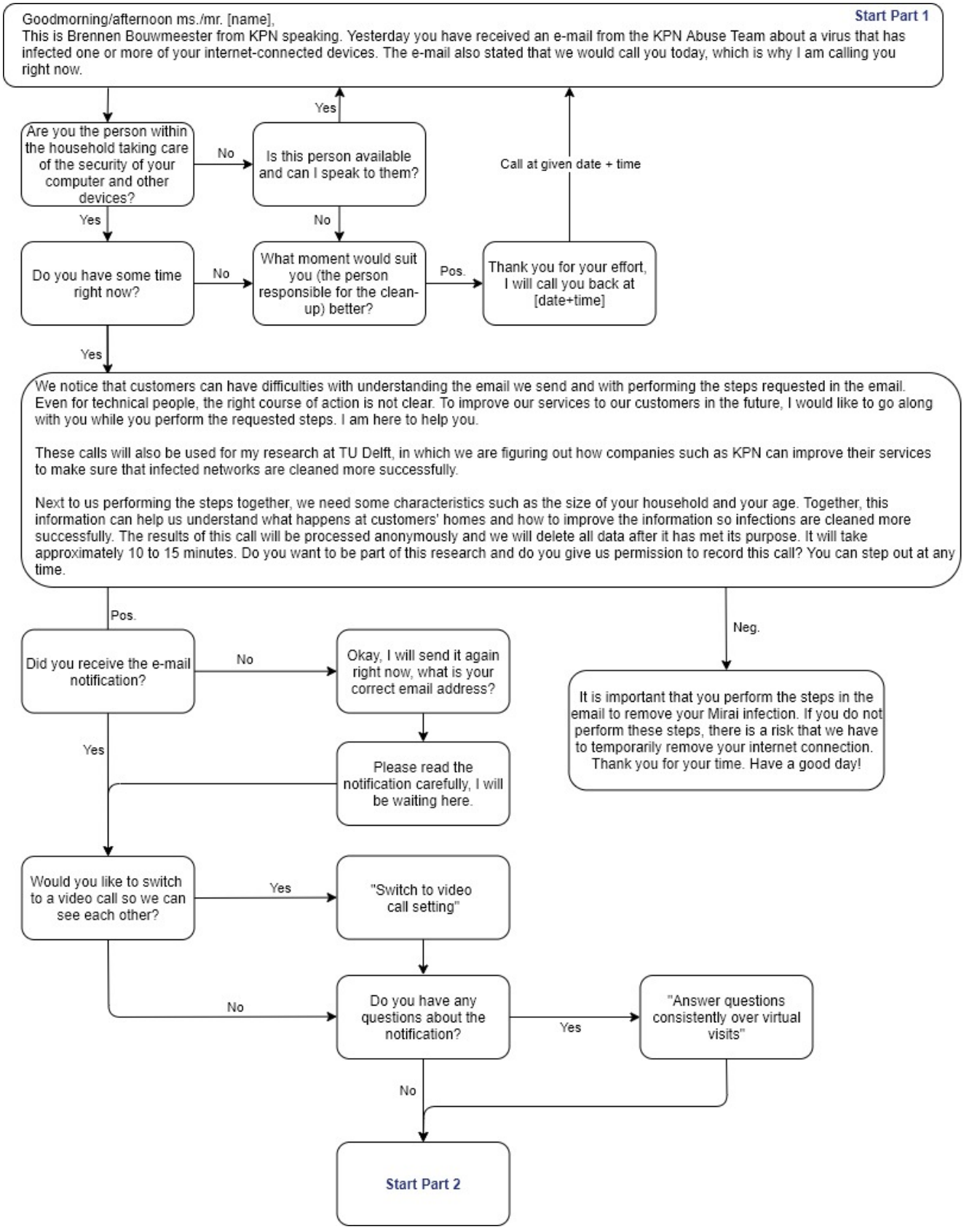
End-users behaviour in IoT security				
How perceived security risk affects intention to use smart home devices: A reasoned action explanation (Klobas et al., 2019)	- Perceived security issues play no role in the purchasing of IoT devices by consumers. - The feeling of having control over the security does affect the usage of IoT negatively	- Users lack knowledge about the security issues of IoT devices, which should be supplied by media: the question is how? - Research and experts on IoT security and risks should guide users: the question is how?	Computers & Security (Elsevier)	- Users do care about having control over their security risks, but lack the knowledge to be able to take an informed decision. - If users are guided through the dangers of IoT, this should have a positive impact on their behaviour.
Making security usable: Are things improving? (Furnell, 2007)	- Security lacks in multiple widely used computer software applications, such as internet explorer and word. - Users are asked to take decisions about their security settings, but lack in supporting users with the needed info.	- How can users be informed supplied with the right information? - How can manufacturers or software creators be incentivized to secure better?	Computers & Security (Elsevier)	- Even though this research took place in 2007, there is a returning issue of lacking information sharing with consumers. Also, the lack of incentive for producers to create secure products is apparent.
Assessing Users' Privacy and Security Concerns of Smart Home Technologies (Zimmermann et al., 2019)	- Users are more interested in the ends of IoT devices than in the means that are used. - Ironically, users seem to purchase several IoT devices for the sake of safety for their homes. These devices, however, lack in security of a different kind. - The consumers that do care about the security issues often lack knowledge about the heterogeneous products to be able to take the needed steps to improve the security level of their IoT devices.	- There is no quantitative data yet on the understanding by end-users on their home environment. - A different scope than Germany could backup the results.	Scopus	- This research clearly shows the need for more knowledge sharing with end-users. Many users seem to lack interest and/or knowledge about a secure IoT environment. - A first step in helping users to understand the security of their devices better is to add labels.
Do or Do Not, There Is No Try: User Engagement May Not Improve Security Outcomes	- This study compared users attitudes and skills to their performed actions. - There is a large variance in engagement to secure ones machines.	- More research is needed into the content, presentation and functionality of security advice (notifications).	Scopus	- Designing a one-size-fits-all notification might be hard due to the large variety in security engagement, awareness and skill.

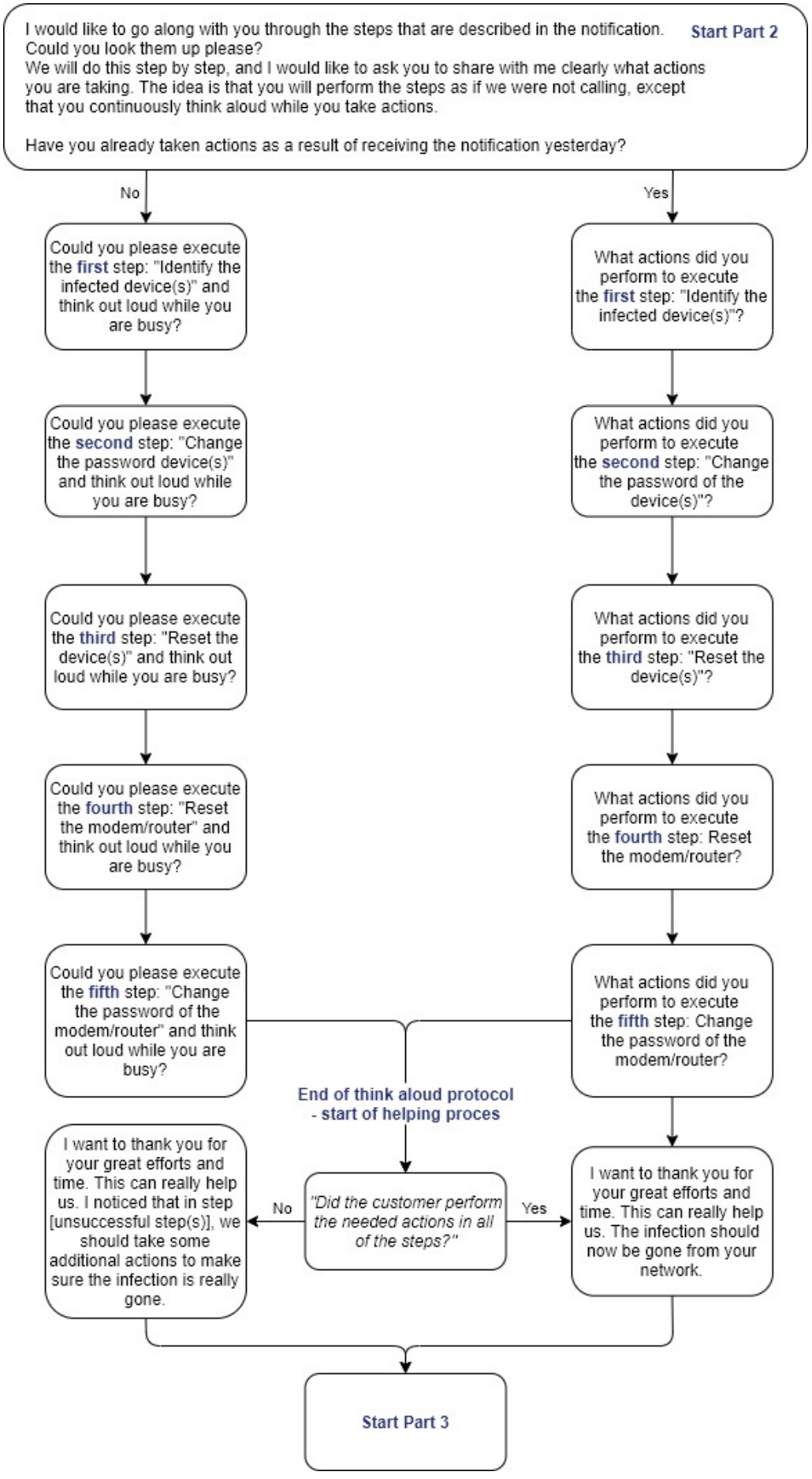
(Forget et al., 2019)	- The security level is dependent on ones awareness of their level of expertise on security issues.			
“I feel stupid I can’t delete...”: A Study of Users’ Cloud Deletion Practices and Coping Strategies (Gibson et al., 2005)	- This study is one of the first to look into users mental models in the area of computer security. It seems that many mental models of users are incomplete, which leads to inappropriate actions.		Scopus	- Although this study took place in 2005, and is about deleting files from the cloud, an interesting overlap can be found in the issues that arise in this article. Due to the wrongly perceived knowledge of end-users, they perform the wrong actions, while thinking that the right and needed actions were taken. This can be a problem in mapping users’ actions and thinking process.
"Should I Worry?" A Cross-Cultural Examination of Account Security Incident Response (Redmiles, 2019)	- The uncertainty of the cause of the security problem can remove effectivity of a notification. As users are unsure about the reason of their security to fail, they are also unsure how to improve it. - A lack of information in the notification might have a counterproductive effect. As users keep on receiving notifications but are unsure about the cause of the solution, they start to believe the notifications themselves are fake, or phishing.	- More transparency might improve the effects of notifications.	IEEE Xplore	- This research shows the importance of humanlike interaction during a notification mechanism. As some users are unaware of security issues, they might see the notification itself as a threat. Transparency in the notification can improve the effect.
I Think They're Trying to Tell Me Something: Advice Sources and Selection for Digital Security (Redmiles et al., 2016)	- This paper is about the best practices for computer security advice in general. The first finding is that users lack the skills to determine whether notifications are useful or not. - The second finding is that the effectiveness of a notification decreases as it holds more marketing content or threatens users’ sense of privacy. -Also, users seem to believe that someone else is responsible for their digital security.		IEEE Xplore	- Mostly the finding about what scares users of is useful in this paper. As users need information about how to clean their infected IoT device, it is necessary to give users this information. However, the more information that an ISP gives the user, the more this user will feel like their privacy is

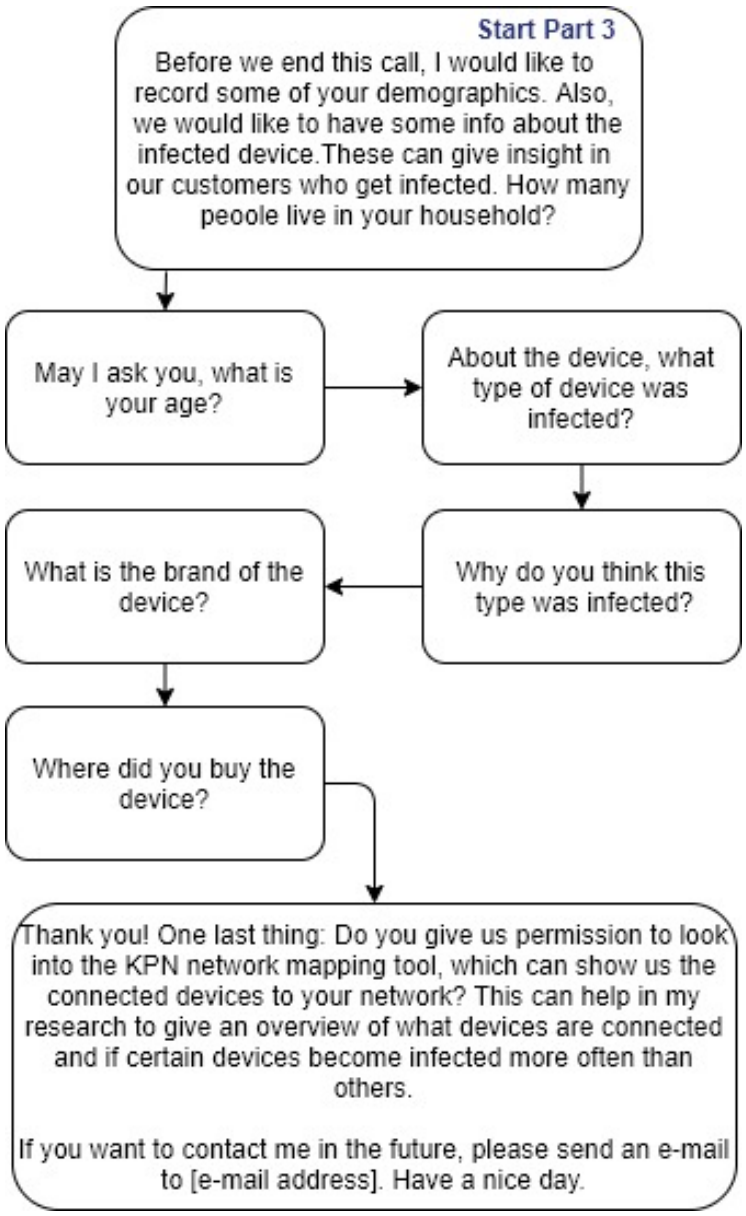
				threatened. This paradox of users needing detailed info, but not feeling like their privacy is invaded can become a problem in finding an optimal notification mechanism.
Users' perceptions and responses to cyber-physical attacks on IoT devices in the home environment: a naturalistic field experiment (Huijts et al., 2019)	- IoT devices have inherent flaws, which makes users believe that their malfunctioning is caused by the bad quality of the device itself. Rarely, users will look at malfunctioning as a result of interference by malware or hackers.		Snowballing	- This paper shows the importance of notifications in the first place, for users to become aware that malfunctioning is caused by something that is not inherent with the device itself. However, it could remain a problem for users to recognize their infected device, as they might believe the quality of the device is poor.
End User Security and Privacy Concerns with Smart Homes (Zeng et al., 2017)	- End users often lack the knowledge to take the right steps in the case where their Internet of Things device has been compromised. They often use heuristics from older technologies to fix their current devices, which is not always the correct way. - Tensions might exist between multiple owners of smart homes, which sometimes makes mitigation of infected IoT devices harder.	- IoT devices and support for security issues should consider multiple users of the devices. - There is a need for a mental model of end-user behaviour when dealing with their IoT environment.	Snowballing	- This research shows the importance of understanding end-user behaviour. Moreover, it seems that some observed behaviour is the cause of the interaction of multiple users of a certain IoT device. This could explain part of the difference between the stated and the observed behaviour of end-users.

Table A2. Identified papers and their relevance.

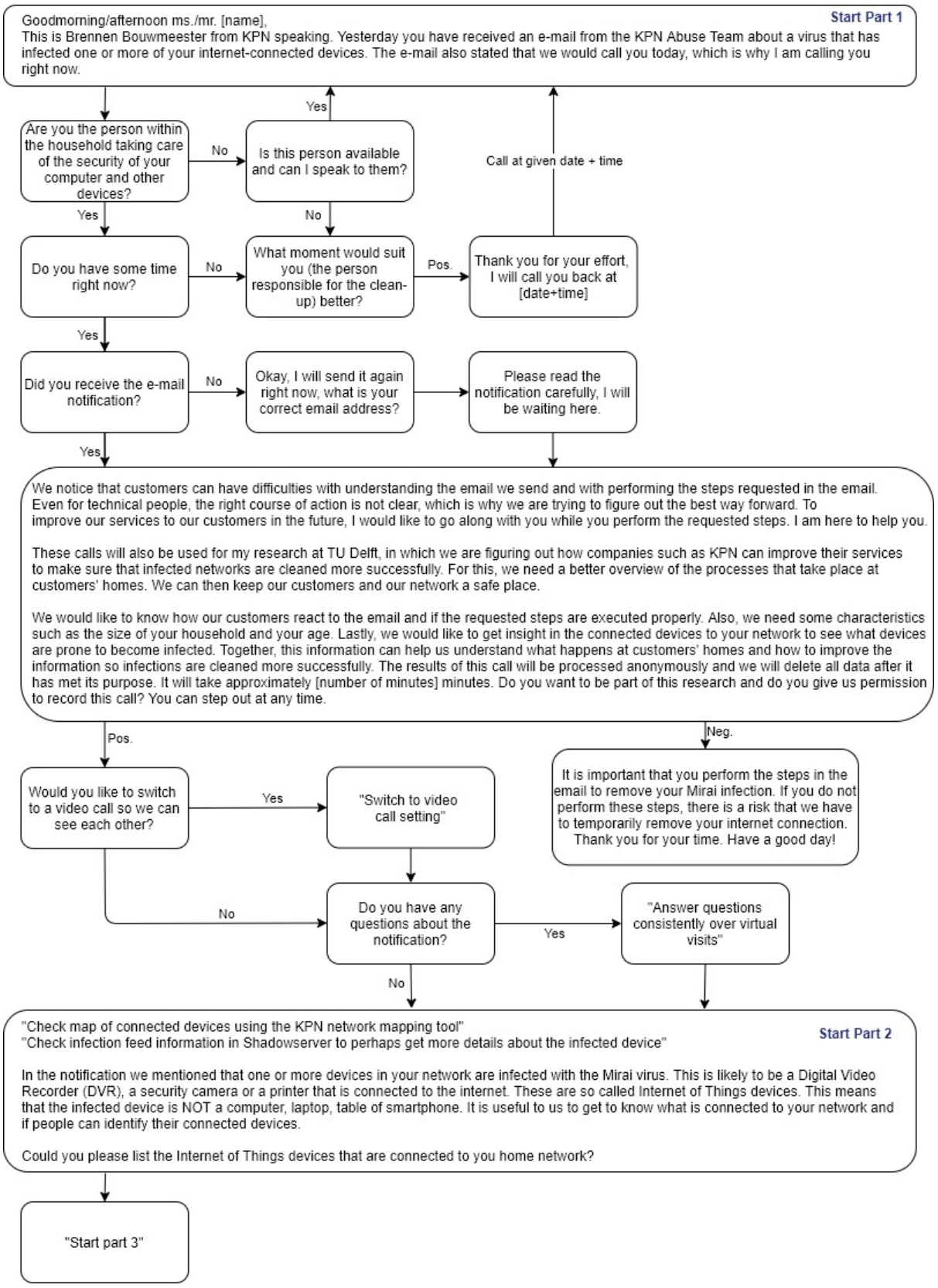
Appendix B. Final virtual visit protocol (after the pilots)

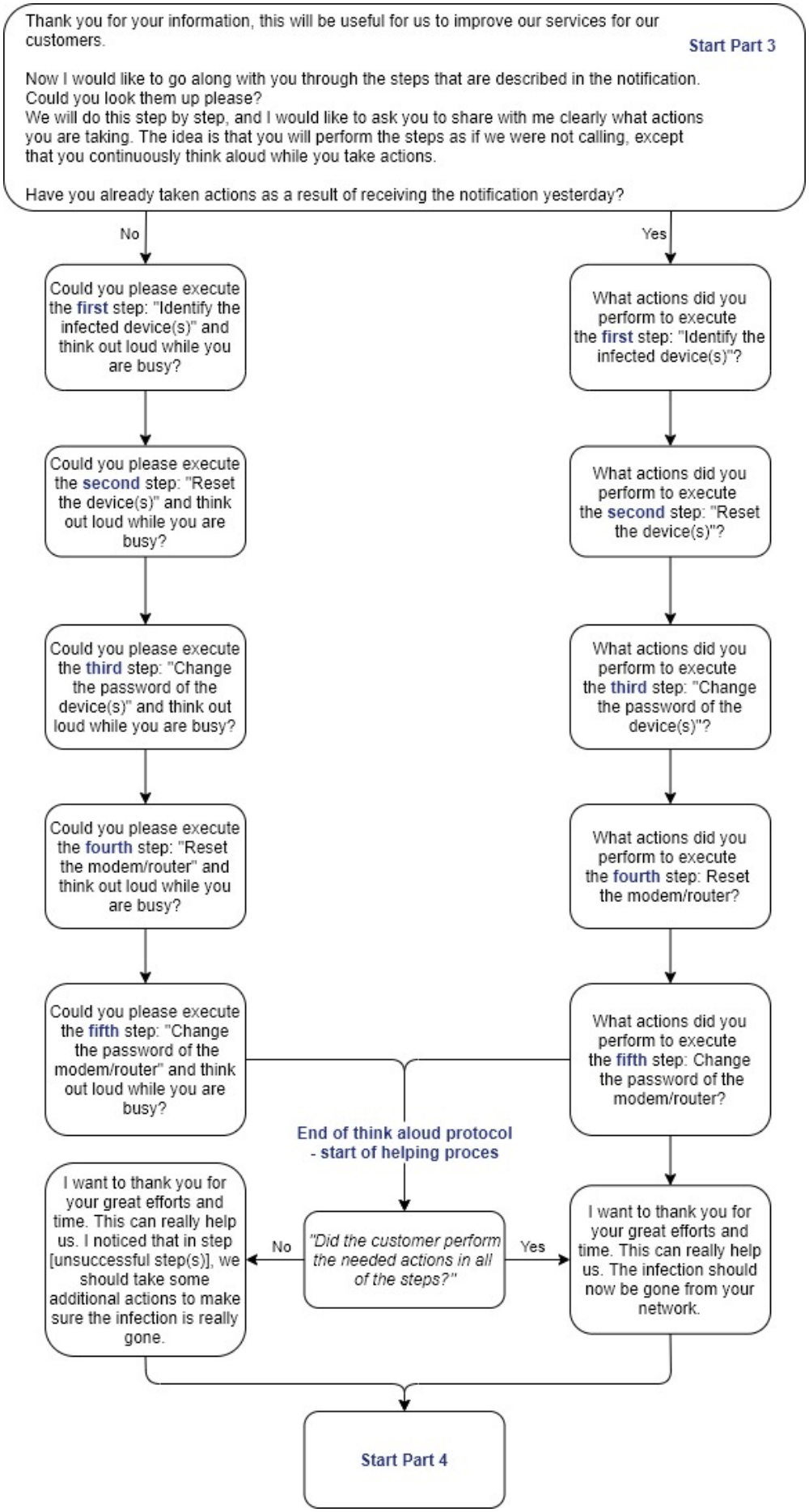


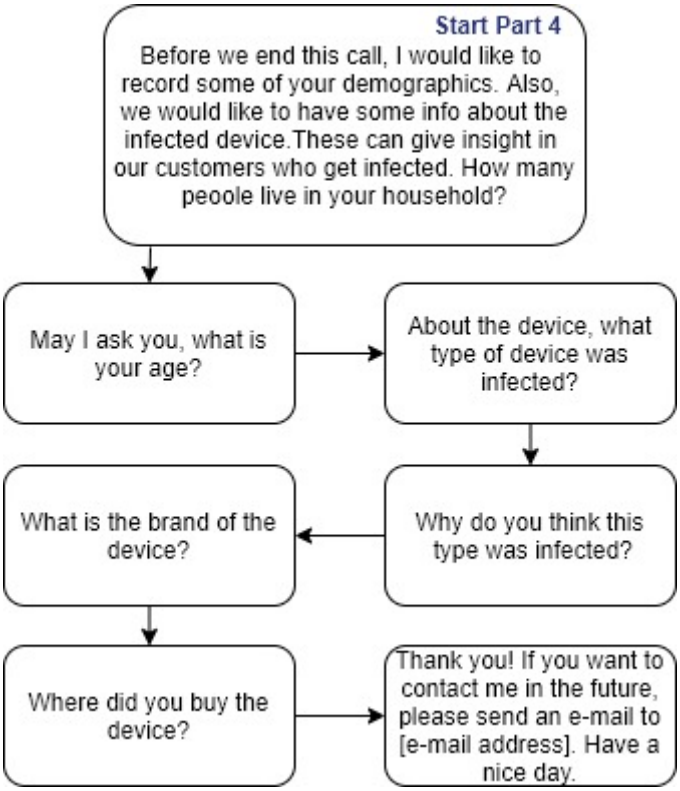

I would like to go along with you through the steps that are described in the notification. Could you look them up please?
We will do this step by step, and I would like to ask you to share with me clearly what actions you are taking. The idea is that you will perform the steps as if we were not calling, except that you continuously think aloud while you take actions.
Start Part 2
Have you already taken actions as a result of receiving the notification yesterday?
No
Yes
Could you please execute the first step: "Identify the infected device(s)" and think out loud while you are busy?
What actions did you perform to execute the first step: "Identify the infected device(s)"?
Could you please execute the second step: "Change the password device(s)" and think out loud while you are busy?
What actions did you perform to execute the second step: "Change the password of the device(s)"?
Could you please execute the third step: "Reset the device(s)" and think out loud while you are busy?
What actions did you perform to execute the third step: "Reset the device(s)"?
Could you please execute the fourth step: "Reset the modem/router" and think out loud while you are busy?
What actions did you perform to execute the fourth step: Reset the modem/router?
Could you please execute the fifth step: "Change the password of the modem/router" and think out loud while you are busy?
What actions did you perform to execute the fifth step: Change the password of the modem/router?
End of think aloud protocol - start of helping proces
I want to thank you for your great efforts and time. This can really help us. I noticed that in step [unsuccessful step(s)], we should take some additional actions to make sure the infection is really gone.
No
"Did the customer perform the needed actions in all of the steps?"
Yes
I want to thank you for your great efforts and time. This can really help us. The infection should now be gone from your network.
Start Part 3


Start Part 3
Before we end this call, I would like to record some of your demographics. Also, we would like to have some info about the infected device.These can give insight in our customers who get infected. How many peoole live in your household?
May I ask you, what is your age?
About the device, what type of device was infected?
Why do you think this type was infected?
What is the brand of the device?
Where did you buy the device?
Thank you! One last thing: Do you give us permission to look into the KPN network mapping tool, which can show us the connected devices to your network? This can help in my research to give an overview of what devices are connected and if certain devices become infected more often than others.
If you want to contact me in the future, please send an e-mail to [e-mail address]. Have a nice day.

Appendix C. Virtual visit protocol during the pilots




Thank you for your information, this will be useful for us to improve our services for our customers.
Start Part 3
Now I would like to go along with you through the steps that are described in the notification. Could you look them up please?
We will do this step by step, and I would like to ask you to share with me clearly what actions you are taking. The idea is that you will perform the steps as if we were not calling, except that you continuously think aloud while you take actions.
Have you already taken actions as a result of receiving the notification yesterday?
No
Yes
Could you please execute the first step: "Identify the infected device(s)" and think out loud while you are busy?
What actions did you perform to execute the first step: "Identify the infected device(s)"?
Could you please execute the second step: "Reset the device(s)" and think out loud while you are busy?
What actions did you perform to execute the second step: "Reset the device(s)"?
Could you please execute the third step: "Change the password of the device(s)" and think out loud while you are busy?
What actions did you perform to execute the third step: "Change the password of the device(s)"?
Could you please execute the fourth step: "Reset the modem/router" and think out loud while you are busy?
What actions did you perform to execute the fourth step: Reset the modem/router?
Could you please execute the fifth step: "Change the password of the modem/router" and think out loud while you are busy?
What actions did you perform to execute the fifth step: Change the password of the modem/router?
End of think aloud protocol - start of helping proces
"Did the customer perform the needed actions in all of the steps?"
No
Yes
I want to thank you for your great efforts and time. This can really help us. I noticed that in step [unsuccessful step(s)], we should take some additional actions to make sure the infection is really gone.
I want to thank you for your great efforts and time. This can really help us. The infection should now be gone from your network.
Start Part 4


Start Part 4
Before we end this call, I would like to record some of your demographics. Also, we would like to have some info about the infected device.These can give insight in our customers who get infected. How many peoole live in your household?
May I ask you, what is your age?
About the device, what type of device was infected?
Why do you think this type was infected?
What is the brand of the device?
Where did you buy the device?
Thank you! If you want to contact me in the future, please send an e-mail to [e-mail address]. Have a nice day.

Appendix D. Randomization protocol for data subjects

```
import random

customers_today = 12
max_participants_per_day = 10


number_of_exceeding_customers = customers_today - max_participants_per_day

possible_participants = list(range(customers_today))

while number_of_exceeding_customers > 0:
    possible_participants.remove(random.choice(possible_participants))
    number_of_exceeding_customers = number_of_exceeding_customers - 1


print(possible_participants)
```

```
[1, 2, 3, 4, 5, 6, 7, 9, 10, 11]
```

Appendix E. Adjusted email notification

**********FOR ENGLISH VERSION SCROLL DOWN **********

Geachte heer/mevrouw [naam],

We hebben een veiligheidsprobleem op uw internetverbinding ontdekt. Dit willen we graag samen met u oplossen. Lees hieronder hoe.

Wat is er aan de hand?
We zien dat een apparaat bij u thuis het mirai-virus bevat. We kunnen alleen niet zien welk apparaat. Waarschijnlijk is het een apparaat zoals een digitale videorecorder, beveiligingscamera of printer die op het internet is aangesloten. Het gaat in elk geval **niet** om uw computer, laptop, tablet of mobiele telefoon. Het virus zorgt ervoor dat een crimineel toegang heeft tot uw apparaat. Dat is onveilig voor u en andere internetgebruikers.

Wij bellen u morgen om het op te lossen
Onze collega Brennen Bouwmeester belt u om samen het virus te verwijderen. We helpen u hier graag bij, omdat veel klanten het moeilijk vinden om zelf te doen. Daarnaast is het een onderdeel van ons onderzoek samen met TU Delft naar dit virus. We stellen u tijdens het gesprek daarom ook een paar vragen om u en andere klanten in de toekomst nog beter te kunnen helpen.

Wilt u zelf het virus verwijderen?
Laat dit dan weten in een antwoord op deze e-mail of tijdens het telefoongesprek. En voer de stappen hieronder zelf uit.

Deze stappen zijn nodig om het virus te verwijderen
Stap 1. Bepaal welke apparaten zijn aangesloten op uw internetverbinding. Het mirai-virus infecteert apparaten die op het internet zijn aangesloten, zoals een digitale videorecorder, beveiligingscamera of printer.

Stap 2. Verander het wachtwoord van deze apparaten. Kies een wachtwoord dat moeilijk te raden is. Weet u uw huidige wachtwoord niet meer? Kijk dan in de handleiding van het apparaat.

Stap 3. Zet de apparaten uit en opnieuw aan. U verwijdert daarmee het mirai-virus uit het geheugen van de apparaten.

Uw apparaten met een internetverbinding zijn nu veilig. Volg de laatste stappen om ook uw modem te beschermen.

Stap 4. Reset uw modem naar de fabrieksinstellingen. Kijk op https://www.kpn.com/service/internet/wifi-en-modems/herstart-reset-experia-box.htm hoe u een Experia Box reset.

Stap 5. Stel het wachtwoord in van uw modem. Kijk op https://www.kpn.com/service/internet/wifi-en-modems/wijzigen/servicetool-wifi-naam-wachtwoord-wijzigen.htm hoe u een nieuw wachtwoord instelt op een Experia Box.

Hebt u toegang op afstand nodig voor een apparaat? Stel dan port forwards in op uw modem. Kijk op https://forum.kpn.com/internet-9/port-forwarding-upnp-watwaarom-en-hoe-322560 hoe u port forwards instelt voor een Experia Box.

Hebt u nog vragen?
Stel deze dan in een antwoord op deze e-mail. Of stel ze tijdens het telefoongesprek.

Met vriendelijke groet,

KPN Abuse Team
abuse@kpn.com

De afdeling Abuse van KPN handelt veiligheidsincidenten af voor KPN. Meer informatie over de afdeling Abuse vindt u op: https://www.kpn.com/abuse

**********ENGLISH VERSION**********

Dear Sir/Madam [name],

We have discovered a security issue on your internet connection. We would like to resolve this issue together with you. The following sections explain how.

What is going on?
We have noticed that one or more internet-connected devices in your home have been infected with the mirai virus. While we cannot exactly detect which one of your connected devices has been infected, it is most likely a device such as a digital video recorder (DVR), security camera or printer connected to the Internet. Devices infected with the Mirai virus are typically **not** computers, laptops, tablets or mobile phones. The infection means that right now criminals have access to your infected device. This is putting you and other internet users at risk.

Tomorrow we will call you to resolve the issue
Our colleague, Mr. Brennen Bouwmeester, will call you within a day to help you remove the virus. We gladly help you with this, as customers find it difficult to resolve the issue on their own. Moreover, the call will be a part of a scientific research that is executed together with TU Delft about the virus. This means we will ask you several questions to be able to help our customers better in the future.

Do you wish to remove the virus on your own?
Please let us know by a reply to this email or during the phone call. After that, please execute the following steps.

These are the steps needed to remove the virus

Step 1. Determine which devices are connected to your Internet connection. The Mirai virus mainly infects Internet connected devices such as a digital video recorder (DVR), security camera or printer connected to the Internet (not computers, laptops, tablets or mobile phones).

Step 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.

Step 3. Restart the Internet connected devices by turning them 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.

Step 4. Reset your modem/router to the factory settings. On https://www.kpn.com/service/internet/wifi-en-modems/herstart-reset-experia-box.htm it is described how you can do this for an Experia Box.

Step 5. Change the password of your modem/router. On https://www.kpn.com/service/internet/wifi-en-modems/wijzigen/servicetool-wifi-naam-wachtwoord-wijzigen.htm it is described how you can do this for an Experia Box.

NOTE: If remote access to a certain device is absolutely necessary, manually define port forwards in your router for the device. On https://forum.kpn.com/internet-9/port-forwarding-upnp-wat-waarom-enhoe-322560 it is described how you can do this for an Experia Box.

Do you have any questions?

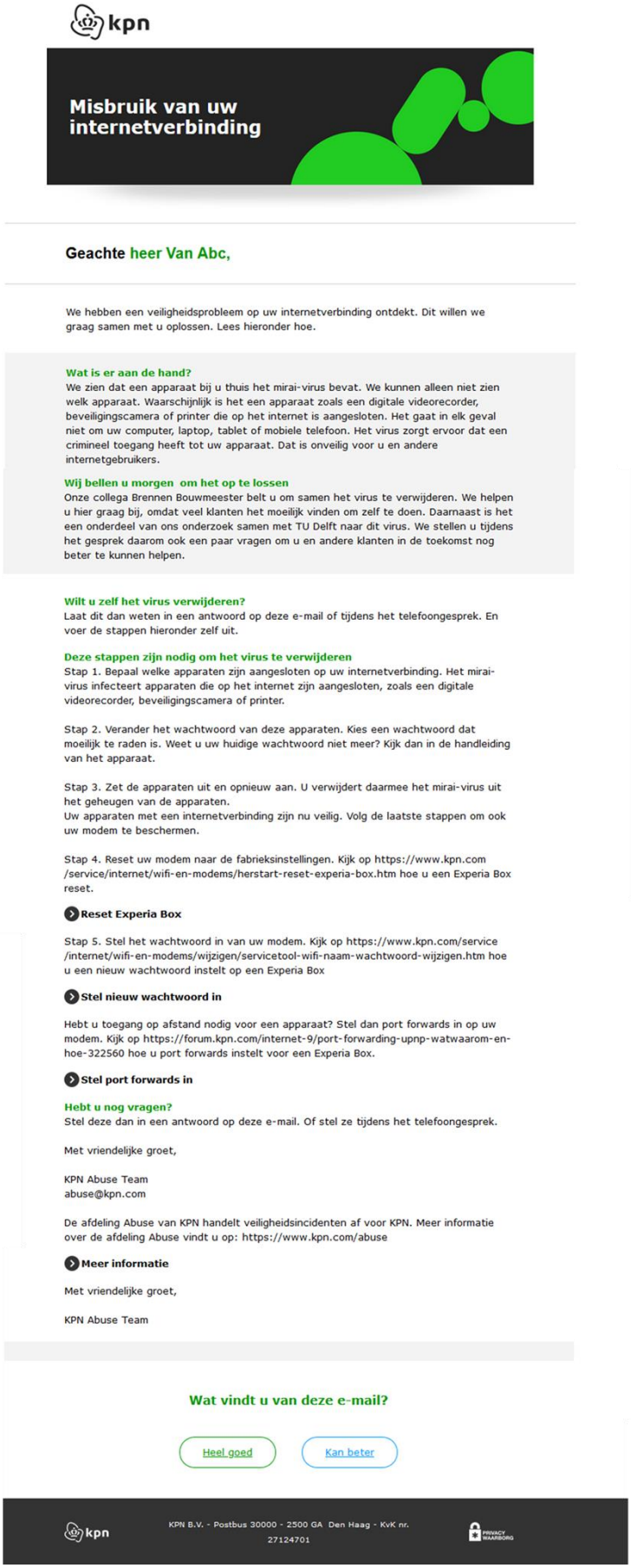
Please ask them in a reply to this email or during the phone call.

Kind regards,

KPN Abuse Team
abuse@kpn.com

The KPN Abuse department deals with security incidents for KPN. You can find more information about the Abuse department on: https://www.kpn.com/abuse

Appendix F. Email notification KPN and Telfort layout



kpn

Misbruik van uw internetverbinding

Geachte heer Van Abc,

We hebben een veiligheidsprobleem op uw internetverbinding ontdekt. Dit willen we graag samen met u oplossen. Lees hieronder hoe.

Wat is er aan de hand?
We zien dat een apparaat bij u thuis het mirai-virus bevat. We kunnen alleen niet zien welk apparaat. Waarschijnlijk is het een apparaat zoals een digitale videorecorder, beveiligingscamera of printer die op het internet is aangesloten. Het gaat in elk geval niet om uw computer, laptop, tablet of mobiele telefoon. Het virus zorgt ervoor dat een crimineel toegang heeft tot uw apparaat. Dat is onveilig voor u en andere internetgebruikers.

Wij bellen u morgen om het op te lossen
Onze collega Brennen Bouwmeester belt u om samen het virus te verwijderen. We helpen u hier graag bij, omdat veel klanten het moeilijk vinden om zelf te doen. Daarnaast is het een onderdeel van ons onderzoek samen met TU Delft naar dit virus. We stellen u tijdens het gesprek daarom ook een paar vragen om u en andere klanten in de toekomst nog beter te kunnen helpen.

Wilt u zelf het virus verwijderen?
Laat dit dan weten in een antwoord op deze e-mail of tijdens het telefoongesprek. En voer de stappen hieronder zelf uit.

Deze stappen zijn nodig om het virus te verwijderen
Stap 1. Bepaal welke apparaten zijn aangesloten op uw internetverbinding. Het mirai-virus infecteert apparaten die op het internet zijn aangesloten, zoals een digitale videorecorder, beveiligingscamera of printer.

Stap 2. Verander het wachtwoord van deze apparaten. Kies een wachtwoord dat moeilijk te raden is. Weet u uw huidige wachtwoord niet meer? Kijk dan in de handleiding van het apparaat.

Stap 3. Zet de apparaten uit en opnieuw aan. U verwijdert daarmee het mirai-virus uit het geheugen van de apparaten.
Uw apparaten met een internetverbinding zijn nu veilig. Volg de laatste stappen om ook uw modem te beschermen.

Stap 4. Reset uw modem naar de fabrieksinstellingen. Kijk op https://www.kpn.com /service/internet/wifi-en-modems/herstart-reset-experia-box.htm hoe u een Experia Box reset.

Reset Experia Box

Stap 5. Stel het wachtwoord in van uw modem. Kijk op https://www.kpn.com/service /internet/wifi-en-modems/wijzigen/servicetool-wifi-naam-wachtwoord-wijzigen.htm hoe u een nieuw wachtwoord instelt op een Experia Box

Stel nieuw wachtwoord in

Hebt u toegang op afstand nodig voor een apparaat? Stel dan port forwards in op uw modem. Kijk op https://forum.kpn.com/internet-9/port-forwarding-upnp-watwaarom-en-hoe-322560 hoe u port forwards instelt voor een Experia Box.

Stel port forwards in

Hebt u nog vragen?
Stel deze dan in een antwoord op deze e-mail. Of stel ze tijdens het telefoongesprek.

Met vriendelijke groet,

KPN Abuse Team
abuse@kpn.com

De afdeling Abuse van KPN handelt veiligheidsincidenten af voor KPN. Meer informatie over de afdeling Abuse vindt u op: https://www.kpn.com/abuse

Meer informatie

Met vriendelijke groet,

KPN Abuse Team

Wat vindt u van deze e-mail?

Heel goed | Kan beter

kpn KPN B.V. - Postbus 30000 - 2500 GA Den Haag - KvK nr. 27124701 PRIVACY WAARBORG