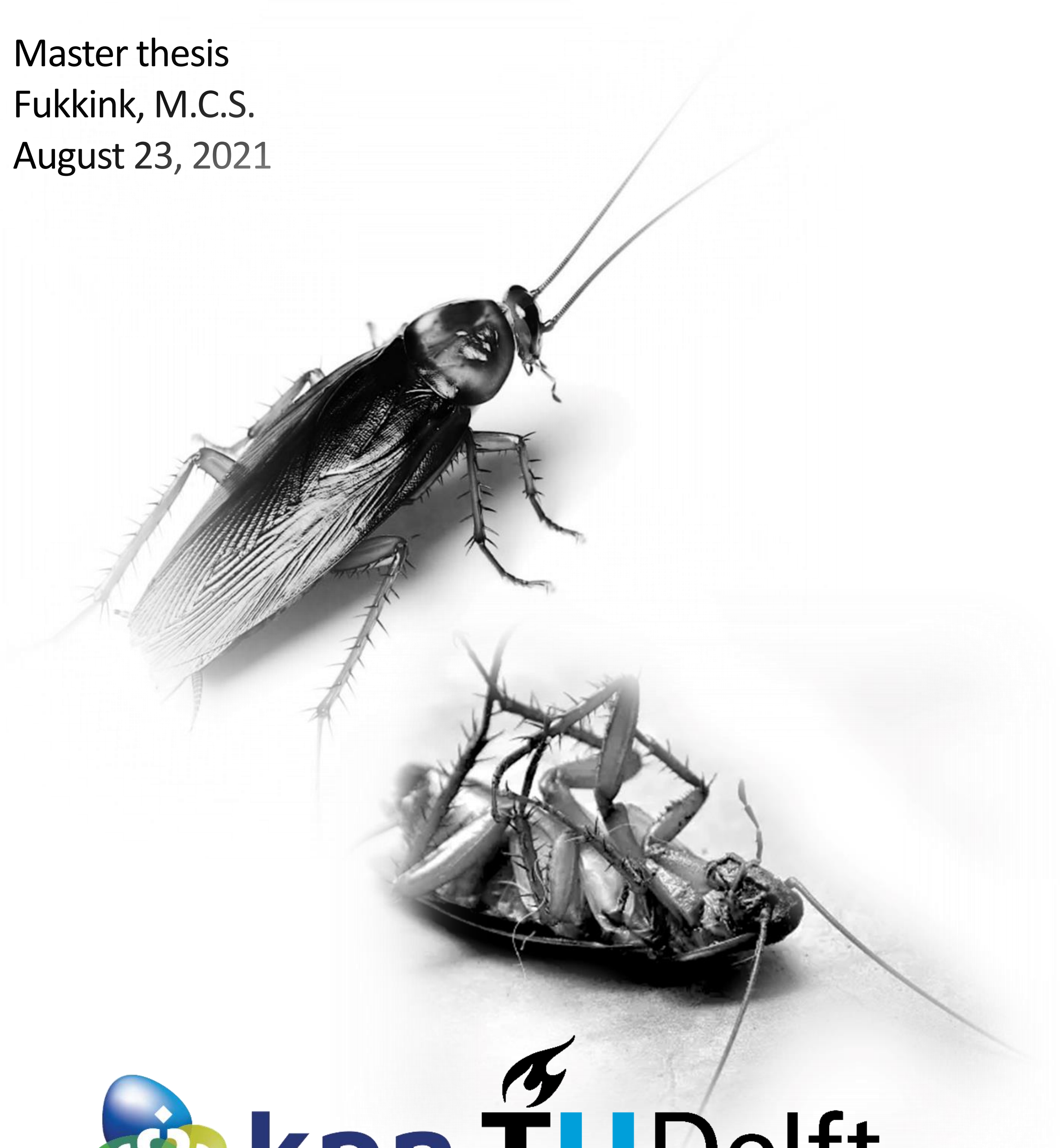
Dear customer, critters are crawling through your precious files

Understanding real-world evidence of QSnatch clean-up results and user experiences after warnings from the ISP

Master thesis
Fukkink, M.C.S.
August 23, 2021



Dear Customer, critters are crawling through your precious files

Understanding real-world evidence of QSnatch clean-up results and user experiences after warnings from the ISP

M.C.S. (Max) Fukkink
Student ID: 4461436
MSc. Management of Technology

to be defended publicly on Monday August 23, 2021

Graduation committee:
Chairperson: Prof M.J.G. (Michel) van Eeten, section Organisation & Governance
First supervisor: Dr ir C.H. (Carlos) Gañán, section Organisation & Governance
Second supervisor: Prof dr ir M.F.W.H.A. (Marijn) Janssen, section Information & Communication Technology
Internal supervisor: Ir E.R. (Elsa) Turcios Rodríguez, section Organisation & Governance
External supervisor: R. (Raymond) Teunissen, Abuse Desk KPN

Acknowledgements

After a long period of researching QSnatch, writing this chapter brings an end to my student days. The last two years of that period have been particularly challenging, but my learning curve has never been steeper. The study has brought me a lot, and my analytical skills have improved enormously. None of this would have been possible without the amazing people around me. This chapter is devoted to all these people, and I would like to thank them now for their unmistakable role in this research. Although it is almost impossible to express in words how grateful I am to you, in the following paragraphs, I will make an attempt.

First of all, I would like to thank Elsa Turcois Rodríguez for her guidance during this research. In weekly meetings, I could ask her all my questions, and she gave sharp analyses of the progress I was making. Without her help, it would not have been possible to achieve this result. You have taught me a lot, and I would like to thank you for the time that you put into this research to help me. I really enjoyed working with you, and I look back on many great conversations and feedback sessions. I also hope that I was able to help you gain new insights for your PhD research. I think that you are already an outstanding researcher and hope to be able to congratulate you soon on completing your PhD.

I would also like to thank KPN for the opportunity to do an internship and do research with them. Without the possibility to analyze their data, it would not have been possible to arrive at the insights that follow from this research. I would like to especially thank Raymond Teunissen and Virgil de Klerk for their guidance as supervisors from KPN. Every week we had useful but enjoyable meetings to discuss results and questions I had. I was welcomed with open arms and had a great time.

Furthermore, I would like to thank Michel van Eeten, Carlos Gañán, and Marijn Janssen for their feedback. Sometimes it was challenging to be almost overwhelmed with feedback. I am really amazed at how analytically you can look at things and how incredibly much you know about the subject. It is truly inspiring to see how good you are in your profession and how much experience you have as researchers. It is indescribable how much I have learned from your comments and how much your insights have improved the quality of this research. I would also like to thank Simon Parkin for his assistance with this research. Although he was not officially part of the committee, he took the time to read the study and provide feedback. I think that the past period has not only shown that TU Delft has excellent researchers to offer but, above all that, its employees provide excellent guidance and are able to offer the right tools where necessary.

Finally, I would like to thank all customers who participated in the interviews. Even though they will probably not read this document, it is of great importance that their contribution is also considered here. It warms my heart that there are people who take the time to participate in research and share their experiences without wanting anything in return. Their selfless contribution to this research will undoubtedly be able to help other KPN customers, and in this way, the world can be made just a little bit more beautiful. Fighting QSnatch is a mission that we can only solve together, and where everyone has to do their part. These customers have more than contributed to that pursuit.

To summarise, I was astonished by the time and effort people were willing to share with me. It cannot be overstated how important the help of others has been to this research. This research's foundation is based on the contribution of people with busy schedules who have taken time off to talk to me, provide me feedback, and contribute to this research.

I sincerely hope you enjoy reading the following chapters of this study.

Max Fukkink

Den Haag, August 17, 2021

Summary

As the IoT is widely deployed in people's homes, adversaries are busy exploiting the vulnerabilities of these devices. One kind of such device is the NAS device made by the company QNAP. Unfortunately, these devices are prone to the QSnatch malware. Unlike previous malware such as Mirai has this nasty habit, it settles deeper into the machine. In this way, the malware gains reboot persistence. Therefore, we consider the malware as persistent IoT malware compared to the non-persistent IoT malware. This affects the clean-up of the virus, as changing the passwords and rebooting the device is not enough to remove the virus. As a result, other steps are needed to get rid of the virus. If we take a look at the NAS device market, we see that the manufacturers of these devices have little incentive to invest a lot in the security of the devices. It is then challenging for the customer to estimate which devices are secure and are mainly tempted by discounts and devices that can be configured quickly.

Then, the ISP is the link in the process that, with the help of the non-profit organisation Shadow Server, can determine which of its customers may be infected with certain malware. Shadow Server uses servers to receive the malicious traffic and forwards the corresponding IP addresses to the ISP. The ISP then knows which customer is dealing with possible infection and can inform them. This also happens for the QSnatch malware. The ISP sends the infected customer a notification informing them about the infection and providing steps to clean their device. These steps are a simplified and Dutch-translated version of the steps provided by QNAP. From that moment on, it is up to the infected customer to take action. Previous research has made a tremendous effort in understanding the efforts of infected customers in remediating the issue and showed that various resources could be used by the ISP to improve the results of this process.

However, the focus of previous research was on non-persistent malware, which, as this research figured, requires different actions to clean an infected device. Therefore, the advent of QSnatch, a persistent malware, raises several questions about the effectiveness of the current process. If customers are less able to remove the malware, they will be more susceptible to data theft and ransomware attacks, among other things. In addition, people with malicious intent have more chances to create a network of bots from individual devices that can then endanger the entire network. However, it is not yet known how effective their efforts are and how customers experience the infection and the clean-up process.

This research was, therefore, specifically interested in how the survivability of QSnatch compares to other malware such as non-persistent malware and Windows malware. That is why the research consisted of plotting Kaplan-Meier curves and calculating logrank tests to determine whether the probability that a virus is still present on a device at a certain time is higher with the QSnatch virus compared to other malware. When the different malware were divided into the groups of persistent malware, non-persistent malware and Windows malware, it became evident that QSnatch does indeed remain longer on the devices. That is bad news for both the customers and the ISP. This shows that the customer is longer susceptible to attacks.

To better understand what exactly happens in the process on the customer side, this research focuses on their experience with the infection and clean-up after they received a notification about the infection from the ISP. Moreover, this research conducted interviews with customers of the ISP with questions based on the COM-B model. This model has proven to be suitable for analysing why certain behaviour occurs. The literature showed that this model was better suited for this than other models. With the insights of the interviews, the research should be able to identify possible improvements to the notification process and provide feedback to the ISP. The interviews were then transcribed and coded to determine the influences of the various factors on the clean-up of the QSnatch malware. The research divided the respondents up into two groups, where one group used the devices for private purposes and the other for at least business purposes or both private and business purposes. It became clear that the business group was relatively more likely to call in help. These had received support from IT professionals and acquaintances with knowledge of IT.

From the analysis, it turned out that the respondents had sufficient physical and psychological capability. Therefore, the study concluded that, in general the respondents had enough capability to carry out the steps. The analysis also showed that the reflective and automatic processes were not barriers preventing the respondents from performing the steps. That is why it seems that there was enough motivation among the respondents to follow the steps. The suggestion that the respondents were motivated was reinforced by the questions they asked at the end and during the interview. Questions about the interview, the research, persistent malware, their security, other steps they could take to become even more secure indicate that even after a 10 to 15-minute interview, some respondents were eager to learn more about the topic. In most cases, there was enough physical opportunity, but social opportunity was not a critical component. That would suggest the respondents had enough opportunity to perform the steps. However, it also became clear that the respondents did not immediately trust the notification. That slowed down the process. Moreover, the steps, in some cases, did not seem to work. That makes it impossible to remove the virus and leads to much irritation among the affected customers. Therefore, the whole process is also dependent on the software provided by QNAP.

Contents

List of Figures

List of Tables

List of Acronyms

ACM	Autoriteit Consument en Markt	13
APT	Advanced Persistent Threat	11
ASN	Autonomous System Number	16
AT	Agentschap Telecom	13
BEREC	Body of European Regulators for Electronic Communication	13
BM	Business Market	13
CC	Civil Code (Burgerlijk wetboek)	14
C-HIP	Communication-Human Information Processing	30
CM	Consumer Market	13
CISA	Cybersecurity and Infrastructure Security Agency	12
CISOs	Chief Information Security Officers	34
CGI	Common Gateway Interface	12
CSR	Cyber Security Raad	15
DDoS	Distributed-Denial-of-Service	11
DKOM	Direct Kernel Object Manipulation	11
DNS	Domain Name System	27
DRDoS	Distributed Reflection Denial of Service	29
EC	European Commission	13
EU	European Union	13
ENISA	European Union agency for Cybersecurity	31
HKC	Healthy Kids Check	38
ICMP	Internet Control Message Protocol	29
IDS	Intrusion Detection System	27

IoT	Internet of Things	11
IRC	Internet Relay Chat	26
ISP	Internet Service Provider	11
IT	Information Technology	13
KDE	Kernel Density Estimation	46
LSTM	Long Short-Term Memory	29
NAS	Network-Attached Storage	11
NCSC	National Cyber Security Center	11
NIST	National Institute of Standards and Technology	16
OPTA	Onafhankelijke Post en Telecommunicatie Autoriteit	13
RAM	Random-Access Memory	26
RAID	Redundant Array of Inexpensive Disks	14
RNN	Recurrent Neural Networks	29
SDN	Software-Defined Networking	28
SME	Small and medium-sized enterprise	13
SOM	Self-Organising Maps	29
SSH	Secure Shell	68
SVM	Support-Vector Machine	28
TAM	Technology Acceptance Model	30
TCP	Transmission Control Protocol	29
UDP	User Diagram Protocol	29
UPnP	Universal Plug and Play	27

1 Introduction

In this section, we will get a better understanding of the QSnatch malware, the users of Network-Attached Storage (NAS) devices, customers of the Internet Service Provider (ISP), and possible incentives to protect NAS devices. In this way, the reader gets a complete picture of the different interests of different actors contributing to the NAS users' security. Thereafter, it addresses the problem statement, research questions, and relevance. It will also set the focus and scope of the research. The section closes by discussing the research's structure.

1.1 Background

COVID-19 is expected to accelerate the investment in the Internet of Things Internet of Things (IoT) [2]. Dorsemaine et al. [3] consider the IoT as *'group of infrastructures interconnecting connected objects and allowing their management, data mining, and the access to the data they generate'*. The IoT has a diverse set of applications and is applicable in sectors as smart cities [4], smart homes, connected cars, and e-health [5]. Gartner estimates 25 billion of such devices will be connected by 2021 [6]. Also, the spending on IoT-related software and hardware is expected to grow from 726 billion dollars in 2019 to 1.1 trillion dollars in 2023 [2]. As a result, society is increasingly installing the IoT in their homes and businesses. IoT devices are, for example, smart home control, smart TVs, cameras, or virtual assistants. A device is considered part of the IoT if it can connect to a network and share data, perceive information from its environment, and perform tasks without user interference [7]. Fundamental challenges of the IoT are data protection and privacy [8]. Much effort is put into the research of those challenges [9, 10, 11, 12, 13, 14, 15, 16] and to put those in legal context [17, 18]. In the study of Atlam and Wills [19], the researchers combine knowledge on IoT privacy, safety, and ethics. Likewise, Sholla et al. [20] consider the ethical consequences of adopting IoT on a society-wide scale. But, taking a moral and legal perspective is not the only concern for the IoT. The security of the IoT has recently come under threat.

Alarming findings are those of Eresheim et al. [21]. Those researchers pointed out that IoT is increasingly becoming a target of Advanced Persistent Threat (APT) actors. Their study specifically focused on Direct Kernel Object Manipulation (DKOM), an obfuscation technique with malicious intent to hide running processes from system monitoring tools [21]. It turned out that Windows 10 IoT Core was vulnerable to this technique. In their scenario, the attacker already has access to the IoT device and wants to achieve long term access to the machine to misuse it for subsequent attacks or other network nodes [21]. Withal, the most significant threat of the IoT are IoT botnets [22]. Those can serve purposes ranging from Distributed-Denial-of-Service (DDoS) attacks to spam and advertisement fraud [23]. In most cases, users are on their own when it comes to the clean-up of their devices. Fortunately, some ISPs are kind enough to notify the user when they find malicious traffic on their network. However, whether users are able to do so is questionable and dependent on human behaviour. Moreover, malware is currently gaining persistence which makes it even harder to remove them from devices [24]. In this research, in collaboration with an ISP, a look behind the scenes will be taken in order to hopefully offer resistance to the worrying developments. A further explanation of this collaboration will be given later.

The next section will first delve deeper into a new danger. This chapter then discusses the different actors involved in the problem. These are connected by being bound by laws and regulations, market participation and the use of IoT devices. Their precise role will become apparent in the upcoming sections. These insights are concluded with an analysis of the actors during a stakeholder analysis.

1.1.1 A new player

An example of new persistent malware is QSnatch. Hackers have infected thousands of IoT devices with the malware QSnatch [25]. The affected devices are NAS devices from the Taiwanese producer QNAP. A NAS is a central storage device that is connected to a local network. This allows the owner of the NAS to save and open files from any computer, laptop, media player, and other devices. Moreover, Qsnatch is next in line of other malware targeting NAS devices. In 2019, there was a ransomware aiming at devices from Synology [26], ransomware eCh0raix [27], and ransomware Muhstik [28]. All of those malware used brute-forcing to track down weak passwords and the latter two both targeted QNAP devices. In 2020, a new type of ransomware was discovered called Agelocker [29]. Then again, at the beginning of 2021, QNAP warns its customers that another malware is targeting its devices to mine cryptocurrency's [30]. In August 2021, Synology warned its customers for brute-force attacks targeted on the Synology devices by the StealthWorker-botnet [31, 32]. In sum, there is much malware targeting the QNAP's NAS devices and NAS devices in general.

It is not yet clear how QSnatch spreads, but it burrows into the firmware to gain reboot persistence [25], takes full control over the device, and blocks future updates [33]. It is still unclear if hackers developed QSnatch to carry out DDoS attacks, perform cryptocurrency mining, or as a backdoor to sensitive files or host malware payloads [25]. The National Cyber Security Center (NCSC) of Finland found out that QSnatch can connect to a Command and Control (C&C) server, download, and run modules. Therefore, the NCSC-FI's theory is that the hackers are currently building the botnet and plan to activate other modules in the future. According to the US Cybersecurity and Infrastructure Security Agency (CISA) and the NCSC-UK, the number of reported infections grew from 7,000 devices in October 2019 to more than 62,000 in June 2020 [33]. Of those devices, approximately 46% are located in Western Europe. The agencies divided the infections up into two campaigns. The first took place from 2014 to 2017 and the second from 2018 until 2019 according to the agencies [33].

The latest version of QSnatch is particularly dangerous and has a variety of functions [33]:

- A Common Gateway Interface (CGI) password logger to install a fake version of the device admin login page, logging successful authentications and passing them to the legitimate login page;
- A Credential scraper;
- A SSH backdoor enabling the hacker to execute arbitrary code on a device;
- An Exfiltration - With this function, QSnatch can steal a predetermined list of files, which includes system configurations and log files. These are encrypted with the actor's public key and sent to their infrastructure over HTTPS; and
- A Webshell functionality for remote access.

Still, in 2020 it was a mystery how QSnatch initially infected devices [33]. Although the second campaign is down, infections remain active on the Internet. Therefore, the CISA and the NCSC-UK advised users to patch the devices, because otherwise the hacker would keep a backdoor into the devices and enable them to get their hands on possibly sensitive data [33]. Fortunately, QNAP, the NAS developer, has also been busy. On October 25, 2019, they got a notification of NCSC of Finland and Germany that they found QSnatch on QNAP devices. On November 1 and 2, 2019, QNAP released an update for their Malware Remover and published an updated security advisory to tackle QSnatch. Thereafter, the company released press releases, updated the security advisory, and emailed possibly affected users [34]. QNAP's security advisory is on their website [a]. For the more nontechnical users, different Youtube accounts provide help. NASCompares [b] explains how to protect a QNAP NAS, Mike Faucher [c] elaborates on how to remove QSnatch and clean a device, and QNAP UK [d] provides a solid walk-through.

1.1.2 Remediation by users

The security advice brings this research to a refreshing thought that IoT users can clean their own devices. The paper of Cetin et al. [35] already found that users can succeed even when they are operating with wrong mental models. Those researchers discovered that quarantining and notifying the users by an ISP resulted in remediation in 92% of the cases. Essential to mention here is that users removed the Mirai virus which is a non-persistent malware.

Quarantining infected customers, or putting them in a so-called wallet garden, is a costly measure. In such a case, the ISP restricts the access of the customer to only white-listed sites [36]. The benefits of a wallet garden are two-fold: it protects the customer from further harm and avoids that the customer's device harms others on the network [36]. Moreover, it is also an effective tool to inform the customer about the malware. Instead of forwarding to the Internet, it directs the customer to a website with information about the malware and how to clean the device [36]. Cetin et al. [36] pointed out that the period in the wallet garden depends on the malware detection process, the infection notification and quarantining process, and the release process. The customers could leave the quarantine by providing proof of the clean-up or they just have to report that they have done so [36]. In both cases, it does not guarantee that the problem is solved. The customer could not have performed steps as some wallet gardens have a self-release option, an expiration period, or ISP staff members releasing the customer; did the steps incompletely; or the device could get reinfected [36].

Again, the research of Cetin et al. [36] focused on Mirai, a non-persistent malware. Currently, QNAP owners have problems with persistent malware. The difference between both is that non-persistent malware is stored in the internal memory of the device. By simply changing the passwords and rebooting the device, the malware should be removed. On the other hand, persistent malware burrows deep into the firmware and becomes resistant to reboots. Therefore, there are other steps required to clean the device. That raises all sorts of questions such as: Could customers also clean their devices from persistent malware?

1.1.3 Users and customers

Customers of the involved ISP are generally Dutch citizens. The ISP can deliver services like mobile, television, and Internet and the ISP's customers can purchase these services. This research is mainly interested in customers that have

[a] https://www.qnap.com/en/security-advisory/nas-201911-01

[b] NASCompares: https://www.youtube.com/watch?v=o5uWH_utEWs

[c] Mike Faucher: https://www.youtube.com/watch?v=C4tbieli6f0

[d] QNAP UK: https://www.youtube.com/watch?v=qoFnhVGPXDM

an Internet subscription at the involved ISP. On the website of the ISP, it is possible to choose from different Internet speeds such as 50, 100, 200, and 1000 Mbit/s. This study considers the customers of the ISP as all persons who have a subscription with the ISP.

Moreover, the ISP also offers Internet service for businesses. The website advertises special subscriptions for Small and medium-sized enterprise (SME)s with services like speeds up to 1Gbit/s, a 4G backup, a fixed IP address, software integration with Microsoft 365, and an optional pin connection. In its systems, the ISP differentiates between Consumer Market (CM) and Business Market (BM) customers. According to employees of the ISP some small companies tend to arrange a CM subscription when the conditions and costs are more favourable compared to a BM subscription. As a result, some customers may appear to be using their subscription for a private network when it is actually being used in a corporate environment.

A subset of the ISP's customer group is the owners of QNAP devices. This research distinguishes them as the users of QNAP devices or, in short, users. It, therefore, concerns the group of customers of the ISP who have a QNAP device. To all subscriptions that the ISP offers applies that no special Information Technology (IT) skills are required to arrange them. After all, the ISP wants to be able to provide these subscriptions to as many people in the Netherlands as possible. In addition, connecting and configuring a QNAP device is also straightforward. It is therefore not a requirement that a user is an IT expert in any way. Users of QNAP devices could just be 'normal' people without a lot of IT knowledge. On the other hand, one could argue that someone with a NAS device is likely to have at least some experience. The reason for this is that an average user does not even know what NAS devices are or even if they exist.

1.1.4 Laws, Regulations, and Regulators

The involved ISP provides Internet connection to its customers in the Netherlands. It thus operates in the Dutch telecom market and is bound by laws and regulations. The market is restricted by European and Dutch laws. The legislators of the European Union (EU) established the Body of European Regulators for Electronic Communication (BEREC) in 2009 [37]. In 2016, the BEREC proposed the Guidelines on the Implementation by National Regulators of European net Neutrality Rules, and these guidelines have been in effect ever since [38]. The BEREC has multiple functions such as fostering independent and consistent regulation, contributing to the development and functioning of electronic communications networks and services, and assisting the European Commission (EC) and the national regulatory authorities in implementing the EU regulatory framework for electronic communications [39].

In the Netherlands, the national regulatory authorities are the Autoriteit Consument en Markt (ACM) (Authority for Consumers and Markets) and Agentschap Telecom (AT) (Telecom Agency). The monitoring of the telecommunications market started with Onafhankelijke Post en Telecommunicatie Autoriteit (OPTA) (Independent Post and Telecommunications Authority). Between 1997 and 2013, the OPTA was a governmental agency enforcing Dutch law on the post, telecommunication, and cable TV services. In 2013, the agency's tasks were taken over by the ACM, and the OPTA ceased to exist. Since then, the ACM enforces the Telecom Act, which promotes opportunities and choices for businesses and consumers; opportunities for innovation, new products, services and companies, and choices for consumers, so that consumers really have something to choose [40]. Originally, there was only one provider in this market. ACM opened up the market by ensuring that other companies could use the network of the initial provider. Today, consumers can choose from several providers. AT, on the other hand, is the enforcement organisation when it comes to licensing radio amateurs, auctioning radio frequencies, and checking compliance with the Telecom Act [41]. The two organisations have a collaboration protocol in place to work together while enforcing the Telecom Act [42].

Within the jurisdiction of the Netherlands, an ISP would also have to comply with laws and regulations, which are the Telecommunicatiewet (Telecommunications Act) [43] the Wet telecommunicatievoorzieningen BES (Telecommunications Facilities Act) [44], General Administrative Measures for the Telecom Market, Ministerial Regulations for the Telecom Market, ACM Policy Rules for the Telecom Market, and Policy Rules of the Ministry of Economic Affairs and Climate [45]. Telecommunications Facilities Act is less vital for this research as it comprises the facilitation of telecommunication to the islands Bonaire, Sint Eustatius and Saba, which are special municipalities within the Kingdom of the Netherlands since 2010. The Telecommunications Act spikes the interest as Article 11.3 second paragraph of this Act is it includes the policy rules on the obligation to provide information about Internet security [43]. According to this article, the ISP should at least provide information on several risks, including botnets, zombies, spyware, trojans, and 'other malware'; and information about means to cope with these risks like a firewall, a virus scanner, the use of legal software, the activation of automatic updates of software, and appropriate caution when opening offered files. The ISP should provide this information on a dedicated page on its website and easy-to-find references. This concerns information provision regardless of whether the customer is infected with malware.

So far, there seems to be no legal obligation for an ISP to notify customers when they get infected. However, the liability of ISPs is usually linked to the real possibility of not taking too far-reaching precautions [46]. This seems to be breaking with the old interpretation of ISP indemnity, whereby ISPs were not liable as long as they were passive. ISPs are exempt from non-contractual liability insofar as they remain passive with regard to the content they transmit under Article

12-15 E-commerce Directive 2000/31, implemented in Article 6:196C Civil Code (Burgerlijk wetboek) (CC) [47, 48]. ISPs actually run the risk of liability if they actively search for malicious content or suspicious communications. In the US, this is called the 'chilling effect'. More extensive liability does not seem a good idea [47]. Most ISPs have an interest in combating cybercrime themselves but are confronted with regulations that make this more difficult. If an ISP would be actively monitoring and blocking internet traffic, then that would easily be in conflict with rules on net neutrality, Internet access, and privacy.

The parties involved have spoken of a 'zorgplicht internetveiligheid' (duty of care internet security). This duty would mean that the ISPs would have been obligated by law to take care of its network and monitor for abuse, so the predecessor of ACM decided in 2007 not to implement policy rules for this duty [49]. The emphasis was put on the provision of information to consumers about internet safety, as discussed in the previous paragraph. The decision would have been revised if the ISPs themselves did not come up with a robust system to increase consumer safety on the Internet. With the knowledge of today, we know that this has not happened. At the same time, Dutch ISPs felt the responsibility to safeguard their networks. Together, they made a gentlemen's agreement to mitigate illegal activities on their networks, such as remediating botnets and malware. The involved ISPs complied with this agreement by notifying customers when they get infected and possibly shut off a customer if he does not remediate the infection.

The involved ISP does cover the legal ground for a potential shut down of the customer's network if he engages in illegal practices. The ISP's general terms and conditions for fixed and mobile telecommunication services describes various unwanted activities such as but not limited to sending spam, spreading (computer viruses) or other files that can damage the (proper) functioning of our software of that of others, threatening people, or hacking [50]. Someone engaging in these kinds of activities does not comply with the ISP's code of conduct, requiring the ISP to take measures such as temporarily stopping the service or terminating the service agreement [50]. Combining the insights of this section would bring us to the formal chart in figure 1. It is evident that the rules and regulations are rather straightforward at this point. Moreover, an essential side note here is that the ISP is not obligated by law to notify and shut down customers when they are infected, as discussed in the previous paragraphs.

On the other hand, a customer can have complaints about the ISP. If that does not work, they recommend writing an official letter, email, or a notice of objection [51]. If the customer and the ISP cannot reach a solution, the customer can choose to take further steps. In that case, ACM recommends that the customer also reports the complaint to them. In this way, they can keep track of which companies do not comply with the rules and take any measures. The customer will then have to submit it to the Disputes Committee or he can choose to go to court [52].

Legislation and regulations are currently somewhat limited with regard to the requirements that NAS producers and NAS retailers must meet. A NAS user is entitled to sufficient information about the product [53]. The ACM investigates whether the NAS providers and retailers deliver sufficient information. The TA is planning to set legal digital security requirements to ensure that unsafe products do not come onto the Dutch market, but it is not that far yet. If a customer makes a report to ACM and it turns out that a NAS producer or retailer provides insufficient information, the ACM can then determine how great the total damage is for consumers, how great the social importance is, and whether ACM has a solution for the problem at hand [54]. This could include providing additional information on the subject, additionally checking a sector, conducting research into a company, forcing a company to find a solution for something that goes wrong, and in the event of a violation, fine a company [54].

1.1.5 Market developments

QNAP is not the only NAS device producer in the market. Companies like Synology, Inc., ioSafe, Asustor Inc., Promise, WD, Seagate Technology, Buffalo Americas, Inc., Western Digital Corporation, and TerraMaster also produce NAS devices. Fortunately for these companies, the NAS market is expected to grow in the upcoming years [55, 56, 57]. This market is highly fragmented and competitive in nature [56]. In the Netherlands, consumers can buy NAS devices at stores such as Coolblue, MediaMarkt, and Alternate. According to PCM, the most popular NAS brands in the Netherlands are Synology, QNAP, Netgear, Asustor, Western Digital, Thecus, and Drobo [58].

Most of those companies use a Linux-based operating system that a customer would install on his NAS [58]. On top of the operating system come extensions or packages. These are separate modules that the customer can install to add functionalities. For example, a customer could decide to install a media centre to stream movies or manage his photo collection. The number and quality of the extensions differ per brand and per model and is constantly changing [58]. NAS devices come in designs for a different number of drives [59]. There are variants for one, two, and more discs. The more drive bays a NAS offers, the more expensive the device is, and the more a customer could store on it. Moreover, a NAS with two or more disks offers the possibility to protect the data against failure of a disk through Redundant Array of Inexpensive Disks (RAID) [59]. To illustrate, a customer could decide to buy a dual-bay NAS and buy two drives with 2TB of storage. Then, there is enough storage for 4 TB of data. The customer could opt for the protection of RAID1, which means that 2 TB of data would fit on the NAS.

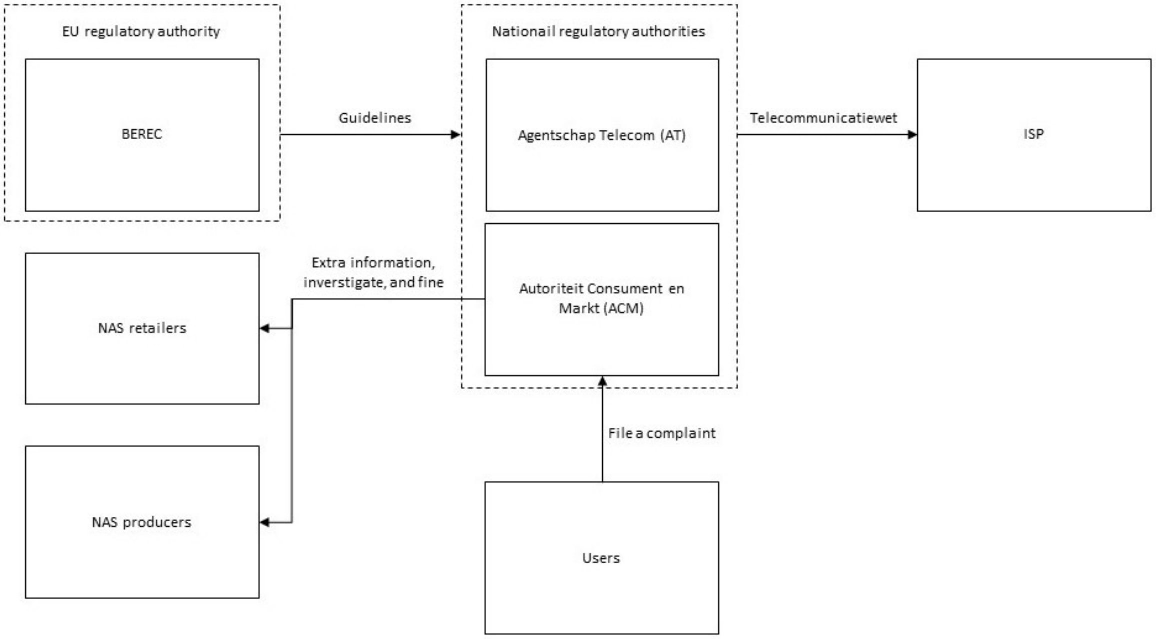


Figure 1: The formal chart

In a broader sense of the IoT, security is not a business driver; A customer does not pay more for a product that claims to be safe [60]. To customers, security is an invisible property. A customer usually does not select on it. In addition, there is an information asymmetry, as a result of which a customer is often unable to include these factors in a choice, and the information that would be required for this is not available [60]. Therefore, consumers simply assume security. Hardware is often a commodity, so the price is a key driver. At the same time, the market is flooded with cheap hardware from low-wage countries. As developing IoT devices often involve combining existing technologies, products can be produced relatively quickly [60].

On the other hand, the Cyber Security Raad (CSR) (Cyber Security Council) of the Netherlands urged businesses to do more to gain digital security [61]. As such, the NAS producers are obligated to provide information and warn the customers if security-related support for a product is to be withdrawn. The organisation also summarised many obligations for companies selling products with an ICT application, including but not limited to [61]:

- take cybersecurity into account at the development stage (security-by-design)
- encourage the development of properly protected products and services through organisational measures
- comply with relevant security standards, codes of conduct, or certifications
- use a system to ensure that security patches are installed by users
- encourage users to report security incidents and investigate them.
- offer users support in the event of a security incident
- warn users if the level of cybersecurity is inadequate

Some argue in favour of a quality mark for IoT. The company Computest analysed the issue and came up with key aspects for it [60]:

- the manufacturer communicates a clear end-of-life date to potential buyers,
- the product comes with a safe standard configuration,
- security updates are applied automatically (without user intervention),
- and the data collected and processed by the device is never sent over a network unencrypted.

In 2020, the AT commissioned research into the requirements for IoT to provide better protection for consumers against cyberattacks [62]. That research described attack scenarios and vital security issues in the context of consumer IoT.

After a literature review and the evaluation of more than four hundred measures, the study summarised the eight best measures and considered the minimum requirements for the IoT [62]. According to the researchers, the requirements are easy to implement, easy to test, clear, and significantly increase the security of IoT products. Those requirements are [62]:

- all passwords must conform to the standard National Institute of Standards and Technology (NIST) SP800-63b Digital Identity Guidelines [63];
- after initial configuration, passwords must be unique for each device, or specified by the user;
- network access to a device in the functional state must be possible only after authentication;
- the device may only provide ports and links that are necessary for normal and intended functionality;
- all network traffic must be encrypted and authenticated using common encryption protocols, such as TLS;
- manufacturers must be able to initiate an update of the software in devices;
- the device must verify the integrity and authenticity of software before installing it;
- and the manufacturer must provide clear information about the end user's responsibilities for using the device safely.

Unfortunately, the producers of NAS do not seem to implement these requirements so far. Therefore, the customers are vulnerable to vicious types of malware infecting their devices. As a result of economic drivers, it does not look like producers will implement the requirements soon, and customers do not seem to care enough about securing their devices. This makes an ISP the only actor left to step in and act as a guardian.

1.1.6 The Abuse Team, Shadowserver, and the clean-up procedure

As mentioned before, this research will collaborate with an ISP. This collaboration was formed in such a way that the principal investigator of this study also completed an internship at the ISP. The involved ISP is one of the ISPs in the Netherlands that inform their customers with notifications. It has a special department that focuses on abuse in its network. This department is called the Abuse Team. The Abuse Team gets alerts indicating which IP addresses in their network are generating traffic affiliated with different types of malware including but not limited to QSnatch. These kinds of abuse on the network can be malware, but the team also gets information about abuse on the network in general, so it can also involve hosting phishing sites, for example.

The Abuse team's system receives this data from The Shadowserver Foundation. Shadowserver is a nonprofit security organisation that runs a sinkhole infrastructure and large-scale sensor networks of honeypots and honey clients [64]. In their pursuit, the organisation collects threat data of malware traffic and shares that with other organisations. The organisation informs ISPs about malicious traffic on their networks. All the data containing the ISP's Autonomous System Number (ASN), Shadowserver, will be sent to the ISP. The ISP receives data about malicious activity every day. This information is then automatically processed in the abuse system. The different forms of abuse are then divided into groups and the system applies different procedures based on scripts that are made to inform customers or take steps to prevent the abuse.

The team sends email notifications to infected customers in which they are urged to solve the problem themselves. However, some customers tend to reply to those notifications and try to get in touch with the team. Unfortunately, the number of the team's employees is limited and, therefore, only a maximum number of notifications are sent per day. The system makes a choice to give priority to different types of abuse over other forms of abuse. So it could be that customers are infected, but do not receive a notification. Those customers are probably still infected the next day or days after and probably then get a notification. But, it could also be that they take measures on their own to solve the problem, that they turn off the device, or that something else happens so that the device no longer generates malicious traffic. In this way, it could be that customers do not receive a notification, but the problem is solved. But as long as the device is infected with malware of which notifications are sent, is turned on, connected to the Internet, and most importantly, keeps generating malicious traffic, at some point the customer will probably receive a notification from the ISP. So customers who appear to be removing the malware without a notification will generally do so quickly.

In 2019, Shadowserver started reporting QSnatch infections to the ISP. The Abuse Team responded by notifying infected customers. Customers would get an email that explains that the ISP found traffic related to the QSnatch malware coming from the customer's IP address and how customers could remove the malware. These steps include:

- Go to the website: qnap.com/en-en/download
- Under "1 - Product type", select the option "NAS / Expansion" and select the number of slots present on the right.

- Under "3-Model", select the type of NAS you are using.
- Under the "Operating System" tab, select the most recent version and download it via the "Europe" button.
- Open the NAS on your PC or Mac and choose -> Firmware update -> Manual update.
- Browse to the downloaded file and update the firmware / operating system.
- Go to APP Center and choose "Malware Remover" and download it on your PC or Mac.
- Click on "manual update" in App center, browse to the download file and update the Malware Remover.
- Run a scan with the Malware remover.

Appendix A presents the notification mail in both Dutch and the translation to English. On Thursday, March 11, 2021, two employees of the ISP, together with the researcher of this study, carried out the steps as an experiment. It took around 40 minutes to perform the steps. This is the total time it took, including time to wait for the downloading, installation, and scanning, which basically run automatically without requiring actions from the user. It became clear that in the step *'Open the NAS on your PC or Mac and choose -> Firmware update -> Manual update'*, it should also state that the customer should open the Control Panel. The type of the QNAP device was a TS-231. Essential to note here is that we performed the steps on a device that was clean. It may well be that it becomes more difficult to perform the steps once the device is infected. The malware could possibly hinder the execution of the steps.

As the method of the ISP differs from the steps provided by QNAP, the method to remove the malware provided in the notification mail are considered the ISP's clean-up steps or the ISP's clean-up procedure. The broader clean-up process also includes the notification, the interaction between the ISP and customers, and the whole activity of identifying, contacting, and informing the customer. Since this research will analyse the behaviour of customers notified with the clean-up procedure of the involved ISP, the remainder of this research will use the clean-up procedure for the involved ISP's clean-up steps.

When taking a closer look at the clean-up procedure of the ISP in appendix A.1 and A.2, we see that the user would need to fill in information about their device and download a firmware update from the website of QNAP. This means that the customer would need to have a computer and an Internet connection available to perform the steps. On that computer, the customer would then have to open the software of the QNAP device and update the device with the downloaded firmware. Thereafter, the customer would have to download the application 'Malware Remover'. Therefore, the steps also require some basic computer handling and skills. At the same time, it would also be impossible for the customer to perform the steps if the site of QNAP is down. Fortunately, a large proportion of these customers have somehow figured out how to stop showing up in the reports of Shadowserver. There are some explanations for this phenomenon. To illustrate, it could be that the customer followed the steps provided by QNAP or the ISP or they just turned off the device, never to be used again. Another explanation would be that the customer followed another procedure and looked on the Internet for help. Maybe, the customer was able to contact an IT professional and asked to perform the required steps.

But, this phenomenon does not just have to be due to changes in user behaviour. It could also be that the attacker makes changes. For example, an attacker may choose to disconnect the botnet or parts of the botnet because he suspects that he will be discovered. Another reason could be that he has looked at the devices and found that they are not interesting nor contain important information. An attacker might be interested in specific information from large companies and hope to access that information through QNAP devices. In that case, it is not interesting to get information from private users. In fact, if you are looking for very valuable information from large companies, an attacker can better ensure that the malware is apparently not dangerous and, above all, stays under the radar for a long time.

Moreover, it becomes evident that the involved ISP decided not to include all the steps that QNAP prescribes. Appendix A.3 shows the steps of QNAP. As described above, the steps of the ISP mainly comprise updating the firmware, installing the anti-malware software 'Malware Remover', and letting the software scan the device for malware. It looks like the ISP has taken a subset of the steps that focus mainly on removing the malware. Updating the firmware and installing the Malware Remover mainly focuses on removing the malware and not yet specifically ensuring that the malware cannot infect the device again. It could be that the updates remedy vulnerabilities in the system but if the user is still using weak passwords, then the device could still be vulnerable. In the end, it is not yet sure how the devices get infected so the steps of the ISP at least cover the most vital steps. In addition, the ISP has translated the selection of steps into accessible Dutch. This makes it straightforward for Dutch customers to read and understand the steps. On the other hand, the steps of QNAP also note that the device still is vulnerable to reinfections. Therefore, it also recommends other steps to be taken. In addition to the steps that the ISP also describes, the company also recommends to:

- Change the admin password.

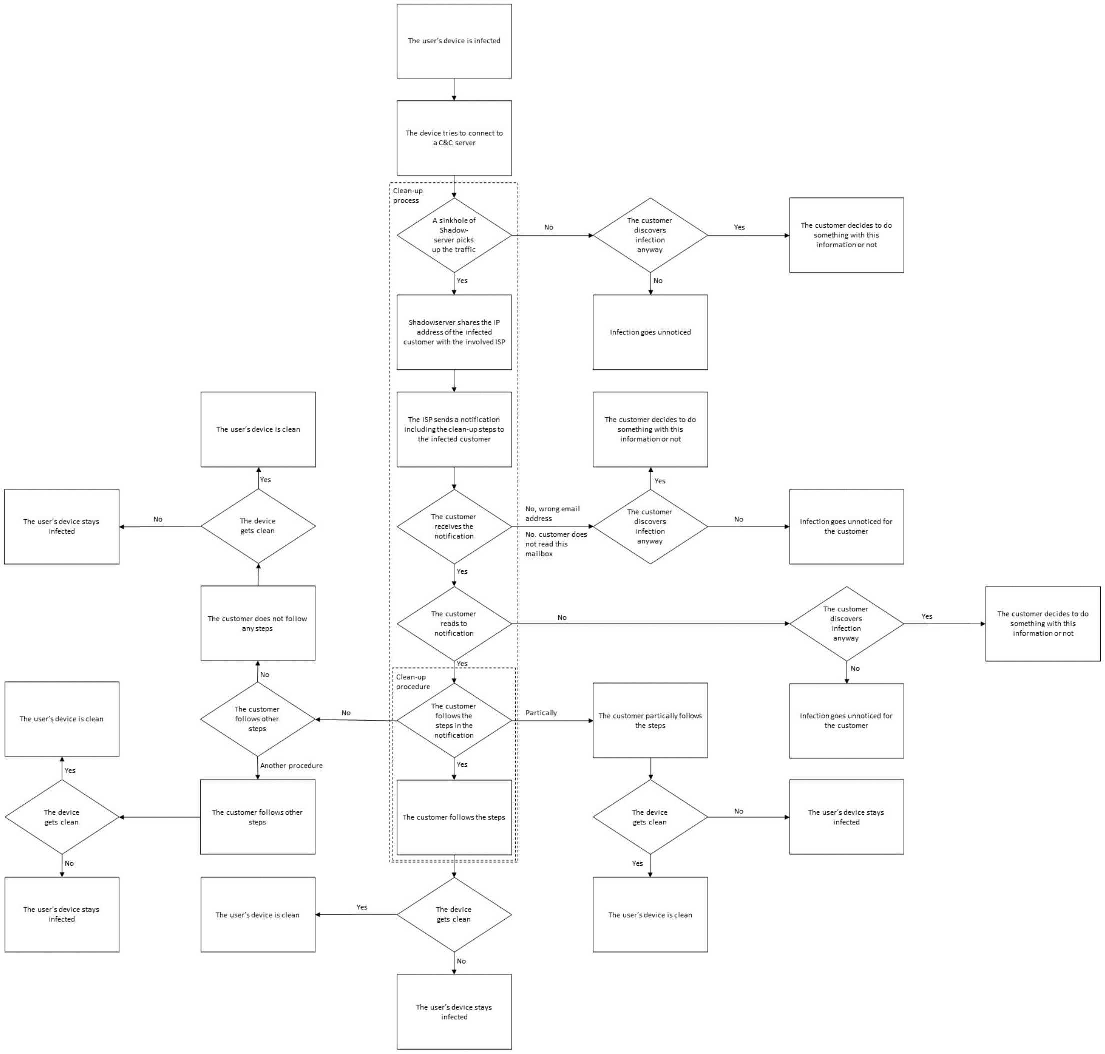

The user's device is infected
The device tries to connect to a C&C server
Clean-up process
A sinkhole of Shadow-server picks up the traffic
No
The customer discovers infection anyway
Yes
The customer decides to do something with this information or not
No
Infection goes unnoticed
Yes
Shadowserver shares the IP address of the infected customer with the involved ISP
The ISP sends a notification including the clean-up steps to the infected customer
The customer decides to do something with this information or not
The user's device is clean
Yes
The user's device stays infected
No
The device gets clean
Yes
The customer receives the notification
No, wrong email address
No. customer does not read this mailbox
The customer discovers infection anyway
No
Infection goes unnoticed for the customer
Yes
The customer does not follow any steps
The customer reads to notification
No
The customer discovers infection anyway
Yes
The customer decides to do something with this information or not
No
Yes
Clean-up procedure
No
The user's device is clean
The customer follows other steps
No
The customer follows the steps in the notification
Partically
The customer partically follows the steps
Infection goes unnoticed for the customer
Another procedure
Yes
Yes
The device gets clean
The customer follows other steps
The customer follows the steps
The device gets clean
No
The user's device stays infected
No
Yes
The user's device stays infected
The user's device is clean
Yes
The device gets clean
The user's device is clean
No
The user's device stays infected

Figure 2: The clean-up process

- Change other user passwords.
- Change QNAP ID password.
- Use a stronger database root password
- Remove unknown or suspicious accounts.
- Enable IP and account access protection to prevent brute force attacks.
- Disable SSH and Telnet connections if you are not using these services.
- Disable Web Server, SQL server or phpMyAdmin app if you are not using these applications.
- Remove malfunctioning, unknown, or suspicious apps
- Avoid using default port numbers, such as 22, 443, 80, 8080 and 8081.
- Disable Auto Router Configuration and Publish Services and restrict Access Control in myQNAPcloud.
- Subscribe to QNAP security newsletters.

If a user would like to comply with those steps, he does not only have to perform more steps but he also has to perform steps that seem to be a lot more difficult than just updating the firmware and installing an application. Even some people with experience with IT might get a little bit nervous when they are ordered to disable SSH and Telnet connections and avoid using specific port numbers. This also provides more insight into the question we posed in section 1.1.2. It seems that there is more than rebooting the system when it comes to removing persistent malware from a device. This could have serious consequences for the success of users in cleaning their devices. To illustrate, removing persistent malware would take more time which will demand more of the user's patience. One could wonder if the user will be willing to spend so much time on the problem. It could also be that the steps to remove the persistent malware surpassed a certain threshold after which most users are no longer able to perform them.

Even more, the user has become dependent on the software of the device producer. Where the user could first perform the steps to remove the virus himself, he can now only hope that the software will do it for him. In addition, malware is known for preferring not to be detected. It is, therefore, more likely that the malware will do everything it can to avoid being detected by the malware remover and hinder the processes of that application as much as possible. This can also have consequences for the behaviour of the users. Because how often will a user perform the steps when they do not seem to work? Moreover, we can all guess how much frustration it causes when the steps that are prescribed do not work and then the Internet connection is shut down by the ISP. It should be mentioned that at the moment, the network of the customers is not shut down if the customer is infected with QSnatch. Due to working from home as a result of the measures to combat COVID, the ISP has decided that shutting down customers is currently too harsh of a measure. Since QSnatch is currently a threat to the customer and not so much to the network, it has been decided not to use this measure. On the other hand, the notification message has not been changed and, therefore, the customer still gets threatened to be shut down if he does not comply. All this makes it intriguing to look more closely at how this process works.

1.1.7 Stakeholder analysis

The previous sections discussed different involved organisations and companies. Table 1 provides an overview of those stakeholders. Moreover, it elaborates on the stakeholders involved, such as the ISP, the QNAP user, NAS producer, NAS retailer, and regulatory authorities.

Stakeholder	What is the stakeholder's primary goal?	What is the stakeholder's primary concern (related to the botnet infections)?
ISP	provide Internet to paying customers	too many bots hindering legitimate Internet use
User	use his NAS device	his data no longer accessible and stolen
NAS producer	sell NAS devices to users	a stigma surrounding the security of its brand decreasing its sales but wants to make a profit and security can be compromised
NAS retailer	sell NAS devices to users	customer dissatisfaction leading to returns of products but wants to make a profit and compromised security
Regulatory authorities	regulating law and regulations	ISPs and NAS manufacturers not obeying the law and leaving users vulnerable

Table 1: Stakeholders

The ISP aims to provide Internet to paying customers. Its customers tend to buy QNAP devices and connect those to their local network. If these devices are subsequently infected by the QSnatch malware, the ISP will receive a notification from ShadowServer, which can retrieve infected IP addresses based on the information from their sinkholes. It is essential to mention that the ISP only receives reports of infections discovered by Shadowserver. This is no guarantee that all infections will be detected and reported. It became evident that an ISP is obligated by law to take action and help infected customers by sending them a notification about the infection and corresponding steps to clean their device. The ISP depends on the customers to solve the issue. The number of employees limits the number of infected customers that can be helped per day. If the infection is out of control and the customer does not remediate the issue, then the ISP could decide to shut down the customer's Internet connection. As mentioned earlier, the measure of shutting down customers their connections is not used due to their customers working from home as a result of the COVID restrictions.

Thereafter, it is up to the QNAP user or infected customer to follow the steps. If the steps successfully clean up the device, the problem will be resolved for the user. If it fails for whatever reason, the virus will remain on the device, and the customer will remain in danger. So far, it looks like an infection with QSnatch is mainly a risk for the customer. At the same time, the customer does not have much information. Infections often seem to go undetected, and so the infected customer will depend on the notification of the ISP. If the ISP does not send a notification, the chances are that the customer would not even know the device is infected. As long as the malware is on the device and the device is connected to the Internet, the customer's data could get stolen or encrypted. The ISP offers different subscriptions for either CM and BM customers. In this way, the customers are also subdivided in the administrative systems of the involved ISP, although it is not excluded that small companies use consumer subscriptions.

The third group is the NAS producer. Since QSnatch targets QNAP devices, we are mainly interested in the producer of those NAS devices. If this research uses QNAP devices, it specifically refers to the NAS devices from the manufacturer QNAP. This company develops NAS devices and also provides firmware and malware scanner updates. QNAP is in a competitive market and will need to be able to offer a device at an attractive price; otherwise, potential customers buying their devices from the competitor. Security is a relatively elusive concept and may therefore might not receive the attention it deserves. This ultimately comes at the expense of the security of the devices. Despite recommendations by CSR and AT, the NAS producers do not seem to comply with all of the demands that are expected of them. As a result, vulnerable products are brought onto the market, ready to be plugged in by unsuspecting users.

QNAP then sells the devices in bulk to retailers. These sellers are mostly technology-oriented and sell their products in physical stores or online shops. There the QNAP devices are then sold to the users and possible customers of the ISP. In the Netherlands, those are companies and websites like CoolBlue, Bol.com, and Mediamarkt. These stores are motivated to sell as many products as possible and will especially suggest the benefits of such devices. Therefore, when purchasing a NAS device, a consumer may not be fully aware of the potential dangers. Retailers cannot change the design of these products but can choose to offer certain products instead of others. Unfortunately, here too, it is the products that are most popular with as much functionality and specifications as possible at the best possible price. The intangibility of security makes it challenging to value the importance of it in the price of the product.

Finally, there are the regulatory authorities such as the AT and ACM. They ensure that all matters in the telecommunications market are conducted in accordance with the law. The moment it appears that an ISP structurally fails to warn customers, the authorities may take action. For example, at some point, they can choose to fine the ISP. Another measure they could take is to have the rules adjusted to give more direction to adequate help from the ISP to protect customers.

1.2 Problem statement

Based on the previous sections, it becomes evident that users of QNAP NAS devices are being targeted by the QSnatch malware. The QNAP's NAS device users are becoming victims of the persistent IoT malware QSnatch. Persistent malware would require more steps to clean it from devices compared to non-persistent variants. Users are in an unfavourable position in that regard: They are in possession of potentially vulnerable devices and cannot rely on a legal process in case their data is leaked or their files encrypted. Fortunately, the ISP notifies the infected users and provides the clean-up steps. It is then up to the user to do something with it. What happens on that side of the story is still unclear. There is a need to understand this essential part of the remediation process. With an exact understanding of what happens when the user gets a notification, this research can determine whether it is feasible for users to go through the steps and address possible improvements of the overall clean-up process. At this point, there are too many different possible scenarios to get to concrete advancements.

1.2.1 Research question and subquestions

As a result of the problem statement, this research aims to answer:
'How does the QSnatch clean-up effectiveness compare to other types of malware, what are the customers' experiences with the clean-up procedure, and to which extent do they succeed?'

This research will address the user behaviour when the infected customer receives a notification and acts upon it. It involves identifying the facilitators and barriers that make the steps doable or not. As described in 1.1.6, there are many possible sequels after the notification. The user may not even read or notice the notification but turn off the device permanently because it started malfunctioning. Without this information, it will not be possible to come to usable insights from the abuse data. In answering this research question, the study would like to unravel four themes: clean-up rates of persistent IoT malware and non-persistent malware or non-IoT malware, user experiences of successful and unsuccessful clean-ups, differences in user experiences between business and consumer market customers, and possible improvements for the clean-up process. Therefore, this research divides the research question into three subquestions:

SQ1) *'How does the QSnatch clean-up effectiveness compare to other types of malware such as non-persistent malware and non-IoT malware?'* This part of the research will address the differences between the clean-up of QSnatch and other malware. It will be particularly interesting to see whether more steps in the clean-up process will lead to differences to the clean-up rates. QSnatch has been around for approximately two years now, and the involved ISP already has notified many infected customers. There have been clean-ups in the past, and some IP addresses stopped showing up in the abuse data. How does this compare to the clean-up of other malware? For example, there is also data available on the clean-up of the Mirai botnet and Windows malware. This research would expect the clean-up of QSnatch to take longer to clean up the devices and at a lower success rate compared to the Mirai procedure as it requires the user to perform more steps. Comparably, this research would also suppose the same when comparing the QSnatch clean-up to non-IoT malware because there are more options when it comes to anti-virus when it comes to computers, for example. But then again, it is not clear what happens on the user's side. The user could just permanently shut down the device and even out-perform others that actually perform the clean-up process. This is still unclear, and it brings this research to the second subquestion.

SQ2) *'What are the customer's experiences of the QSnatch clean-up after they received the infection notification?'* The current abuse data shows whether there is malicious traffic associated with an IP address. When the data shows that a customer of the involved ISP was infected with QSnatch and later did not show up on the data anymore, there are multiple explanations for this phenomenon. A few explanations are user behaviour, attacker behaviour, and simply a device that breaks. This part of the research will investigate what is actually leading it. Do the users receive the notification, and do they also read it? Do the users comply with the prescribed clean-up process or do they just simply shut down the device? Do they follow the procedure themselves or let someone else do it? All of these issues can only become known if this research delves deeper into the user experiences and their behaviour. But, this question does not end there. This research would also provide a better understanding of the underlying processes that make the behaviour of the users and their experiences. It will analyse the user behaviour based on a model from the literature.

SQ3) *'What are the differences in user experiences of the QSnatch clean-up of customers that use their devices for private and business purposes after receiving the infection notification?'* The abuse data and ISP allows this research to analyse the differences between CM and BM customers in the clean-up of QSnatch malware. However, the contact information of the BM customers turned out not to be available to the Abuse Team, and the BM customers are not being informed about QSnatch-infections. However, it will be particularly intriguing to test this as users in private and business environments have different resources to deploy in the process. To name a few: businesses presumably have more money to spent to fix the device and could potentially more easily put a professional on the problem. At the same time, the stakes are much higher in a business environment as the infected devices could be an essential part of the business process or there could be business secrets and customer information on the devices. For these reasons, this research would expect differences between the clean-up experiences of users in the business. As it turned out, there are smaller businesses getting CM subscriptions. In a later stage of this research, we will explain how we will differentiate between those groups and analyse the difference between their experiences. Although the differences may not be as great as with CM and BM customers, if there are differences, it can provide insights into how the ISP should deal with those groups.

1.2.2 Research Scope

This research will analyse the possible clean-ups of QNAP users in the Netherlands. The data on the infections come from a particular ISP and is limited in the sense that it only contains infected IP addresses from the ISP network. There are probably also Dutch QNAP users who also have infected devices but who are not on the network of the involved ISP. Moreover, the data contains infected IP addresses of devices that tried to get in contact with Shadowserver's sinkholes. Presumably, there could be infection going unnoticed as infected devices could normally connect to different C&C servers. This research will go deeper into the sampling sizes and methods in the methodology, but in this section, it is relevant to note that this research's subjects do not contain all QSnatch-infected users in the Netherlands. However, it might be possible to generalise the findings to the whole population.

1.3 The relevance

This section addresses this study's relevance. We will look at how this research will try to contribute. Subsequently, it will discuss the societal, scientific, and managerial relevance.

1.3.1 Societal relevance

As discussed in subsection 1.1, the IoT adaption by society is on the rise and unlikely to halt in the upcoming years. Even more, the IoT is only expected to increase. This development has profound implications for data protection and privacy [8] and enables the establishment of astronomically large botnets, the main threat of the IoT [22]. To make matter even worse, IoT malware has evolved from non-persistent, such as Mirai, to persistent variants. An example of persistent IoT malware is QSnatch. The main threat of this specific malware is that hackers can get their hands on possibly sensitive data or perform a ransom attack [33]. At first glance, the malware may, therefore, seem to harm only the individual user. However, this does not guarantee that the hackers will not update the malware and use the botnet for DDoS attacks and spam. Thus, the malware poses a threat to the community as a whole and is it not only in the users' interest to clean up the malware.

1.3.2 Scientific relevance

As became evident from Appendix C, current literature does not provide much understanding of persistent IoT malware. Even more, it does not address the possibilities of botnet remediation and QSnatch. Therefore, this research will fill the lack of understanding on this topic. Previous research already addressed the clean-up by users and the ISP's role in assisting them, as discussed in subsubsection 1.1.2 [36, 35]. Those studies focused on Mirai, a non-persistent IoT malware. The research's novelty will be the analysis of persistent IoT malware.

1.3.3 Managerial relevance

Naturally, questions can be asked about the course of events in the case of a QSnatch-infected device. Is it the most desirable way to first sell vulnerable devices to customers and later notify customers of infection via malicious traffic while the customer is already at risk? Is it even possible to guarantee that all infections are detected and reported to the customer in the current system? But those are questions that will fall outside the scope of this study. The ISP takes on the task of informing and helping customers, and there are costs involved. For example, the ISP has to hire employees to send the notifications and to answer feedback and questions. Therefore, it will be vital to see if the notification and remediation process is effective or if there are elements in the process that would benefit from improvements. That way, more customers could be supported with the same amount of resources or the same number of customers with fewer resources. Moreover, the ISP also does not want to just simply put its customers in quarantine as it can frustrate them. It may very well be the case that a customer wanted to arrange various matters via the Internet that day but is unable to do so due to the shutdown, with all the nasty consequences that could entail. It could even get so bad that the customer decides to switch to another ISP in the hopes of not being ’harassed’ all the time. In other words, this research will analyse the removal of QSnatch, aiming to get a better understanding of the user response. This would enable the ISP to implement improvements in the notification process. In this pursuit, it would allow a streamlined process and advances in the management of technology.

1.4 Conclusion on the introduction and Structure of the report

The current chapter, section 1 has outlined that the IoT is on the rise and that its mass use has made it an interesting target for hackers. This has an impact on the security of such devices. The IoT devices can be infected with malware and become part of a botnet. Manufacturers of these devices do not always have an equal interest in supplying devices that meet the most stringent safety requirements. Similarly, QNAP devices are susceptible to the QSnatch malware. We saw that QSnatch is a new threat with persistent properties and that users of the affected QNAP devices play an important role in cleaning up the malware as it is ultimately up to them to remove the malware.

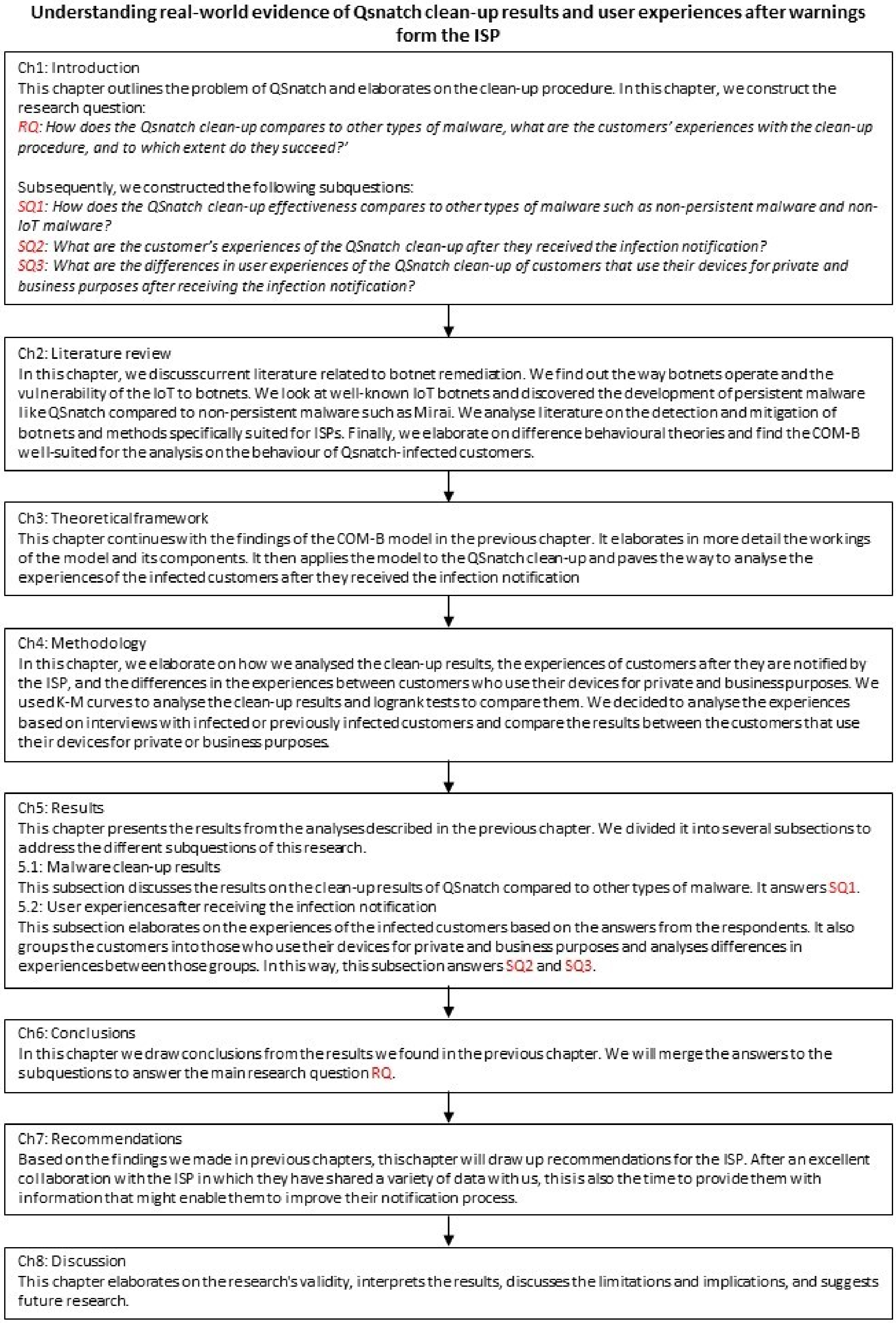
Then, the chapter covered the process of a device becoming infected and starting to generate traffic associated with QSnatch. We saw that Shadowserver captures this traffic and reports infections to the ISP, which in turn is bound by law to take care of its network and customers to send victims a notification. Some customers manage to remove the virus and others have more trouble doing so. Based on this problem, the research poses several questions that it will try to answer. Finally, the relevance of these answers was discussed, and the chapter provided insight into their relevance. Figure 3 shows this structure and organisation of this research. That figure provides a solid overview of how the study will proceed.

This research will continue with a literature review in section 2. This will put the research into a scientific context. It will become clear which dangers lurk when it comes to malware and botnets. We will also gain more insight into the different methods of combatting these dangers. Finally, that part of the research further explores different behavioural theories to understand better what drives users to certain behaviours. This may provide us with more tools to gain new insights regarding the clean-up of QSnatch. In the subsequent chapter, this research offers a theoretical framework on which the research will base part of the analysis. This framework is known for providing insight into which factors are specifically important when it comes to achieving certain behaviours.

Thereafter, the study will elaborate on the method to come to the answers to the research subquestions in the methodology section, section 4. First, it will specifically describe how we will apply a statistical analysis to compare the QSnatch clean-up results with the clean-up of other malware and groups of malware. In that pursuit, it will describe how we answered the first subquestion. Thereafter, we will explain how we come to analysing the user experiences and the differences between the experiences of customers that use their devices for business or private purposes. As such, it will explain how we will answer the second and third subquestions of this research.

The chapter that follows the methodology provides the results. We will first discuss the clean-up results of the malware and find out how the QSnatch clean-up compares to the clean-up of other malware in section 5.1. Thereafter, we will elaborate on the user experiences in section 5.2. In that same section, we will differentiate between the user experiences of customers who use their devices for business and private purposes.

We provide the conclusions of the study in section 6 and, as subsequently, various recommendations for the ISP in section 7. In the final chapter, section 8, we will discuss the validity and limitations of this research and come to suggestions for future research. Finally, we reflect on the MOT program, its relevance to this study, and the course of the research process.



Understanding real-world evidence of Qsnatch clean-up results and user experiences after warnings form the ISP

Ch1: Introduction

This chapter outlines the problem of QSnatch and elaborates on the clean-up procedure. In this chapter, we construct the research question:

RQ: *How does the Qsnatch clean-up compares to other types of malware, what are the customers' experiences with the clean-up procedure, and to which extent do they succeed?'*

Subsequently, we constructed the following subquestions:

SQ1: *How does the QSnatch clean-up effectiveness compares to other types of malware such as non-persistent malware and non-IoT malware?*

SQ2: *What are the customer's experiences of the QSnatch clean-up after they received the infection notification?*

SQ3: *What are the differences in user experiences of the QSnatch clean-up of customers that use their devices for private and business purposes after receiving the infection notification?*

Ch2: Literature review

In this chapter, we discuss current literature related to botnet remediation. We find out the way botnets operate and the vulnerability of the IoT to botnets. We look at well-known IoT botnets and discovered the development of persistent malware like QSnatch compared to non-persistent malware such as Mirai. We analyse literature on the detection and mitigation of botnets and methods specifically suited for ISPs. Finally, we elaborate on difference behavioural theories and find the COM-B well-suited for the analysis on the behaviour of Qsnatch-infected customers.

Ch3: Theoretical framework

This chapter continues with the findings of the COM-B model in the previous chapter. It elaborates in more detail the workings of the model and its components. It then applies the model to the QSnatch clean-up and paves the way to analyse the experiences of the infected customers after they received the infection notification

Ch4: Methodology

In this chapter, we elaborate on how we analysed the clean-up results, the experiences of customers after they are notified by the ISP, and the differences in the experiences between customers who use their devices for private and business purposes. We used K-M curves to analyse the clean-up results and logrank tests to compare them. We decided to analyse the experiences based on interviews with infected or previously infected customers and compare the results between the customers that use their devices for private or business purposes.

Ch5: Results

This chapter presents the results from the analyses described in the previous chapter. We divided it into several subsections to address the different subquestions of this research.

5.1: Malware clean-up results

This subsection discusses the results on the clean-up results of QSnatch compared to other types of malware. It answers SQ1.

5.2: User experiences after receiving the infection notification

This subsection elaborates on the experiences of the infected customers based on the answers from the respondents. It also groups the customers into those who use their devices for private and business purposes and analyses differences in experiences between those groups. In this way, this subsection answers SQ2 and SQ3.

Ch6: Conclusions

In this chapter we draw conclusions from the results we found in the previous chapter. We will merge the answers to the subquestions to answer the main research question RQ.

Ch7: Recommendations

Based on the findings we made in previous chapters, this chapter will draw up recommendations for the ISP. After an excellent collaboration with the ISP in which they have shared a variety of data with us, this is also the time to provide them with information that might enable them to improve their notification process.

Ch8: Discussion

This chapter elaborates on the research's validity, interprets the results, discusses the limitations and implications, and suggests future research.

Figure 3: The structure of the report

2 Literature review

This chapter summarises and discusses the existing literature. The following subsections will underline what IoT botnets are and their operations, the challenge of IoT botnets, well-known IoT botnets and current threats, available detection methods, available mitigation methods, and ISPs mitigation opportunities. Those sections are based on a traditional literature review. The methodology of those sections is in (Appendix B). To better understand customers' possibilities to remove QSnatch from their devices, this chapter will also consider the literature on persistent malware and behavioural theory. Appednix C describes the methodology of that part of the literature review. In the following subsections (subsubsection 1.1.1, subsubsection 1.1.2), the findings do not come from the traditional literature review, but it does lay the study's foundation to the research questions.

2.1 Botnets and how they operate

A compromised computer that carries out the commands of a master is a bot [65]. A botnet stands for a bot network and consists of several bots, a Command and Control (C&C) server, and a botmaster [66]. A botmaster communicates with and controls the bots through a command and control (C&C) network [67]. The C&C server is the central rallying point for this network [66]. With a botnet, a malicious actor, the botmaster, could pursue a variety of attacks such as DDoS, port scan, remote exploits, phishing, spam, spyware, and identity theft [65, 66]. As those attacks are harmful to legitimate users of the Internet, the identification and mitigation of botnets' effects is a hot topic in literature. This chapter will cover those more thoroughly in subsection 2.5 and subsection 2.6.

On the other hand, a botmaster tries to evade those efforts by designing botnets in such a way that they become stealthy [65]. To understand stealthy botnets, the researchers of Leonard et al. [65] suggested a graph-based model and stealth measures to capture botnet (Command & Control) C&C mechanisms. Based on those, their research came to exciting findings on the impact of when the botmaster abandons a bot, the topology, fragmentation, and attack sophistication on the botnet's stealthiness. The research also made suggestions that could improve the understanding of the workings of C&C mechanisms. Based on the work of Vormayr et al. [66], it became evident that the research of Khattak et al. [68] presented a complete overview of the botnets. That research included the use of proxy machines to the other elements described above. The botmaster can optionally employ a number of those machines, called stepping-stones, between the C&C server and itself.

Moreover, it is essential to explain different botnet behaviour features for a full understanding of a botnet [68]. The first step is to transport the bot code to the victim machine via *propagation* mechanisms. Then, the victim machine announces its presence to the C&C server. This is called call-home mechanism or *rallying* [68]. The rallying initiates the C&C channel over which the C&C communication takes place. The botnet has a specific *topology* based on this communication. The bot could implement the C&C channel by using different network paradigms, for exmaple, p2p networks, central, or fluxing servers [69]. Moreover, it could use different protocols such as HTTP, plain TCP or UDP, or custom. The bot and bot owner need the connection because if the connection is lost, the control over the bot is also lost. Therefore, the bot needs to receive commands via the C&C service repeatedly in time [69]. This element is a vulnerability for the botnet. For example, a single, fixed C&C server is detectable by binary or traffic analysis and security researchers or law enforcement can easily sink-hole, blacklist, or take the C&C server down [70, 71]. For this reason, botnets use Domain Generation Algorithms (DGAs) to prevent take-down measures [70, 71].

The research of [70] proposed HYDRA, a comprehensive dataset of Algorithmically-Generated Domains (AGD) enabling differentiation between benign requests and malicious ones out of DGAs-generated traffic. Another approach is that of Upadhyay and Ghorbani [71]. Their study proposed a framework to detect DGAs in real network traffic. That framework uses studied features from legitimate domain names in static and real traffic by considering feature extraction [71]. After that, the victim machine officially has become a bot and awaits further instructions to serve its actual *purpose* [68]. Until that time, there is also the option that the bot spreads the malware to other machines via, again, *propagation* mechanisms. Finally, the botnet carries out operations to ensure that the bot code, C&C communication, C&C servers, and the botmaster are not detected. Those are considered *evasion* operations [68].

To summarise, this subsection addressed the workings and components of a botnet. Also, it presented the effort of botnets to evade detection and the different steps of the infection of devices.

2.2 The challenge of IoT botnets

Besides the traditional security issues, IoT must account for additional factors [72]. The research of Ray et al. [72] named a long device life of over ten years compared to 3 to 4 years, energy efficiency as most IoT need to run on energy budgets disabling security solutions, and configurability as IoT devices require dynamic configurability of security requirements on-field and during execution. Singh and Singh [73] mentioned 6 'challenges' but did not manage to substantiate those properly. For example, the researchers note insurance concerns: *'The autonomous cars are adding insurance industry concerns. But data will make it easier to assess risks & it provides an opportunity for new pricing* *models. For example, insurance premium tuning based on health and driving data.'* Although it is possible to estimate the researchers' intentions, this text alone is not enough to substantiate the issue. Therefore, this chapter will only include a subset of the challenges described in the research of Singh and Singh [73]. Those challenges lack common standards in the IoT and the capability to handle high volume and density of the devices and data.

Kolias et al. [74] pointed to challenges specific for the IoT [74]: wireless connections and their open nature, incompatibility in Internet and IoT application domains, hardware diversity, IoT devices and enterprise networks due to employees bringing their unmanaged, employee-owned devices into managed corporate environments, the identification of every 'thing', and data in the cloud. The study of Zhou et al. [75] discussed features of IoT devices that contribute to IoT botnets' problem. The first feature is that the devices are *constrained* which is comparable to the requirement for IoT to run on energy budgets of Ray et al. [72] but this also encapsulates limited computational capabilities and storage resources of IoT devices. Secondly, *diversity* is a feature of IoT, like the hardware diversity aspect provided by Kolias et al. [74]. Thirdly, the feature *mobile* is in line with the security requirements on-field and during executing of Ray et al. [72]. The other IoT features which make them vulnerable are that they are *unattended*, *interdependence*, *myriad*, *intimacy*, and *ubiquitous* [75].

Due to market forces pushing down the costs of IoT devices, the devices enter the market with many vulnerabilities [76]. Due to the lack of security measures and persistent online connection, IoT devices are a prime opportunity for adversaries to target and abuse them [77] Most IoT devices have no user interface, no security protocol, and no computing and storage capacity to enable firewalls and diagnostic tools [78]. Another problem lays in the diversity of the devices. Due to variations in data formats, frequency of operation, or type of supported communication protocols, it is challenging to implement even basic security measures [79]. As a result, a lot of IoT devices come short in security [80]. Chawathe [81] even concluded that improving the security of an IoT device by hardening its software is not a realistic option, especially in the cost-sensitive consumer market or in legacy-bound industrial settings. Already in 2014, Zhang et al. [82] underlined these vulnerabilities could lead to backdoor problems. Attackers can implement a backdoor in vulnerable devices and take over control. Unfortunately, this has not gone unnoticed by the hacking community and IoT devices are already involved with significant security incidents [76].

The research of Kolias et al. [83] pointed to the ease of infection and stability of the generated bot population caused by the constant and unobtrusive operation, feeble protection, poor maintenance, and noninteractive or minimally interactive user interfaces. According to Dange and Chatterjee [22], the difference between the IoT botnets and traditional botnets is in the underlying devices. They label the detection of IoT botnets as *'difficult'* whereas the other botnets get *'easy'*. With this statement, an explanation and thorough comparison is unfortunately missing.

In 2016, Symantec identified eleven types of IoT malware [84]. They concluded that IoT malware's main purpose is to perform DDoS attacks, and little is required to exploit an IoT device. With the rapid growth of IoT and processing power, Symantec expected attackers might change tactics to cryptocurrency mining, information stealing, and network reconnaissance [84]. During that same year, the world had to deal with a DDoS attack on a DNS service provider [85]. The victim company, Dyn, had to handle over 1 Tbps of data generated by a botnet of primarily IoT devices [86]. Many popular sites such as Amazon, Netflix, and Twitter used Dyn as their authoritative DNS provider and, therefore, the attack resulted in the in-availability of those sites for several hours [87]. In this sense, the DDoS made an organisation inoperable and, thus, had cascading effects on other firms [88].

Before a DDoS, the attacker infects millions of computers worldwide with malware [89]. This infection of locally-separated devices is the distributed element and enables the attacker to take over the device. With the control, the attacker could launch an attack by overloading the victim's server with communication requests [89]. The attacker exhausts the computing resources of the victim and obviates others to access the server [90]. Therefore, this kind of attack impacts the availability of the CIA trait (confidentiality, integrity, and availability) [91] and is a problem as minutes of downtime or latency hurts the customers' satisfaction, the victim's reputation, and eventually the victim's revenue [92].

Due to the minimum user interventions that IoT devices require, users are not likely to note the infections [83]. And even though the IoT devices are in their users' houses, the infections tend to go unnoticed [93]. ISP play an essential role as malicious traffic of the IoT travels through their network [93]. Therefore, ISPs and network operations need to identify vulnerable devices in their network. According to Saidi et al. [94], major ISPs are developing strategies to deal with large-scale coordinated attacks from IoT devices. However, ISPs have not yet succeeded in mitigating IoT botnets. The aim is, therefore, to come up with methods to do so. Subsection 2.7 specifically addresses methods for ISPs.

In conclusion, this section laid out the challenges of botnets. IoT devices turned out to be prone to botnet infections. The characteristics of those tend to make them a suitable target for attackers. However, a point of attention is the detection of IoT and the conclusion of Dange et al. [22] stating that this is difficult compared to the detection of traditional botnets. Further examination is needed to substantiate this claim.

2.3 Well-known IoT botnets

The work of Wainwright and Kettani [95] depicted the history of botnets. The first Internet Relay Chat (IRC) bot, Eggdrop, was non-malicious and developed in 1993 to manage and protect a chat channel [95]. Unfortunately, soon after its publication, other bots began attacking IRC users and servers. Then, the GTbot became a well-known malicious IRC botnet for launching DDoS attacks. In 1999, the first large-scale DDoS attack occurred causing the unavailability of the University of Minnesota computer network for two days [95, 96]. The first botnet to require little programming ability was the Agobot. This collection of worms could perform packet sniffing, keylogging, rootkit installation, and DDoS attacks.

Nonetheless, these botnets were primarily not IoT botnets. According to Mahjabin et al. [97], the first IoT malware was Linux.Hydra in 2008. On the other hand Zhang et al. [82] claimed that the first IoT malware, Linux.Darlloz, was discovered in 2013 by Symantec. This worm's first purpose was to make the IoT perform DDoS attacks, and a later variant made the infected IoT devices mine cryptocurrencies [98]. Zhang et al. [82, p. 231] specifically stated: *'In Nov. 2013, Symantec confirmed the finding of the first IoT malware, Linux.Darlloz, which brings up the malware issue for IoT security.'* The work of Ngo et al. [99] also underlined that the first IoT malware was Linux.Hydra. It seems that the research of Zhang et al. [82] is factually incorrect on this point. Ngo et al. [99] provided an overview of IoT malware. Those researchers addressed that malware writers generate new malware based on variants of previous malware. After Linux.Hydra, malware writers developed the variants Psybot, Chuck Norris, and Tsunami.

The research of Salim et al. [96] described three recent popular botnets: Bashlite, Mirai, and Reaper. Bashlite is a prevalent malware with Linux-based IoT devices as its primary target [96]. The ancestor of Bashlite is Tsunami [99]. From Bashlite, malware writers developed the Mirai malware [99]. The Mirai botnet performed a massive DDoS attack on Dyn with a volume of 1.1 Tbps using 148,000 infected IoT devices [95, 96]. subsection 2.2 already discussed this attack. A variant of the Mirai code is the Reaper botnet. This code is even more dangerous as it can exploit other security vulnerabilities in the victim devices' code rather than using only default credentials [96]. The most recent IoT botnets are Wicked, Satori, Jen X, and IoTrojan [97]. Those malware are all variants of Mirai. The malware are evolving over time, increasing in botnet size, and the malware tend to use brute force methods to infect IoT devices [97]. This trend in IoT malware shows that the world is certainly not freed from IoT botnets and can even expect larger attacks to be forthcoming.

In short, this research examined other researchers' work to get a better understanding of IoT malware. This section discussed the development of botnets and IoT botnets, and it became evident that malware tends to evolve from previous malware. The way IoT malware evolves could indicate where future dangers will lurk.

2.4 Persistent malware

Until recent times, the security community has been fighting malicious programs for Windows-based operating system [100]. Due to the adoption of the IoT, hackers are increasingly shifting their targets to Linux-based devices. The research of Cozzi et al. [100] consisted of a comprehensive study on Linux malware. The research of Chen et al. [101] proposed a method to identify malware based on software genes. The researchers referred to biological knowledge and built a gene model of malware by calculating a similarity score.

Most IoT malware variants cannot gain persistence [24]. Those non-persistent malware are stored and executed from within temporary filesystems in the Random-Access Memory (RAM). This type of memory is volatile, and stored programs and data are lost when the device loses power [24]. As discussed in subsection 1.1, some malware will scan the Internet for other vulnerable devices [24]. If the device loses power or is rebooted, the device can be easily reinfected within minutes. However, some IoT malware families are persistent is some form. The malware would then be harder to remove because this would require the user to modify the device's flash memory [24]. According to Brierly et al. [24], this is not possible for the average user. Their study also examined two persistent IoT malware: Torii and VPNFilter. Torii is a variant of Mirai that has six techniques to gain persistence. Secondly, VPNFilter is a complex IoT malware believed to be developed by Fancy Bear, a Russian-based hacker group. So, beyond QSnatch, there is more persistent IoT malware currently active.

VPNFilter is specifically designed to target SOHO routers [102]. The malware can steal typed passwords and create fake copies of a page so that the victims do not know that they are being hit. It can also destruct the device, leave it unusable, and switch off internet access [102]. The router platforms Linksys, TP-Link, Qnap, Netgear, and MikroTik are more susceptible to the malware as they implement home networks on internet gateways.

Brierly et al. [103] analysed feasible ways for infecting IoT devices. Those researchers also examined potential methods for gaining control and applying a persistent change to the devices. Then, they developed a proof of concept ransomware and tested it against six IoT devices. This ransomware also enabled the researchers to determine limitations that may discourage attackers from developing IoT-specific ransomware and discuss workarounds [103]. Moreover, the researchers also proposed countermeasures to their proof of concept attack. Those measures are making partitions read-only, restoring the bootloader through a factory reset, implementing the principle of 'least privilege', and device updates.

To better understand persistent IoT malware, it is also essential to look at persistent Windows malware. The work of Gittins and Soltys [104] analysed five persistent Windows malware: Emotet, OceanLotus Symantec DLL Sideloading, TrickBot OceanLotus - Explorer-COM Hijack, and Agent Tesla. The researchers discussed the file type that carriers the malware and the malware's persistent mechanism. Botacin, De Geus, and Grégio [105] proposed a dynamic analysis system for categorising different types of malware. More relevant for the current study, they described how DarkKomet gains persistence over a Windows system by adding Registry entries to enables automatic execution at every system startup.

2.5 Detection methods

The paper of Dange et al. [22] did not only lay out a challenge to IoT botnets; it also pointed to IoT botnet detection methods. According to this paper, the detection methods are categorised into host-based and network-based detection methods. With the host-based detection technique, the focus is on the victim machine or host machine. All the activities performed on this machine are tracked [22]. Whereas with network-based detection, the aim is to track the network traffic. Network-based detection comes in several flavours. There are active and passive monitoring techniques. The passive methods are signature-based botnet detection techniques, Domain Name System (DNS)-based botnet detection techniques, and anomaly-based botnet detection techniques [22].

The signature-based technique detects a botnet based on an Intrusion Detection System (IDS) [106]. This system monitors networks for known malicious activities and policy violations based on matching attack signatures [106]. A signature is a footprint or pattern of a series of bytes in network traffic. To communicate with the C&C server, bots need to send DNS queries [22]. Other techniques, DNS-based techniques, use these queries to detect a botnet. According to Dange et al. [22], this technique is the most famous and easy for botnet detection. The anomaly-based method detects IoT botnet behaviour based on the profile of normal behaviour. Therefore, it is required to have this profile in advance [106]. The study of Dange et al. [22] also provided thorough explanations on these techniques and examples of other research proposing specific methods for each category. In the sake of keeping an overview, this research only mentioned the techniques but gladly refers the reader to the study of Dange et al. [22] for more information.

Al-Duwairi et al. [106] underlined two additional detection techniques: specification-based and hybrid-based. The specification-based technique is similar to the anomaly-based technique. Besides the anomalies, it also takes system specifications into account [106]. When deploying two detection approaches, this would be a hybrid-based technique. For example, if the detection involves both signature-based and anomaly-based techniques. The researchers Li et al. [107] elaborated on the development of signature-based detection using blockchain technology. Their solution could deliver a verifiable manner in distributed architectures without the need for a trusted intermediary.

The work of Dwyer et al. [108] provided a DNS-based profiling scheme over real datasets of Mirai-alike botnet activity. Then, the research evaluated its suggestion over various Machine Learning classifiers and demonstrated its scheme's applicability. This approach resulted in a reduction of the detection time whilst maintaining a high level of accuracy. Comparably, the study of Xu et al. [109] combined static and dynamic analyses of Mirai behaviour. They propose an application of Threat Tracer, an information system simulator, for Mirai malware's operating processes. In this way, the researchers could simulate the malware's processes and reveal a system's vulnerabilities. Kurniabudy et al. [110] provided a testbed topology for anomaly detection research. This topology resulted in a global framework for anomaly detection in IoT and proposed a distributed preprocessing framework.

In conclusion, this subsection addressed various detection techniques. There are host-based and network-based techniques, and of the latter, this chapter discussed an example with suggestions for each of the three subdivisions.

2.6 Mitigation methods

The research of Bertino and Islam [111] addressed several protection techniques: ensuring that all default passwords are changed to strong passwords, updating IoT devices with security patches, disabling Universal Plug and Play (UPnP) on routers unless necessary, monitoring IP ports 2323/TCP and 23/TCP for attempts to gain unauthorised control over IoT devices using the network terminal (Telnet), and monitoring for anomalous traffic on port 481011, as infected devices often attempt to spread malware by using this to send results to the threat actor. Frankly, this is to some extent contradictory to the statement of Chawathe [81] that improving the security of an IoT device by hardening its software is not a realistic option as described in subsection 2.2.

Moreover, the view of Bertino et al. [111] only considered the security of the individual devices. This research also wants to cover the mitigation of botnets' effects. The research of Wainwwright and Kettani [95] covered mathematical models to allow assumptions about botnets and tests of possible defences. It described seven different models: epidemiological models, machine learning models, stochastic models, game theory models, nonparametric Bayesian models, graph models, and economic models. The model of Leonard et al. [65] discussed in the subsection 2.1 is a graph model.

Salim et al. [96] focused on DDoS attacks by IoT botnets. The researchers divided the defence against this attack into prevention, detection, and mitigation. Although this research already discussed their proposed detection mechanisms in the previous section, the prevention and mitigation mechanisms are worth explaining. The researchers divided different approaches provided by other studies in these categories.

Prevention helps ensure a DDoS attack does not result in disabling or takeover of the system [96]. The researchers put a mutual authentication scheme, MECshield, NBC-MAIDS, MAEC-X, Fast path, Classifier system, and a honeypot in this category. On the other hand, mitigation reduces the severity of an ongoing DDoS attack [96]. A honeypot, algorithm to calculate cosine similarity of vectors, Software-Defined Networking (SDN) and Support-Vector Machine (SVM), SDN and Fog networking, Network Functions Virtualisations, and Deep reinforcement learning are suggestions in this category [96]. However, the research of Salim et al. [96] primarily focused on DDoS attacks, and the presented prevention and the mitigation mechanisms are also based on this type of attack rather than remedying botnets.

Fortunately, the study of Salim et al. [96] also suggested changing passwords on the devices instead of keeping the passwords the same on entire product lines. This suggestion is the same as the first protection mechanism provided by Bertino et al. [111]. Other recommendations of Salim et al. [96] are to implement a firewall, update firmware, implement a centralised body to provide certification and implement government law that requires ISP providers to implement preventive measures such as Ingress filtering.

The study of Mahjabin et al. [97] took a different approach. Those researchers proposed a load distributed mitigation process and a benign-bot mitigation method based to cover DNS flood attacks. With load distributed mitigation process, DNS service providers make agreements to deal with DDoS attacks. For example, one or more other suppliers will take over the load when a DDoS attack is involved [97].

The benign-bot method uses a bot program installed on customers' DNS local servers to allow IP addresses of a list of paid businesses' websites to maintain in the caches [97]. The benign bot in the resolver makes an automatic request for a particular domain to keep the record always fresh. In this way, the customer can access the websites even when the DNS servers are down.

To test their approach of load distribution, Mahjabin et al. [97] simulated DDoS attacks. Their testing used a C-based simulation library simlib and covered normal to extreme cases of incoming traffic. Besides this explanation, the methodology is rather limited. It is challenging to determine what the researchers did. Moreover, the research did not cover a simulation on the benign-bot method.

In summary, this subsection discussed mitigation methods for botnets and DDoS attacks. There are ways to make IoT devices less vulnerable to malware, DDoS attacks less effective, and ensure that DNS requests remain possible during a DDoS attack. Although the research seemed to be persistent, examinations of mitigation approaches on real DDoS attacks is missing and does not get further than simulations. These are all techniques that apply in general, but do not yet represent well what ISPs can do. The research by Bertino and Islam [111] mainly focused on the protection of devices, which the ISP has little control over. ISPs have a central role in solving botnets and the next section discusses the measures specific to the ISP.

2.7 ISP's mitigation opportunities

The last suggestion of Salim et al. [96], implement government law, is vital for this research as it involves ISPs. The suggested filtering technique helps detect attacking traffic and halt ongoing traffic by blocking all packets with spoofed IP addresses. The ISP can implement this on edge devices like a router. The edge device will then deny access to malicious traffic and allows traffic to pass through. According to Salim et al. [96], a mandated law passed by the government is essential for this filtering technique to be effective.

Comparable to the statement of Van Eeten and Bauer [93] described in the subsection 2.2, Ko et al. [112, p. 53] stated that *'Since internet service providers (ISPs) connect the internet with users, the mitigation system should be deployed within the ISP domain to deliver a more efficient solution.'* The researchers meant more efficient compared to the *'inefficient solutions'* of utilising auxiliary servers at the host site, in the cloud, or at dedicated data scrubbing centres provided by other researchers. By either blocking or dropping the malicious traffic, the ISP prevents it from propagating any further on the network to reduce the number of resources wasted. Another argument is that by deploying the mitigation within the ISP domain, the rerouting of malicious traffic causing delays is no longer required [112].

Before providing its contribution, the research of Ko et al. [112] addressed the related work and pointed to other research on the signature-based and anomaly-based detection methods. Although this research already covered the differences between those methods in the subsection 2.5, the reason to put the study of [112] here is that it put the techniques in the perspective of the ISP domain. The researchers point to other work on Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, Boltzmann machines, and Self-Organising Maps (SOM).

Then, the paper of Ko et al. [112] proposed a 3-layered SOM. The researchers tested their method based on four types of DDoS attacks (Internet Control Message Protocol (ICMP) flood attack, Transmission Control Protocol (TCP) flood attack, User Diagram Protocol (UDP) flood attack, and Distributed Reflection Denial of Service (DRDoS) attack). The researchers [112] provided explanations of those attacks and the testing methodology. The researchers also compared the results of their method with other unsupervised and supervised learning algorithms and showed that their SOM performed better at separating malicious from normal traffic.

Finally, Ko et al. [112] underlined the need for future research. Based on their SOM, the mitigation could improve by increasing the number of SOM dynamically. The researchers also pointed to improving the efficiency of dynamic feature selection as this process is computationally expensive. Thus, their suggestions are two-fold: increasing the number of SOM dynamically and improving the dynamic feature selection.

To conclude, this subsection underlined detection and mitigation approaches specifically for ISPs. In combination with the previous subsection, it showed there is a myriad of approaches for ISPs to take. Like in the previous section, research is missing data from real-life incidents. Moreover, little research specifically aimed at mitigation from the perspective of ISPs as only Ko et al. [112] dared to do so. Moreover, the previous subsections laid a foundation for an understanding of botnets and the mitigation by ISPs. However, the described methodologies do not involve users as stakeholders. That is strange considering the users are currently the only one able to remove the malware from their devices. Moreover, it also did not discuss persistent malware, which will be considered next.

2.8 Behavioural theory

At some point, the centre of attention becomes the infected customer. As discussed earlier, they are the group that needs to get the malware off their devices. To better understand what drives them to certain behaviours, this chapter discusses current theories on the subject. Ultimately, the ISP wants to bring about a certain change in the behaviour of the customer. It is important for the ISP that the customer will carry out the steps and that is why this chapter will take a look at theories about behavioural change.

One of the first papers on attitude change is that of Rogers [113]. That researcher discussed the protection motivation theory. It is based on three components of fear appeal influencing the cognitive mediating processes, which result in attitude change. Those components are the magnitude of noxiousness, the probability of occurrence, and the recommended response efficacy. In turn, those components affect the appraised severity, the expectancy of exposure, and the belief in the efficacy of coping response, respectively [113]. Those cognitive mediating processes, in turn, influence the protection motivation. Finally, the protection motivation results in adopting the recommended response [113].

Beyond the theory of Rogers [113], there are other theoretical approaches to understand how risk shapes risk perceptions discussed by the research of Van Schaik et al. [114]. Their paper described the risk compensation model of Adams [115], the presentation of risk information by Gigerenzer and Todd [116], availability of risk information of Kahneman [117], affect in risk perception of Finucane et al. [118], revealed risk-related preferences of Starr [119], and expressed risk preferences by Slovic [120]. According to Gollwitzer [121], goals or resolutions stand a better chance of being realised when they are furnished with implementation intentions that link anticipated suitable opportunities to intended goal-directed behaviours. By implementing intentions to delegate goal-directed behaviours, this strategy uses the automatic control of the action. This strategy helps people effectively meet their goals while dealing with initiating goal-directed actions, tempting distractions, bad habits, and competing goals [121].

Hadlington [122] studied the attitudes of 515 employees to cybersecurity. The researcher found that aspects such as lack of skills, knowledge, and awareness were the critical barriers for the employees to engage in active cybersecurity. Moreover, the study of Thompson et al. [123] consisted of think-aloud interviews with students about cybersecurity. That study found that the students had misconceptions and analysed where those arose. The examined students made overgeneralisations; had conflating concepts, biases such as user bias, psychical bias, and personal bias; and incorrect assumptions.

The research of Van Schaik et al. [114] also involved students' perceptions on cybersecurity. From that study's respondents, it became evident that the perceived risk was highest for identity theft, keylogger, cyber-bullying, and social engineering. Significant predictors of perceived risk turned out to be voluntariness, immediacy, catastrophic potential, dread, the severity of consequences and control, and Internet experience and frequency of Internet use. Finally, control was a significant predictor of precautionary behaviour. Another approach to influence users to implement more security is the involvement of an authoritative figure. A famous and above all extreme example of an authority's effect is in the study of [124]. That research showed that subjects obeyed orders to administer deadly shocks to another person.

The study of Tversky and Kahneman [125] showed that people rely on a limited number of heuristic principles for reducing complex tasks of assessing probabilities and predicting values to simpler judgemental operations. That research admitted that these heuristics are quite useful and underlined that they could lead to severe and systematic errors [125]. The heuristics are representativeness, availability, and adjustment and anchoring.

Representativeness means that people evaluated probability to which A is representative of B or by the degree to which A resembles B [125]. The second heuristic, availability, has to do with the ease with which instances or occurrences can be brought to mind. Finally, the researchers speak of anchoring when people start with an initial value and adjust their estimation along the way [125]. Then, adjustments tend to be insufficient, and the outcome is biased toward the initial value. After that, the researchers explained the different biases that result from these heuristics [125].

According to Pfleeger and Caputo [126], security is intertwined with the way humans behave when trying to meet a goal or perform a task. Moreover, security is not the primary but a secondary task in most cases. That leads to the person trusting that the system assures safety, and when security slows down the primary task, a user subverts the security [126]. Those researchers were gathering this insight while conducting interviews.

Other insights are how limitations on memory or analytic capability (cognitive load) interfered with an analyst's ability to perform, inattentional blindness is a particular aspect of cognitive load that played a role in each scenario, there is a significant bias in the way each interviewee thinks about security, and there is a significant element of risk in each scenario and decision-makings have a difficult time both understanding the nature of the risk and balancing multiple perceptions of the risk to make the best decision in the time available [126].

Furthermore, the research of Pfleeger and Caputo [126] also pointed to the studies of Mayo and Hollander [127] and Slovic [128] as essential papers on risk perception and risk communication. Their research offered even more vital insights because it discussed relevant areas of behavioural science next [126]. Useful contributions for cybersecurity are that recognition is easier than recollection, and frequent changes to a memorised item interfere with remembering the new version. Then, Pfleeger and Caputo's research [126] continued with findings on cognition, biases, heuristics, and health-related behavioural models. This research will not discuss them in detail but recommends the reader to take a look at the paper of Pfleeger and Caputo [126]. Firstly, the insights on cognition are the identifiable victim effect, the elaboration likelihood model, cognitive dissonance, the social cognitive theory, and the bystander effect.

Secondly, the discussed biases are the status quo bias, framing effects, the optimism bias, the control bias, the confirmation bias, and the endowment effect [126]. Moreover, the heuristics are the affect heuristic and availability heuristic. Finally, the paper discussed the health-related behavioural models the health belief model, the extended parallel process model, the illness representations, the theory of reasoned action [129]/theory of planned behaviour [130], the stages of change model, and the precaution-adoption process theory.

Yildirim and Ali-Eldin [131] reviewed the literature to come up with a conceptual model regarding the acceptance of technologies and influencing factors such as risk and trust. Those researchers explain how the Technology Acceptance Model (TAM) of Davis [132] evolves from the theory of reasoned action [129]. The TAM discussed two factors: perceived usefulness and ease of use to determine an individual's intention or acceptance of using technology. It later became one of the most used technology models [133]. TAM2 of Venkatesh and Davis [134] is the improved model of TAM and added five factors norm, image, job relevance, output quality, and result demonstrability [131]. Then, the unified theory of acceptance and use of technology is another upgrade of the TAM and includes performance expectancy, effort expectancy, social influence, and facilitating conditions as the primary constructs to a user's intention of using an information system [133].

The work of Redmiles et al. [135] analysed the quality of security advice. Those researchers determined to what degree 374 unique recommended behaviours were comprehensible, actionable, and effective. The actionability or perceived actionability was determined by the confidence, time consumption, disruption, and difficulty. The study showed that the majority of advice is perceived by the most users to be at least somewhat actionable and somewhat comprehensible. On the other hand, both the users and experts of that study struggled to prioritise the advice [135]. Then, the Communication-Human Information Processing (C-HIP) approach described in the study of Conzola and Wogalter [136] addressed behavioural compliance to warnings as information that flows from a source through different stages to a receiver. Those stages are the source, the channel and the receiver. At all of those stages, the information could flow through, or it could bump into obstacles. If the information does flow through the system, the receiver will change its behaviour [136].

Although this research would not dispute the elements of the C-HIP approach, however, it is a shame that the model does not discuss the work of previous studies. We could still gain a lot from previous work. For example, as this research described above, there are many heuristics, biases, and other factors involved in the process, while the C-HIP model did not address those.

A research that did involve literature is that of the European Union agency for Cybersecurity (ENISA). This study involved 688 publications on behavioural sciences constructs that influence human behaviour in cybersecurity [137]. The report argued that previous research ignored the difference between correlation and causality. Based on the report's view, the studies neglected the possibility that other underlying factors influence both the measured constructs and the security behaviour [137]. For example, it discussed the theory of planned behaviour of Ajzen [130]. That theory suggested that the attitude and the subjective norm result in motivations to perform a certain behaviour. It also includes self-efficacy, a construct that Bandura claims to be the most important indicator of behavioural change [138]. According to Bandura [138], it initiates coping behaviour and it is, therefore, the precondition for change.

However, this still does not capture the full picture according to ENISA's report [137] because the study of Karlsson et al. [139] showed that awareness, intention, and self-efficiacy only leads to secure behaviour if the subject, in this case an employee, finds that such behaviour does not conflict with other organisational values. Therefore, ENISA suggested to look beyond the tunnel vision on security which ignores other factors driving security behaviour [137].

ENISA's report [137] concluded that most current metrics are not suited to measure human behaviour or indicate how to influence behaviour. According to the report, the theory of planned behaviour ignores the wider contextual factors and the current models tend to assume that compliance is a positive outcome. Another conclusion is that there is increasing evidence that increasing users' understanding of cybersecurity threats or fear of the consequences is not an effective tool for changing behaviour [137]. Finally, there is a moderately reliable link between people's ability to cope in the face of threats and their cybersecurity behaviour where the report divides coping into the effectiveness of the response and the user's ability to carry out the response [137].

Therefore, the report proposed two models that, in the eyes of the writers, are particularly suitable for cybersecurity. Those models are the COM-B and Fogg's behaviour model [137]. According to the COM-B model [1], whether a behaviour is enacted depends on the interrelated factors capability, opportunity (physical and social), and motivation (automatic and reflective). The various factors are then divided into subcomponents. The work work of Michie et al. [140] then used this COM-B model to construct the behaviour change wheel. That model uses the the sources of behaviour, described above, as the first layer of the wheel. The other two layers of the wheel consists of nine interventions and seven types of policy.

Fogg's model [141] or B=MAT model assumes that the likelihood of a behaviour (B) occurring is a product of motivation (M), Ability (A), and the appropriate trigger (T). The report of ENISA suggested to use the COM-B model to identify why a desired behaviour may be performed and Fogg's model for thinking about possible interventions [137]. This gives a significant advantage compared to the other models described in this chapter. If we use the COM-B model to identify why a certain behaviour may be performed, then we could use the Fogg's model to come to improvements we could make to the notification process. Since the other models cannot offer this advantage, this study will use the COM-B model to determine the factors that may or may not lead to the cleaning of infected devices.

2.9 Summary on the literature review

In this section, we tried to get a grip on the current literature on botnets, botnet mitigation, and behavioural theory. In this way, we were able to ensure that we incorporated the knowledge obtained by previous researchers. We found out the way botnets operate and that bots need to get connected with a C&C server once in a while to receive commands from the hacker and that the hackers use GDAs to prevent people from finding out those servers. It also became evident that the IoT is prone to botnet infections. We then looked at well-known IoT botnets, such as Mirai and the development of malware to become persistent. We discovered that Mirai is a non-persistent IoT malware, while malware such as QSnatch, VPNFilter, and Torii are persistent IoT malware. The latter group gains reboot persistence and does not get removed after an infected system is rebooted. This affects the clean-up process when compared to non-persistent IoT malware. We then discussed the various detection and mitigation methods and specifically looked at the methods that are suitable for an ISP. Finally, we compared different behavioural theories and assessed that the COM-B model would be suitable for analysing customer behaviour after they received an email notification from the ISP. With the COM-B model, we would be able to identify why a customer would perform the steps. In the next section, we will discuss the COM-B model in more details and apply it to the QSnatch clean-up process.

3 Theoretical framework

In section 2, this research already touched upon the COM-B model of Michie et al. [1]. Previous researchers used this model to analyse the behaviour of individuals [142, 143, 144, 145, 146, 147, 148, 149]. Those researchers are mainly active in the field of health behaviour. Although highly suggested by ENISA [137], the model has not been applied to cybersecurity to the best of our knowledge. This research would like to make this connection, but it first dives deeper into the model. We want to use the COM-B model to analyse the experiences of QSnatch-infected customers after they received the ISP's infection notification. Therefore, this chapter will explain the components of the model and apply it to the QSnatch clean-up.

3.1 The COM-B model

As previously discussed, the model has three interacting components, capability, opportunity, and motivation, that result in behaviour [1]. Not only do the components influence each other, but behaviour can also affect the components [1]. These relationships make it a dynamic model [142].

Capability is the *"individual's psychological and physical capacity to engage in the activity concerned"* [1]. It includes having the necessary knowledge and skills and consists of physical capability and Psychological capability. The Physical capability is the capability to engage in essential physical processes. Examples from the medical dictionary are grip strength, walking speed, chair raising, and standing balance times. The Psychological capability is *'the capacity to engage in the necessary thought processes - comprehension and reasoning'* [1].

With opportunity, Michie et al. [1] meant *'all factors that lie outside the individual that make the behaviour possible or prompt it'*. The researchers distinguished Physical opportunity, *'afforded by the environment'*, from Social opportunity, the cultural environment that directs the way an individual thinks.

The motivation of a user is *'all those brain processes that energise and direct behaviour'* such as *'goals and conscious decision-making'*, *'habitual processes, emotional responding'*, and *'analytical decision-making'* [1]. This component is divided into Reflective processes (*'evaluations and plans'*) and Automatic processes (*'emotions and impulses that arise from associative learning and/or innate dispositions'*). Figure 4 combines the components of the COM-B model.

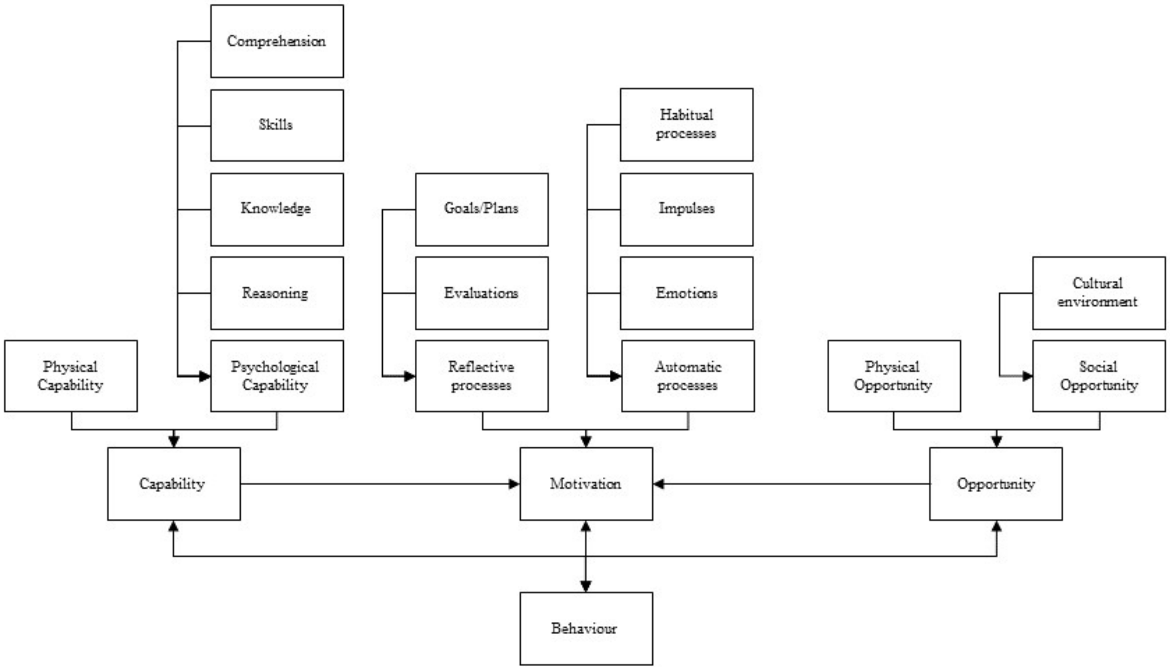


Figure 4: COM-B model [1]

3.2 Applying the model to the clean-up

In terms of the QSnatch clean-up procedure, capability is the individual's psychological and physical capability to perform the required clean-up procedure. The physical capability would mean that the user would need some physical capabilities to perform the steps. The steps require the user to use a computer to make adjustments to the device. Since that is not different from the capabilities for setting up the device, this research predicts that no relationship between physical capability and the clean-up experience exists. Assuming that other subcomponents stay the same, this study would expect users to perform different steps when the described ones are physically not possible.

Then, the psychological capability involves the comprehension and reasoning of users. This part is about the user's understanding of the malware and the remediation. As discussed earlier in this study, the clean-up of persistent malware ought to be more problematic as it requires more remediation steps. Therefore, this research would argue that the users' psychological capabilities are influencing the clean-up by the users.

This research considers opportunity as all of the factors that lie outside the user that make the clean-up possible or prompt it. Physical opportunity is all the things afforded by the environment, such as receiving the clean-up notification, the source of the notification, and other obligations of the user. Since the malware operates relatively stealthy, this research would assume that not receiving the notification would have a tremendous impact on the clean-up. In that case, it is likely that the customer does not perform the steps. At the same time, if the user could also not trust the notification of the ISP. However, this is unlikely, and this research would hypothesise that the source of the notification would not have a significant impact on the clean-up.

Moreover, social opportunity is the cultural environment that directs the way an individual thinks during the clean-up process. Maybe there are stigmas of having malware on a device. Also, it could cause fear when other people might find out about the infection. On the other hand, this research would assume that social opportunity does not influence the clean-up. Especially in this component, this study expects differences between business and consumer market customers. Businesses could get help from third parties by paying for their services and have a heavier impact on their reputation when data is breached.

The motivation of users includes all the brain processes that energise and direct the clean-up. This study denotes the reflective processes of motivation as the users' evaluations and plans during the clean-up process. This research assumes that those impact the clean-up as users are unlikely to follow the procedure immediately. This provides users with time to plan and evaluate the procedure.

Also, automatic processes include users' emotions and impulses. It will not be surprising if users are shocked if they receive a notification from their ISP informing them about an infection. This effect would at least trigger some effects during the clean-up. For this reason, this study would assume that the automatic processes affect the clean-up.

Then, the behavioural change wheel would suggest defining the problem in behavioural terms and pick a target behaviour [140]. In case of a QSnatch malware infection, the desired behaviour of the affected customers is that they perform the steps. The work of Jackson et al. [142] aimed at relating the components of the COM-B model to medication adherence. Therefore, that study did not specify a specific target behaviour. Those researchers covered the problem by conducting a literature review and determined which factors were relevant based on that review. This research will take a different approach. The next chapter will provide insight into the methodology.

3.3 Summary on the theoretical background

This chapter aimed to provide a more in-depth understanding of the COM-B model. We found that preferred behaviour might be performed according to three components that, in their turn, are influenced by subcomponents. In addition, we applied the COM-B model to the QSnatch clean-up and put the components into the perspective of the clean-up. As such, we are able to analyse the experiences of the customers after they receive the ISP's infection notification based on the COM-B model. We will use these insights in the following chapter. That chapter will describe the methodology of how we will analyse the clean-up results, user experiences, and the differences between the customers that use their devices for private or business purposes.

4 Methodology

This section addresses the nature of the study, data collection methods, sampling design, and analysis. In the previous chapter, we applied the COM-B model to the QSnatch clean-up enabling the analysis of the user experiences of the infected customers after they receive the ISP's notification. The current chapter will analyse how to measure the clean-up results of QSnatch and compare those with the results of other types and groups of malware. It will also elaborate on how we analysed the user experiences after the customers receive the notification of the ISP and compare the results of customers that use their devices for private or business purposes. In the following sections, we will discuss our decision to analyse the QSnatch clean-up using a mixed-methods approach consisting of quantitative data analysis and interviews. The following sections will explain those methods in more detail.

4.1 Clean-up effectiveness and User experiences

This research would like to unravel the performance of the involved ISP's customers in cleaning QSnatch from their devices. For that, we want to compare the results of the customers who remove QSnatch with those of customers with other malware. Moreover, if those results turn out to be unsatisfactory, we also want to take a deeper look into what is happening at the side of the customers. In the best case, we can point to several factors that hinder the clean-up process considerably. With that information, we could draw the ISP's attention to those factors and make proposals to adjust certain aspects in the notification process. So, we want to look at the results of many different customers as well as the specific experiences of customers in their efforts to remove the malware. This would make a mixed-methods approach preferable for this research. A mixed-methods approach is not new in cybersecurity literature.

To analyse user behaviour in the context of cybersecurity, Kim et al. [150] conducted trend analysis, semi-structured interviews, and an online survey. By building on the trend analysis and interviews' insights, the researchers could determine different user characteristics. The research of Kabanda [151] used focus group discussions and an experiment to evaluate artificial intelligence paradigms for network detection and prevention systems. Dupuis et al. [152] conducted a survey and interviews with Chief Information Security Officers (CISOs) to examine whether veterans make better cybersecurity professionals. The study of [153] provided an overview of the mixed methods approach as the third paradigm for research starting at the work of Campbell and Fiske [154], and continuing with the studies of Eugene et al. [155], Denzin [156], and Jick [157] for the developments on the notion of triangulation.

In essence, the mixed methods approach consists of quantitative and qualitative methods within the same research project [153]. Previous mixed methods studies have either used textual or small-sample (qualitative) data to interpret the results derived from large-sample (quantitative data) and used large-sample (quantitative data) to test the results derived from small-sample (qualitative) data [158]. This research aims to accomplish both.

Moreover, the three subquestions (SQ1, SQ2, and SQ3) lend themselves for a concurrent design. A concurrent design is when the methods are conducted more or less simultaneously, which is especially useful when the insights of the methods would help during the analysis of the other methods. As answering the subquestions requires different research methods, the following sections will explain those approaches. There, we will discuss the statistical method to uncover any different performance in the clean-up and how we use interviews based on the COM-B model to gain insights on the customer experiences after they receive the infection notification of the ISP. Ultimately, this study tries to determine the clean-up performance of the larger group of customers of the ISP who have been or have been infected with the QSnatch virus and what the experience of these customers was after they received the notification.

4.2 QSnatch clean-up effectiveness compared to non-persistent and non-IoT malware

This research will first try to measure the performance of the QSnatch clean-up and how it compares to the clean-up of other malware. Therefore, the first method of this mixed-methods approach is quantitative data analysis. The overarching aim of this research is to develop explanations for various aspects of human behaviour. Social and behavioural sciences, in general, are in the same pursuit, and data can show either consistency or inconsistency with such explanations [159]. This analysis will cover the first research question.

This part of the research will address possible differences in the clean-up of malware. It will analyse the success rate and the duration of the clean-ups. This research will divide the users from the abuse data up into two distinct groups. Those who have successfully cleaned their devices and those who did not. If no more malicious traffic is leaving from the customer's IP address, this does not automatically mean that the malware is removed from the device. There could be several explanations for this to happen, and one could only be sure by examining the device.

Therefore, this research will have to make an assumption on this and figures that users successfully cleaned their devices after some time. Inspired by the approach of Cetin et al. [36], there are three possible scenarios: 1) the user successfully performed the clean-up and did not show up in the abuse data for the rest of the study period, 2) the clean-up was successful, but the device got reinfected and started showing up in the abuse data within 30 days after the last reporting of the user, or 3) the clean-up was not successful, and the user shows up in the abuse data within 30 days after the last reporting of the user. For the sake of simplicity, this research will only differentiate between successful and not successful clean-ups. Thus, when a device becomes reinfected, this research would assume the clean-up was unsuccessful.

The research used a dataset containing malicious traffic from the network of the ISP's customers from May 15, 2020, to May 19, 2021. There were customers in the dataset that were considered infected at the time of the analysis, according to the 30-day rule of thumb like in the study of Cetin et al. [36]. What is also important to mention is that some cases do not appear in the data until much later. After all, the infections are not timed and thus will appear in the dataset at a different moment. To prevent that we follow some cases for much longer than others, we have to choose a certain period in which we analyse each case. That is why some of the cases that appear late in the dataset should not be included in the analysis. Ideally, you want a period that is as long as possible to get a good picture of the lifespan of the malware, but on the other hand, more and more cases are dropped that cannot be analysed long enough. Later in this research, in section 5.2, it will become clear that most of the interviewed customers tend to get their devices cleaned from QSnatch after 150 days. That is why we choose to use a time span of 150 days in this study. Since this research aims to unravel the differences between the remediation of malware, we had to come up with a method to approximate the time when those infected customers would remediate the malware on their devices.

4.2.1 Kaplan-Meier curves

Kaplan-Meier (K-M) curves deal with incomplete data on survival times (times-to-event) [160]. These curves are essentially determined by the survival function and the hazard function. De survival function is given as: $S(t) = P(X > t)$. In this equation, S is the probability that the malware is still on the device and is generating malicious traffic. The hazard function provides the probability that the malware will not survive for an additional time: $h_Y(y) = f_Y(y)/S_Y(y)$. In this formula, $f_Y(y)$ is the probability density function of survival time Y and $S_Y(y)$ is the survivor function (the probability of surviving longer than a certain time. The K-M Estimator is a non-parametric statistic used to make estimations about the survival function. The function is defined as: $^{S}(t) = \Pi(1 - d_i/n_i)$. d_i is the number of death events at a certain time t and n_i is the number of instances at risk of death just prior to that time. In K-M curves, times-to-events may be vital end-point variables such as cancer survival times, or more generally, time from diagnosis to death or time from treatment to relapse [161]. However, survival times does not have to relate to actual survival, with death being the event. Nonmedical disciplines also use K-M analyses, for example, subscription time of a newspaper's customers [160]. In the case of QSnatch infections, the deaths are a clean-up of a device. Therefore, this research considers a device that does not produces QSnatch-related traffic for 30 days to be clean.

In the survival curves, the starting point is also essential. Ultimately, we want to determine the duration of a particular infection, and we need the infection date in addition to the time of removal. These two determinations ultimately allow us to make a determination about the overall course and associated probabilities related to an infection on a device. Unfortunately, it is not always the case that these two values fall in the period of study. For example, some of the malware may still be on the device at the end of the study. This does not mean their clean-up will not happen in the future. But, we will not be able to determine it. Despite this limitation, we do want to determine the lifetime of malware. That is where censorship comes into play.

There are different types of data that require censorship. The most common type is right-censoring. Then future determinations cannot be made. As in the example just given, it is not yet clear when the malware will be removed but it is also possible that the device no longer participates in the study for whatever reason and it is no longer possible to determine whether the malware is removed from the device. Right censoring is taken into account in this study. Another form is left censoring. This form of censorship mainly concerns information that has taken place before the study. In the case of malware, one can think of the moment of the infection before the starting date of the analysis. Left censorship is not taken into account in this study.

It is then possible to plot the curves and compare them visually. They represent the probability of malware remaining on a device after a certain amount of elapsed time. To statistically prove the potential differences, this study used the Log Rank Test. This is a statistical method to compare survival functions between different populations. It compares the estimates of the hazard functions of two groups at each observed event time.

4.2.2 Kaplan-Meier curves on persistent, non-persistent, and non-IoT malware based on abuse data

This study analysed the remediation of persistent IoT malware such as QSnatch and VPNFilter. Torii is not taken into account as the ISP did not have data available on that malware. To analyse the remediation of those malware, this research compared those results to other malware types, i.e. non-persistent IoT malware and non-IoT malware. The research used a dataset subtracted from the abuse data provided by the ISP, which spans a time from 15 May 2020 to 15 May 2021. In the data, there were several malware to be analysed. Note that the data depends on the information provided by parties like Shadowserver. Therefore, the data does not contain all possible malware out there. Moreover, the ISP does not inform its customers of every infection but limits its efforts to a subset of those. The dataset contains

QSnatch and VPNFilter, so those malware are included as persistent IoT malware. However, as it turned out, the ISP does not notify customers infected with VPNFilter. So, it was not included in the K-M curves. Moreover, the set has also data on Mirai infections, which is a non-persistent IoT malware. Therefore, this malware is included as such. The ISP also informs its customers about a Mirai infection, and appendix D shows this notification.

The dataset contains different malware targeting Windows systems. This research analysed this last group initially with the 10 most common Windows malware in the data of the ISP. Those malware are Avalance, Bladabindi, Conficker, Gamarue, Necurs, Sality, Gozi, Sirefef, Caphaw, and ZeroAccess. At first, this research included these malware including customers that are notified and not notified about the infection by the ISP. Later, this research included only malware that the ISP does notify as it would not be fair to compare data of malware that are and are not notified by the ISP about the infection. If we would include both those categories, we would also measure the effect of the infected customer being notified rather than just measuring the stubbornness of the malware. Therefore, we decided to perform the analysis on a group of Windows malware that the ISP does notify. This second group consists of Ramnit, ZeroAccess, Kovter, Sality, Zeus, Downadup, Conficker, Tinba, Cutwail, Gozi, Nymaim, Nivdort, Gamarue, Necurs, Caphaw, Citadel, Qrypterrat, Sirefef, Pushdo, and Emotet. The customers infected by these malware all get the same notification. Appendix E.1 displays this notification.

As discussed before, although the ISP does try to provide notifications to all infected customers, there will still be, for whatever reasons, customers that did not get a notification. As this research is not interested in these cases, those will be filtered out. On the other hand, those cases could be interesting as they could be considered as a control group. The problem with this reasoning is that the ISP's system automatically sends a specific number of notifications per day. If a case of, for example, a QSnatch-infected customer does not receive a notification because the system already sent the maximum number of notifications for that day, the system will try to send a notification on the next day. This process goes over and over until the customer gets a notification. Therefore, the group will be based in some way because they are part of the process, although they seem to be a control group. In short, they are not excluded from receiving the notification but are part of the process of notifying customers, although they did not get a notification. On the other hand, it is intriguing to see if there are differences between customers that are infected but do not receive a notification for different types of malware as this would indicate a difference in prioritising of the ISP in notifying customers with different malware. We will compare the K-M curves of the groups persistent IoT malware, non-persistent IoT malware, and Windows malware from the notified group, i.e. the second group.

As it turned out, there are no BM customers that have gotten notifications about QSnatch. It is for this reason that it was not possible to include their performances in the notified section of the analysis. And while following the reasoning in the previous paragraph, we decided not to compare their results with the results of not warned CM customers. For this reason, the BM customers were not specifically addressed in this part of the research.

4.2.3 Data processing

The data that is being used is downloaded from the ISP's online abuse system as a csv file called 'comparisonwithwarnedwindows.csv' and ordered from the most recent case to the eldest case. The dataset can be provided with this report. And it contains several columns. Much of those are not relevant for this study. We will briefly explain those columns that are relevant:

- 'hashed_id': This is the hashed value of the client's id. This value has been hashed to ensure customer privacy. More on this is discussed in section 4.7. We use this value to differentiate between different customers. We specifically choose this value instead of the IP address because the IP address occasionally changes between customers. This does not happen often, but because we use a relatively long period for the analysis, it is wise to exclude this possibility. Otherwise, we might consider the wrong customer as infected.
- 'first_event_date': this column contains the moment that the infected customer first shows up in the data. We will consider this as the birth of the malware or the moment the customer's device is infected.
- 'last_event_date': This is the last time the customer shows up in the data. Under certain conditions that we will talk about in a moment, this value could lead to the determination that the device is no longer infected, or death of the malware.
- 'previous_state': We used this value to determine which customers have received a warning. At some point in the processing, we only keep the instances that are labelled 'WARNED' in this column, indicated that the customer is warned. The data also contains the column 'state'. Although it seems more obvious to use this column to determine which customers have been warned, the system only keeps track of the most recent updates of the cases. Therefore, the data does only contain several cases that at that point have the status warned. We do no use these cases for the analysis. We will substantiate this decision later in this section.

- ’malware_types’: This column indicates the malware that is associated with the generated traffic. We used this to group the different types of malware.

The following paragraph describes the programming process on how we came to the K-M curves, plots, and logrank test results. This is an explanation of the actual program provided in appendix H. To construct the K-M curves, we used the programming language Python and the packages pandas, numpy, lifines (KaplanMeierFitter, logrank test, and add_at_risk_count), and matplotlib. We first loaded the csv file into a pandas DataFrame. We then loaded the last_event_date and first_event_date as pandas date-times to enable analysis on them as they were initially stored as strings. Then, we determined the date of anlaysis or ’analysis_date’ by taking the ’last_event_date’ of the first row and made a copy of the DataFrame and for later analysis with and without warnings. Then, we dropped the cases that did not have ’WARNED’ as the previous state. We also dropped the duplicate hashed ids and kept the first appearing value of all the duplicates. In this way, we made sure we did not include the same customer twice. We also determined the duration per case by subtracting the first event from the last event and the moment of analysis by taking the first event of the first case.

Thereafter, the duration and the time between the first event and the date of analysis of every case is calculated. We then filtered out the cases that were shorter than 150 days in the dataset. We decided to analyse each case for 150 days and those cases could not be analysed long enough. We added a new column and calculated in it whether the death of a case has been observed. Before that, the difference between the start of the infection and the last day of the dataset had to be 150 or more and the case had not appeared for at least 30 days. Then the case was assigned value 1 in this column. In the other cases, the case was assigned a 0. Subsequently, the durations of the infections were rewritten to the data type numpy timedelta64. This is because the package of lifelines can only accept that data type. The code then filters out different types of malware combinations as those were only duplicate cases.

Subsequently, it creates DataFrames containing all the cases of each specific malware. The malware Downadup, Tinba, Nymaim, Sirefef, and Emotet are commended out and are not included in the analysis because the data did not contain any cases that meet the conditions that we imposed earlier. We then loaded the KaplanMeierFitter() function for each malware and plotted the K-M curve for all of the cases with the indications about when certain cases were at risk, censored, or events have happened. We also calculated the p-values resulted from the logrank tests. In the next cell, we grouped the malware as presented previously and also plotted their K-M curves as wel as the calculated the p-values resulted from the logrank tests.

In the next part, the code analysed the cases of the QSnatch-infected customers that were interviewed and compared those with the QSnatch-infected customers that we will describe in the next section, section 4.3. Thereafter, we again performed the steps of calculating the duration, time to last and first event, filtered out cases younger than 150 days, and determined whether an event has happened as we described earlier. Again, we create DataFrames for each type of malware and divide them up into persistent, non-persistent, and non-IoT malware but now we plot the differences between the notified customers and those who are not notified per type of malware group and calculate the p-values of the logrankTests.

4.3 User experiences after the notification

After gaining an understanding of how the clean-up effectiveness against QSnatch compares to other clean-ups, this research will examine the user experiences of the QSnatch clean-up after the customers get the notification from the ISP. This research aims to fit those user experiences into the factors of current literature. Interviews would best fit this pursuit. Kvale [162, p. 1] put it in the following words: *’Through conversations we get to know other people, get to learn about their experiences, feelings and hopes and the world in which they live. In an interview conversation, the researcher asks about, and listens to, what people themselves tell about their lived world, about their dreams, fears and hopes, hears their views and opinions in their own words, and learns about their school and work situation, their family and social life.’*

More specifically, this research will conduct semi-structured interviews. This type of interview incorporates open-ended and more theoretically-driven questions [163]. It allows data grounded in the participant’s experiences and constructs from the literature [163]. Since this research aims to find (in)consistencies between the user experiences and current literature on behavioural science, the specific approach of semi-structured interviews is suited.

Conducting interviews based on the COM-B model is not new in the literature. For example, researchers from different fields used this approach [143, 144, 145, 146, 147, 148]. The researchers used guidance from the work of Michie et al. [140]. That research provided a table discussing per factor of the COM-B model what needs to happen for specific behaviour to occur. The current research also used the framework of Michie et al. [140] to construct the interview questions.

4.3.1 Interview questions

The first step of the interview was to discuss the informed consent. This research got informed consent over the phone, but the respondents would have received the official form if they wanted to examine and sign it. Appendix G shows the informed consent form. Thereafter, the interviewer further informed the respondent about the infection date and the notifications whenever the respondent was interested in that information. The first question involved the physical opportunity and, more specifically, whether the customer received the email notification. If the customer did not receive the email, then it would not make sense to continue the interview. In such a case, the customer was asked whether he would like to have another notification send to the same or another email. When the customer agreed, he got the possibility to perform the steps and got a new invitation to participate in the research after he tried to perform the steps.

Then, it was interesting to see what the users did. Again, even though a customer stopped showing up in the abuse data, this does not mean he performed all the steps. Likewise, it could be that the customer kept showing up in the abuse data although he performed the steps. As discussed in subsection 1.1, the infected customer could perform the steps, perform them partially, or not perform any steps. The customers that performed different steps were considered to be in the last category. The remaining questions needed to be different between these possibilities. For example, it would have been strange to ask someone about his motivations to perform the steps when he did not perform the steps. Therefore, from this point, there are three different scenarios for the interview depended on what the customer did. Figure 5 illustrates the described process of questions.

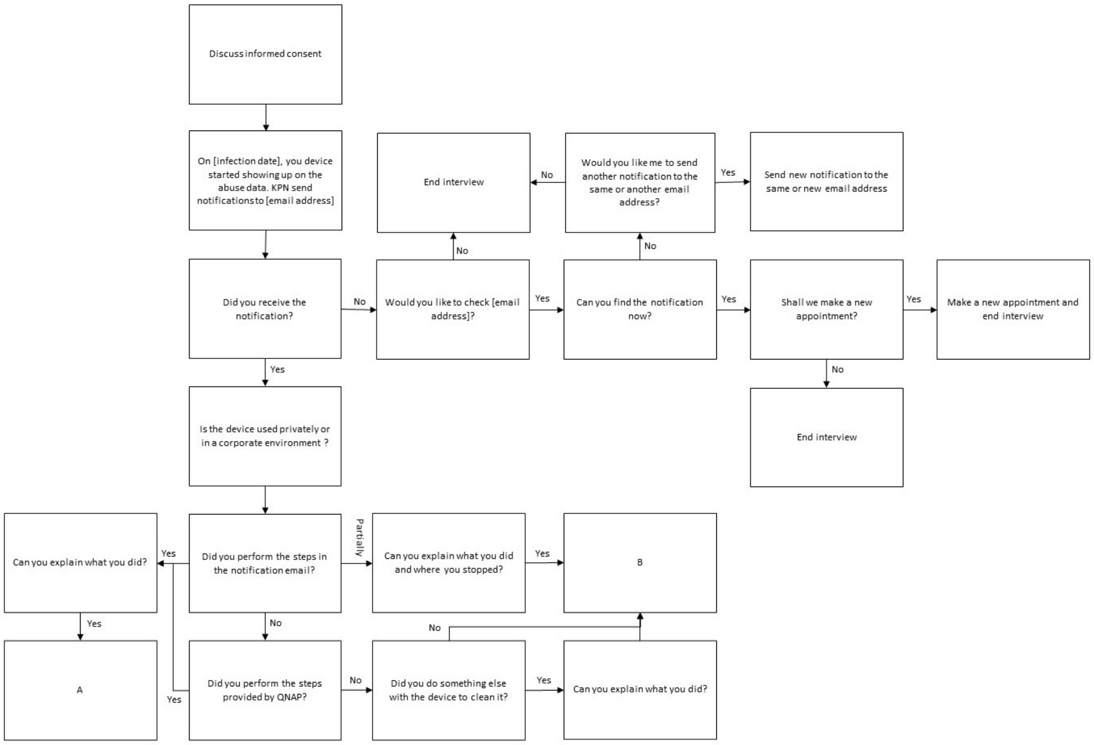

Discuss informed consent
On [infection date], you device started showing up on the abuse data. KPN send notifications to [email address]
End interview
No
Would you like me to send another notification to the same or another email address?
Yes
Send new notification to the same or new email address
No
No
Did you receive the notification?
No
Would you like to check [email address]?
Yes
Can you find the notification now?
Yes
Shall we make a new appointment?
Yes
Make a new appointment and end interview
Yes
No
Is the device used privately or in a corporate environment ?
End interview
Can you explain what you did?
Yes
Did you perform the steps in the notification email?
Partially
Can you explain what you did and where you stopped?
Yes
B
Yes
No
No
A
Yes
Did you perform the steps provided by QNAP?
No
Did you do something else with the device to clean it?
Yes
Can you explain what you did?

Figure 5: The beginning of the interview

For the customers that did perform the steps, Table 49 in appendix I provides an overview of the different questions relating to the concepts constructed in the section 3. The interview also included questions for those customers who have not or partially performed the steps. Table 50 in appendix I presents all of those questions.

In the following sections, we will explain the decisions that lead to the interview questions relating to the COM-B model. These interview questions are inspired by the work of previous researchers [144, 149, 145, 143]. We would, therefore, like to discuss those works first to make their questions more understandable. The work of Ekberg et al. [144] analysed barriers and facilitators to implementing family-centred care in adult audiology practices. Alexander et al. [149], on the other hand, investigated the barriers and enablers to delivery of the Healthy Kids Check (HKC). The HKC was a one-off health assessment aimed at Australian preschool children. The delivery of the service remained low and the researchers wanted to know why this was happening. The researchers of [145] analysed the barriers and facilitators to breaking up sitting time among desk-based office workers. Although their questions were in many cases highly suggestive and advocating reducing sitting time, we were inspired by some of their questions albeit we got rid of the suggestive phrasing. Finally, the work of Barker et al. [143] used the COM-B model to improve hearing-aid use in adult auditory rehabilitation.

For every question in this interview, there are follow-up questions depending on the response of the interviewee. For example, an interviewee could respond to the question *'do some people you know have a strong opinion on performing the steps?'* by answering *'yes'*. In that case, the interviewer will continue by asking: *'What is their opinion?'*. Figures 20 and 21 in appendix J provide the procedure and structure of the interview questions and follow-up questions in flow charts. Note that the question about whether the customer received the notification email is not included in the chart since the interviewer already asked that earlier in the conversation.

For physical capability, the research wanted to know whether the interviewees were physically limited to perform the steps. Like the research of Barker et al. [143] questioned, *'To create a plan with every patient, would you have to overcome physical limitations e.g. get around problems relating to disability?'* Inspired by this question, the first question based on the COM-B model was: *'Did you have any physical or bodily limitations that made the steps challenging?'*. Measuring psychological capability is a bit trickier and cannot be conducted by asking just one question such as; *'Were you smart enough to perform the steps?'* This question would lead to biased results as respondents would not likely indicate that they are not smart enough to perform the steps. Moreover, we wanted to measure the capacity to engage in the necessary thought processes - comprehension and reasoning. The component psychological capacity is built up from four subcomponents: comprehension, skills, knowledge, reasoning. This required several questions. The research measured the psychological capability with five different questions. Together, the answers to these questions formed the indication for the component capability. The first question out of the five was inspired by the question of Ekberg et al. [144]: *'How easy or difficult is it to do a HKC?'* convinced us to question: *'Did you find the steps challenging?'*. The second and third questions were inspired by the work of Alexander et al. [149] as they questioned: *'Do you know about the mandatory and non-mandatory components of HKCs?'* and *'Do you know about the RACGP guidelines for child preventive health?'* Likewise, Ojo et al [145] questioned: *'Can you start by telling me your understanding of current advice by experts about how much sitting time is okay?'* We wanted to examine whether the respondents had any practical knowledge on the subject and subsequently asked whether they knew what are malware and persistent malware. The question *'How have you learned how to do a HKC? Have you had any training for HKCs?'* of Ekberg et al. [144] inspired us to question the respondents' previous experience in IT.

Reflective processes were measured by asking the respondents whether they found the steps useful, what they thought would happen if someone does not follow the steps, and whether they thought they are responsible for following the steps. We decided to use these questions as the answers of the respondents would trigger the respondent to share his thoughts on the steps. If they found the steps useful, then we wanted them to reason about why they thought the steps are useful. Next, if they could not reason about why they thought the steps are useful, we wanted them to reason about what would happen if someone does not follow the steps. In this way, we could measure the reasoning behind how the customers estimate the usefulness of the steps. Comparably, Ojo et al. *'What do you think are the consequences of sitting for long periods?'* And when the customer found them useful or not, it could be that they thought someone else should perform the steps, for example, the ISP or QNAP. We wanted to make sure this possibility is covered in the interview, so we included a question about it. Subsequently, the respondents were asked what they felt while performing the steps and whether an impulse helped them perform the steps to measure the automatic processes. While this study questioned *'Did any feelings help you perform the steps?'* and then prompt different feelings if the respondent was not sure what to respond to that question, the study of Barker et al [143] took a more direct approach. Their study questioned: *'To create a plan with every patient, would you have to feel you want to do it enough e.g. feel more of a sense of pleasure or satisfaction from doing it?' 'To create a plan with every patient, would you have to* and *'To create a plan with every patient, would you have to feel that you need to do it enough e.g. care more about the negative consequences of not doing it?'* The questions on these answers together provided the measure for the component motivation.

Moreover, measurement of the physical opportunity was inspired by the work of Ekberg et al. [144]. That research questioned whether there were factors in the respondent's environment that were facilitating or hindering the preferred behaviour. To get a more valid answer to this question, we have chosen to split up the question and make it more concrete. We split the question into four questions and focused on the clean-up of QSnatch. We questioned whether they received the notification email (already asked at the beginning of the interview), whether they used any tools, whether they had enough time to perform the steps, and whether the location of the device was a problem to access the device. Already at the beginning of this research, we argued that it could be challenging for users to know whether there is malware on their device. In addition, respondents may have used other tools to perform the steps or used other tools instead of the steps. That is why we asked about it. Comparably, the research of Alexander et al. [149] and Barker et al. [164] questioned whether the respondents had the right equipment or necessary materials. At the same time, it might be the case that the respondents wanted to carry out the steps, but other tasks were more important and that the implementation of the steps was therefore postponed. This can be an issue, especially in business environments. Finally, the devices are not tied to a physical location because they are connected in the customer's network. As a result, the device may be in a place that is difficult to reach. That is why this question was asked to rule out that this did not lead to problems. To measure social opportunity, this research questioned whether any people helped the respondents perform the steps and if they knew people that have a strong opinion about performing the steps. With social opportunity, it is about the cultural environment of the respondent and how that directs the way an individual thinks. For this sub-component, too, we have chosen not to ask directly about the cultural environment but whether they have received help and whether there are people who have a strong opinion. If there are people with strong opinions, this can be evidence of a cultural environment that has influenced the respondent's actions. Together, the answers to the questions of these sub-components formed the indication for the component opportunity.

Thereafter, the interview will discuss demographics such as the gender and age of the customer, the brand and type of the device, and the place where the device has been bought. Finally, the interview ends by questioning whether the customer has anything to add that was not covered in the interview so far. By asking this question, it can be prevented that the thinking process is limited by the theoretical model and provides the opportunity for the customers to include anything that seems essential to them.

4.3.2 Sampling and procedure

Although most researchers advocate stopping data gathering only after theoretical saturation, i.e. the process of continuing to sample relevant cases until no new theoretical insights are gained, Baker [164] advises aiming for a medium-size subject pool of 30 respondents. This pool would penetrate beyond a minimal number of people without resulting in endless data gathering [164]. However, this research used a sampling procedure that we will explain later and kept pursuing with this strategy until all customers in the dataset had at least one opportunity to pick up the phone.

In the abuse data, there are customers who stopped and those who kept showing up. Comparable to the Quantitative Data Analysis in section 4.2, this research would also consider those who have successfully cleaned their devices as those who stopped showing up in the data for minimal 30 days as one group and considers their cases as remediated. The other group are those who kept showing up in the data and are considered not remediated.

This part of the study included customers from a database from April 5, 2020 until April 21, 2021. The data contains a different period than the one used for the analysis to determine the results of the clean-up. That is because the data to make these determinations comes from the ISP's systems. These systems do not store the information about the infected customers forever, and over time, the data will be deleted again. The incremental nature of this research has ensured that the data for the different forms of research were not extracted from the system on exactly the same day. That is why there are differences between the start and end dates of the datasets. This also resulted in different numbers of cases between the analysis of the clean-up results and the user experiences.

In that time period, there were in total 189 customers who were not already contacted in the pilot study. More on the pilot study can be found in section 4.3.3 and all of them got a call. Of those customers, 175 were considered remediated and 14 not remediated. A total of 58 customers wanted to participate in the interviews. Fifty-five respondents were considered remediated and 3 not remediated. To find out whether those interviewed customers were representative for the wider population of the ISP's customers infected with QSnatch, this research included an analysis on K-M curves of those groups and performed a logrank test.

An overview of the contacted individuals is presented in table 2. In this table, we successively see how many customers were seen as remediated or not remediated, how many did not answer the telephone, how many answered the telephone but did not want to participate in the survey, how many did participate and finally, how many the telephone number was not in use. It becomes evident that most customers who were considered not to have remediated their devices were enthusiastically included in the pilot. This left less not remediated customers to be potential respondents for the official interviews.

Moreover, the table shows that for the pilot study as well as the official interviews, around one-third of the respondents did not pick up the phone, around one third did not want to participate, and around one third wanted to participate. If a customer indicated that he did not have time at that moment, he was offered to call back at another time. If the customer did not answer his phone at the time of the appointment, it was still counted as not answering the phone. A number of customers indicated that they did not have time at that moment, but also did not want to meet at another time that was convenient for them. This research assumes that they had no desire to participate.

Every working day, the selected customers got an email informing them about the research, including an invitation for an interview on the next day or Monday if the notification was sent on Friday. On which days the customers received their invitations and when each interview took place can be found in appendix L in table 51. That overview also shows the days on which malicious traffic was discovered from the network of the respondents. It should also be mentioned that there could be some time between the last time that the customer is shown in the data and the interview. As a

	remediated	not remediated	*total*	not answered	not participated	participated	not in use	*total*
official	168	14	*182*	62	57	58	5	*182*
pilot	20	25	*45*	14	14	14	3	*45*
total	*188*	*39*	*227*	*76*	*71*	*72*	*8*	*227*

Table 2: Customers contacted for the interviews

result, some respondents have forgotten certain parts of their clean-up experience. When this has been the case, specific attention will be paid to them in section 5.

The selection procedure is shown in Figure 6. This selection procedure was used to select and subsequently invite customers to the interviews. Customers that still showed up and those who stopped showing up got a different invitation. Customers whom we suspect based on the data will receive a message stating their expectancy and an invitation to participate in an interview to discuss their experiences. In the not remediated group, we indicate that we suspect that the device is still infected and that we would like to know exactly what happened and whether they are interested in participating in the research in the form of an interview. Appendix F shows both emails and their translations to English.

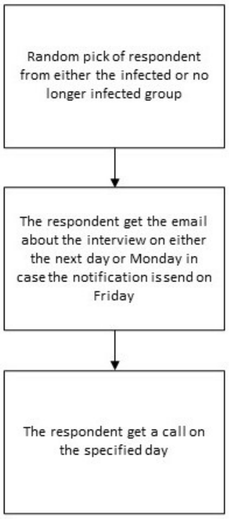

Random pick of respondent from either the infected or no longer infected group
The respondent get the email about the interview on either the next day or Monday in case the notification is send on Friday
The respondent get a call on the specified day

Figure 6: Respondent selection procedure to participate in the interviews

Then, the customer got a call. At the beginning of that call, the interviewer checked whether he had the right person on the line, provided a description of the research, and asked whether the customer wanted to participate in the research. The flowchart in Figure 7 provides this procedure.

After a few days, no new insights emerged from the interviews. Respondent 53 was the last respondent that presented new insights on the clean-up experiences of the infected customers. However, we decided to keep doing 5 more interviews. That was convenient because at that time, all the customers of the ISP that had ever been infected had been contacted. So, we called them at least once. In principle, we could have called all customers again who had not answered the phone, but that was not necessary because saturation was reached.

4.3.3 Pilot study

Previous research described the importance of a pilot study [165, 166]. Two of the arguments in favour of a pilot study are to determine the feasibility of the study protocol and testing the measurement instrument [165]. This research would also need to grasp the feasibility and test the measurement instrument.

For this pilot, this research contacted 45 customers on April 8 and April 9, 2021; 25 who did not show up in the records anymore and 20 who still did. On April 8, 10 customers got a call and on April 9, 35 customers. 28 of those people answered the call, and 14n wanted to participate in the research. Out of those people, 14 participated in the study; 7 who did not show up in the records anymore and 7 who still did. On average, the interviews took six minutes and fifty-one seconds.

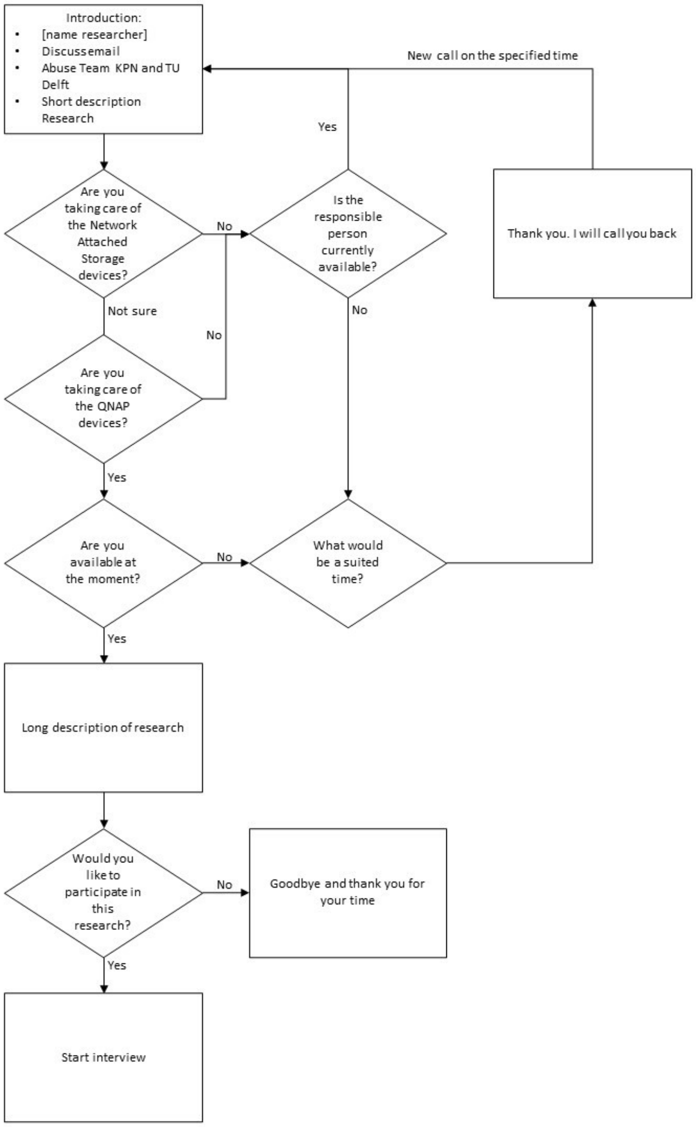

Introduction:
• [name researcher]
• Discuss email
• Abuse Team KPN and TU Delft
• Short description Research
New call on the specified time
Yes
Are you taking care of the Network Attached Storage devices?
No
Is the responsible person currently available?
Thank you. I will call you back
Not sure
No
No
Are you taking care of the QNAP devices?
Yes
Are you available at the moment?
No
What would be a suited time?
Yes
Long description of research
Would you like to participate in this research?
No
Goodbye and thank you for your time
Yes
Start interview

Figure 7: Call procedure

As a result of the pilot, this research included a question of whether the device is used for a business or in a private setting. As described earlier, it was not possible to find out the user experiences of BM customers as they did not receive notifications, and the Abuse Team did not have the right contact information to invite them for the interview. It was, on the other hand, interesting to see whether there are differences between the user experiences of customers that use their devices in a private of business setting. However, based on the ISP's data, it was not possible to determine what are the purposes of the devices. Another finding of the pilot was that some respondents told about their IT experience, although this was not questioned. At the same time, it is an essential part of the psychological capability of the respondent. This was not covered in the pilot but added for the following interviews.

4.3.4 Transcribing and Coding

All of the interviews were recorded and textual versions were created. One coder then analysed the documents to find key phrases and clarify patterns and relationships between patterns with the use of data analysis software program Atlas.ti as suggested in the work of Cope [167].

During the coding process, several codes have been created. The following paragraphs elaborate on the coding process more in-depth and go into which criteria are decisive for getting coded in certain ways. In many cases respondent 26 is labelled as 'not discussed'. The respondent indicated that she had little time for this interview, but could briefly explain what had happened. The interviewer tried to extract as much information as possible from this respondent, but many questions were not addressed.

The first code is ’received’ and it indicated whether the respondent told us he received the notification mail from the ISP. This research differentiated between respondents who indicated who did receive the notification and did not indicate as such by labelling those ’received: yes’ or ’received: no’, respectively. If a respondent explained how he was triggered to perform the steps other than the notification of the ISP, and claiming he did not remember getting a notification, then the case would also be labelled as ’no’.

Then, the respondents indicated whether the device is used for corporate or private purposes. The corresponding codes are ’use: corporate’ or ’use: private’. As it turned out, several respondents use their QNAP devices for both corporate and private purposes. Those were labelled ’use: both’. Other respondents indicated they mainly use their devices for corporate purposes and a little bit privately or vice versa. In those cases, this research considered them as being used for both purposes.

The respondents were also asked whether they performed the steps in the notification email. The respondents either that they did perform the steps or did not perform the steps with the codes ’performed: yes’ and ’performed: no’. Obviously, there are more possibilities that the respondents could have responded with. We coded the answers of the respondents that updated the firmware, installed the malware remover tool, or those who did both as ’updated’, ’installed’, or ’updated and installed’. Those respondents that executed the steps of QNAP were labelled as ’steps qnap’. Other respondents simply turned off their devices and this research coded those as ’turned-off’. This research used the code ’someone else’ for those respondents that got help from someone else. Note that we will explain in more detail who they specifically got help from during the code ’help’ later in the section. One respondent indicated that he did no longer remember what he did and this case was labelled ’do not remember’. Then, there were also additional codes for customers that indicated they tried but failed to complete the steps, those who performed an extra step, and those who experienced setbacks. Those cases were labelled ’tried but failed’, ’extra step’, and ’setback’, respectively. These codes came on top of the codes we mentioned above.

To address the steps the respondents took to clean their devices, this research asked the respondents to elaborate on them. The responses on these lines of questioning were coded with ’done: [what the respondent had done]’. As such, if a respondent would indicate that he had updated the firmware, installed the malware remover, or a combination of both, those cases were labelled as ’done: updated’, ’done: installed’, ’done: updated and installed’. Then, if a respondent would indicate he tried the steps but failed or performed the steps by QNAP, those would be coded as ’done: tried but failed’ and ’done: steps qnap’. Moverover, statements were coded as ’done: extra step’, ’done: setback’ or ’done: someone else’, if the respondent performed extra steps, run into setbacks, or let someone else perform the steps. Note that someone else could be a friend or acquaintance, an IT professional, or employees of QNAP. This research will explain more about who specifically helped the respondents in section 5.2.9. Finally, if a respondent did not remember what he had done, then this research coded his statement as ’done: do not remember’.

The following questions were based on the COM-B model and also the coding is based on this model. More specifically, we used the interview questions to come up with the codes. For example, if a respondent provided an answer to the question ’Do you knew what malware is?’ by confirming that he did and providing the right answer, this was coded as ’malware: yes’. A more in-depth explanation will be provided in the upcoming sections, but this provides an illustration of the coding process. The first question concerning the COM-B model was about the physical capability. The respondents were asked whether they had any physical limitations that prevent them from performing the steps. The responses of the respondents were either coded ’physical: yes’ or ’physical: no’, indicating their either had or had not physical limitations. If there were some respondents indicating they had limitations, that would also have been discussed. However, as it turned out, none of the respondents had physical limitations.

Next, the respondents indicated whether they understood the steps and their responses were coded ’understand: yes’, ’understand: to a certain extent’, and ’understand: no’. These codes pretty much speak for themselves. On the other hand, some responses were coded as ’understand: did not look’ in case someone did not look at the steps and ’understand: not remember’ in case someone did not remember whether they understood them.

When the respondents did not find the steps challenging, their responses were coded as ’challenging: no’. Likewise, if a respondent indicated to find the steps challenging, those cases were labelled as ’challenging: yes’. Thereafter, the respondent would discuss their reasoning for why they did or did not find the steps challenging. Those responses were coded as ’challenging: argumentation’. In one case, the respondent did not remember what he did and this was coded as: ’challenge: not remember’.

The respondents then estimated the time it took to perform the steps. Those estimations were labelled as ’duration: ’ followed by the time they thought it took them to indicate that the respondents estimated the duration of performing the steps. Respondents provided answers of varying sizes and units. Some also gave an estimate of a time span. We chose to express all indications in minutes and chose the largest value if someone indicated a time span. Moreover, some respondents indicated the time in days. In those cases, we assumed the respondents did not work on it from the early morning until late in the evening but we figured that they would approximately have spend a working day of 8 hours per day they indicated.

Moreover, respondents shared a little bit of their knowledge about the topic and elaborated on their understanding of malware. If a respondent indicated that he did not know what malware is, then this was labelled as 'malware: no'. Many respondents, on the other hand, indicated that they knew what malware is. However, this research also asked the respondents to elaborate on their understanding and their responses were only coded 'malware: yes' if they roughly provided the correct definition. According to NIST, malware is 'Hardware, firmware, or software that is intentionally included or inserted in a system for a harmful purpose'. This research, therefore, considered respondents that explained malware while using the terms hardware, firmware, or software, and harmful, malicious, or harmful intent. Moreover, this research also considered that the respondent would understand what malware is if he would elaborate on a specific harmful purpose. For example, one respondent argued that malware is software that encrypts a victim's data and then asks for a ransom and was labelled as 'malware: yes'.

This research also asked the respondents to share their understanding of the difference between persistent and non-persistent malware. Those people who indicated that they knew what persistent malware was were only coded as 'persistent: yes' if they elaborated on that persistent malware, compared to non-persistent malware, is stored on a different part of the device and will not be removed after a reboot of the device. Some respondents translated the word persistent and elaborated that the persistent malware would be harder to remove or that one would rather want non-persistent on his device were coded as 'persistent: no' because they essentially did not know what the difference is. Respondents that indicated that they did not know what the difference is were also coded as 'persistent: no'.

When the respondents were asked whether they thought they could perform the steps, one group indicated they thought so and were labelled 'self confidence: yes'. Another group said they did not think so and were labelled 'self confidence: no'. Others no longer knew or indicated that they could not give a meaningful answer. Their responses were labelled 'self confidence: other'.

Then, the respondents elaborated on their experience with IT. If the respondent would indicate to have work experience, had education in IT, or both, those cases were labelled as 'experience: work', 'experience: education', 'experience: work and education'. In cases when the respondents were talking about that they build websites and do networking but did not specifically address that it is their profession it was coded as 'experience: networking'. Moreover, the instances in which the respondent would indicate that IT is a hobby for him, it was coded as 'experience: hobby'. Finally, if the respondent would indicate not to have any specific experience, this research labelled it as 'experience: no'.

The respondents also discussed their views on whether they thought the steps are useful. Respondents that thought they were useful, were not useful, or indicated they could not provide a meaningful answer were labelled 'useful: yes', 'useful: no', or 'useful: not sure'. If a respondent would indicate that he found the steps limited or useful to a certain degree, those were labelled as 'useful: limited'. If the respondents initially did not give a detailed answer, they were specifically asked to explain their opinion. Their responses were coded as 'useful: argumentation'.

This research was also interested in what the respondents thought would happen if someone does not follow the steps. The respondents argued that there would be harm to the user of the device, the malware would spread to other devices, it would harm the network with illicit and overloading traffic, and the customer's network connection would be shut off. The responses were according coded as 'not follow: harm to user', 'not follow: spread', 'not follow: harm to network'. Some respondents indicated that would happen multiple things. As such, there were respondents indicating that the malware would do harm to the user and to the network, and that the malware would harm the user and spread to other users. Those cases were labelled as 'not follow: harm to the user and network' and 'harm to the user and spread'. Other respondents indicated that the malware would stay on the device, nothing would happen, or they did not know. Those were labelled 'not follow: malware stays', 'not follow: nothing', and 'not follow: does not know'.

Another factor questioned during the interviews was whether the respondents thought they were responsible for performing the steps. If a respondent would indicate that he thought he was responsible for performing the steps, than this research would code it as 'responsible: yes'. Others that thought they were not responsible were coded as 'responsible: no'. A final response was coded differently, namely if a respondent would indicate that he thought he was responsible if he was notified or at least aware of the infection. Those were coded as 'responsible: yes, if you know'.

The feelings of the respondents were very different. If a respondent did not feel anything in particular, this was coded as 'feelings: nothing'. If the respondent would discuss a particular feeling that would be coded as 'feelings: [specific feeling]'. For example, if someone felt angry, this was coded as 'feelings: angry'.

If an impulse helped the respondents to perform the steps, this was coded as ’impulse: yes’. Initially, the coding was more in-depth and also differentiating between what the respondents were specifically describing as being the impulse. Later, the research did not differentiate between those specific impulses specific impulses were only mentioned at maximum twice. For example, two respondents indicated that the impulse was to get rid of the device, but those were both coded as ’impulse: yes’. An exception is the code ’impulse: email’. In those cases, the respondent indicated that the email was the impulse. Many respondents argued about it being an impulse but this research elaborates more on that in the results section as well as about the specific impulses. If a respondent indicated that no impulse helped or that they did not know, those cases would be coded as ’impulse: no’ and ’impulse: not sure’, respectively.

Thereafter, the respondents would discuss whether they had enough time to perform the steps. If the respondents indicated they had, the research labelled those cases as ’time: yes’. On the other hand, if a respondent did not have enough time, this was labelled as ’time: no’. Finally, if a respondent would indicate that he did not remember, this was labelled as ’time: not sure’. On some occasions, the respondents would also provide a reasoning behind their view. This research labelled those parts ’time: argumentation’.

The physical location of the device could have been a reason why respondents had difficulties while performing the steps. If a respondent indicated as such or not this was coded as ’location: yes’ or ’location: no’. Thus, if the location was not a problem for performing the steps, it was labelled ’location: no’.

When the respondent would indicate that he had not had help while performing the steps, this would be coded as ’help: no’. Otherwise, if the respondent did got help from someone, it was labelled as ’help: [the helping person]’. This research differentiated between different types of people that helped out. A number of respondents got help from an IT professional, a friend or acquaintance, or an employee of QNAP and those instances were coded as ’help: IT’, ’help: acquaintance’, and ’help: qnap’, respectively.

Then the research would move on to the question of whether the respondent knows any people with a strong opinion on performing the steps. The respondents indicated they did know, did not know, or did not talk about it with others. Those were coded as ’opinion: yes’, ’opinion: no’, and ’opinion: not talk’. In case the respondent stated that he was not sure, then this was coded as ’opinion: not sure’.

The answers to the questions on demographics and the type and location where the device has been bought were coded as ’gender: [gender]’, ’age’, ’type’, and ’shop’. The responses on gender were coded as ’gender: male’, ’gender: female’, or ’gender: -’ if someone did not want to answer that question. If the respondent still knew at which store he bought the device, this was coded as ’shop’. If the device was bought from another person, it would be coded as ’shop: person’. On the other hand, if the respondent did not remember where he bought the device, the response was coded as ’shop: not sure’.

The previous codes were following from specific questions discussed in the previous sections. However, as became evident during the interviews, the respondents also touched upon insights that were not specifically addressed by the interviewer. Those were divided up into expressions and suggestions. If a respondent expressed that he was happy with the notification or gave his compliments on the steps, this was coded as ’expression: happy with the notification’ and ’expression: nice steps’. If the respondent would indicate that he did not trust the notification at some point, this was labelled as ’expression: distrust’. If a respondent himself about the sanction of disconnecting the customer, it would have been labelled as ’expression: sanction’. In other cases, the respondents expressed hope, motivations for performing the steps, and opinions about other ISPs. Those are labelled as ’expression: hope’, ’expression: motivation’, ’expression: other ISPs’.

Moreover, if the respondent would suggest something to improve the process of notifying customers or the steps, this would be labelled as ’suggestion’. If the respondent made an observation, this would be coded as ’observation’. This research considers an observation when the respondent had observed something on his device, on the Internet, or of something else. Specific instances of those groups are discussed in section 5.

At the end of the interview, there was also the possibility to ask questions. Respondents made frequent use of it as well as, in some occasions, during the interview. The questions were coded as ’question: [subject]’. Thus if a respondent asked a question about how the ISP would determine whether a customer is infected, it would be labelled as ’question: determination’.

4.4 Interviewed customers and abuse data

Before participating in the survey, the respondents were, like the others, QSnatch infected customers of the ISP. This meant that there is data on when they were infected, on what days they popped up in the data, on which days they were notified, and on which day they were last seen. In Table 51, we included the dates on which every respondent was detected and when they were warned or notified by the ISP. We also added the date on which the respondents were invited for the interviews and when they got interviewed. In Table 52, we summarised the data on which the respondents showed up to periods of infection.

Based on this information, it was possible to determine the time in between the first time the respondents showed up in the abuse data and when they got notified as well as the time from the first notification to the day that the customers were last seen in the data.

Unfortunately, the data of respondents 12, 13, 43, and 44 were no longer available. This part of the analysis was gathered relatively late compared to the other datasets. The ISP deletes abuse data after one year. So, it could be that the respondents were infected a long time ago en deleted for that reason. Another explanation for the missing data could be that the respondents were no longer customers of the ISP and switched to another provider.

4.5 Demographics

This section covers the demographics of the respondents. First, figure 8 shows a Kernel Density Estimation (KDE) plot of the ages of both the interviewed customers and the population of the ISP as a whole. The mean of the ages of the interviewed customers is 50.2 years, with a standard deviation of 10.33. The mean of the wider population is 53.75 years with a standard deviation of 17.21. It becomes evident that the means of both populations are not far apart, however, the ages of the interviewed customers is a lot more centralised. Welch's unequal variance t-test estimates a value of 2.55 with a probability of 0.014. With an α of 0.05 we reject the hypothesis that the means are equal and, therefore, the group of interviewed customers is not representative for the wider population of ISP's customers when it comes to their age.

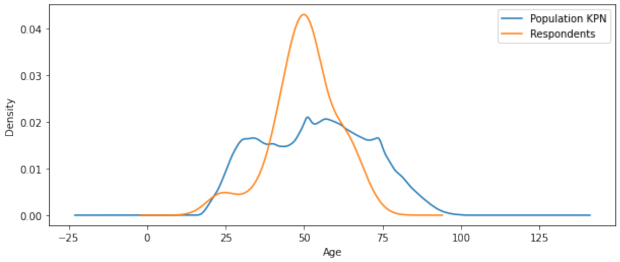


Figure 8: The respondents' and the ISP's wider population's ages

Subsequently, it becomes evident from figure 9 that the respondents were primarily male. Only three of the respondents were female, and one did not want to answer the question. In the wider population, this is much more equally distributed. The N-1 Chi-squared test with Yates correction test gives a value of 27.9599 with a corresponding p-value of < 0.00001. With an α of 0.05, this research concludes that the repondents are also not representative for the wider customer population of the ISP when it comes to their sex.

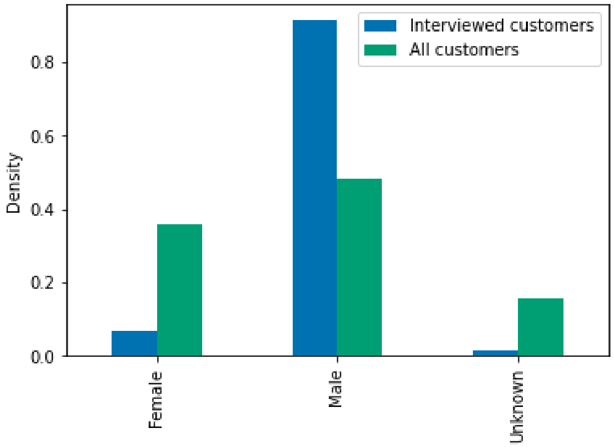


Figure 9: The respondents' and the ISP's wider population's sex

If we take these findings together, we must conclude that the selection of respondents is not representative for the larger population of ISP's customers. For this reason, we cannot generalise the findings of the interviews to the wider population. We can, however, hypothesise that the characteristics we found are representative for the population of NAS users, or even more specifically, QNAP users. But we lack the data from this population to draw conclusions.

4.6 Devices with business and private purposes

As described earlier, it was not possible to contact the BM customers but we could ask the respondents being CM customers, whether they used their devices for business or private purposes. Based on the answer of the respondents, we got three groups of purposes for the devices: 'private', 'business', and 'both'. As will be shown later, these groups consisted of 45, 5, and 6 respondents, respectively. The interviewer did not ask 2 respondents this question. That was the case for respondents 28 and 53. In total, that made 58 respondents.

Then, we combined the groups of 'business' and 'both' and thus grouped all the respondents with devices that are used for at least business purposes. In this group, there are respondents that also use their devices for private purposes but business purposes anyway. This group is called 'business/both' and consisted of 11 respondents: respondent 1, 4, 6, 8, 12, 16, 25, 27, 40, 54, and 57. We have created tables and each lines shows one of the possible groups: 'business/both', 'private', and 'not discussed'. The columns represent the coded answers the respondents gave. For both the rows and columns, we added 'total' to check if the numbers are correct. During most questions, there was also at least one respondent that was not asked the question. In most cases, that was respondent 26. In any way, respondents 28 and 53 were grouped as 'not discussed' and got a specific row assigned to them and left out of the column 'not discussed'.

We decided not to perform statistical analysis on these groups but rather show the parts of the groups that gave a specific answer to the questions and their contribution to the group as a whole. We have chosen this because we do not know how large the groups of private and business purposes are in the total of the ISP's infected customers. In addition, we opted for a definition in which the respondents themselves had to indicate to which category they belong, and we do not know which factors may play a role. For example, it could be that some respondents would not have dared to indicate that the device was used for business purposes, because they did not want to cause reputation damage to their company or because they use a CM subscription for a business and might be afraid that telling the ISP would lead to trouble. Likewise, we do not know when the customer considers the purpose of a device to be for business. It could be that a customer only used the device for business purposes once years ago, deleted those files, and still considers the device being used for business purposes. Therefore, we will not attempt to generalise the results on this part of this study to a broader population of the ISP's QSnatch-infected customers or statistically prove differences between the results of 'business/both' and 'private' purpose cases.

4.7 Ethical Considerations

During the research, it was possible to associate infections with specific customers. A lot of personal data was available from these customers and we have taken measures to prevent their privacy from being compromised. For that reason, the data is only handled on a designated laptop borrowed from the ISP. It was decided not to remove the data from the laptop for the analysis of this study. In order to still share the data, IP addresses have been removed from the data. An extra column has been added to still be able to determine whether the data concerned customers who also participated in the interviews. Finally, the client ids are hashed with SHA-2. Each interview has been given a number as references to the interviews. On the basis of those numbers together with information that is on the laptop, it is possible to trace specific customers. This is not possible without that information.

4.8 Summary on the methodology

In this chapter, we laid out the foundation on how we analysed the clean-up results, user experiences after the customer received the infection notification, and the differences between the experiences of customers that use their devices for private or business purposes. This chapter explained that we used K-M curves and logrank tests to compare the clean-up results of different malware. It provides the specific data processing, analysis, and code on how we come to our results. The chapter also elaborated on how we grouped the types of malware in persistent IoT, non-persistent IoT, and Windows malware and compared the results of those groups. We then explained more about the user experiences after the notification and that we wanted to analyse those based on interviews with the infected or previously infected customers. Subsequently, we showed how we derived the interview questions from the COM-B and how we invited the customers to the interviews. Before the official interviews, we first conducted a pilot study and made some adjustments. The chapter also elaborated on the transcribing and coding process that took place after the interviews were done. It also discussed the demographics of the interviewed customers and concluded that the group is not representable for the wider population of the ISP's customers. We then explained how we differentiated between the customers who have used their devices. Finally, we discussed the ethical considerations and how we made sure the privacy of the ISP's customers is guaranteed. With all this information, we know enough to move on to the next chapter. In that chapter, we will discuss the findings on the research we conducted as described in this chapter.

5 Results

This chapter discusses the results of the methods described in the previous section. The chapter will start elaborating on the results of the clean-up performance of QSnatch and how it compares to other types of malware. In that pursuit, it will provide the answer to research question 1. Thereafter, it will go into the user experiences after they received the notification. In that section, we will be able to provide the answer to research question 2. As we will not be able to differentiate between CM and BM customers as we could not involve those groups of customers in the research, we will provide indications to differences between the user experiences between customers that use their devices for business or private purposes. Therefore, that section will also provide the answer to research question 3.

5.1 Malware clean-up results

This section elaborates on the performance of the botnet remediation of QSnatch compared to other malware such as non-persistent IoT malware and Windows malware. This section will answer the first subquestion. The differences between persistent and non-persistent malware and Windows malware that are being notified are analysed. In this section, we will provide the answer to research question 1.

5.1.1 Comparison between the clean-up performance of notified customers infected with Persistent, Non-persistent IoT malware, and Windows malware

Figure 10 presents the curves of the different types of malware. The curves indicate that QSnatch is one of the most stubborn malware. In the period from 0 to around 25 days, Caphaw, Gozi, and Qrypterrat all have a higher probability of survival but their survivability quickly drops after this period. Only after 120 days QSnatch gets surpassed by Gozi again. In the period of 25 to 120 days QSnatch has the highest probability of survival. However, it is challenging to come to harsh conclusions based on this figure. For example, the curves of Caphaw, Citadel, and Ramnit are based on respectively 1, 2, and 3 cases. Moreover, there is a lot of uncertainty in the figure. Figure 22 in appendix K also shows the number of corresponding cases at risk, censored, and events.

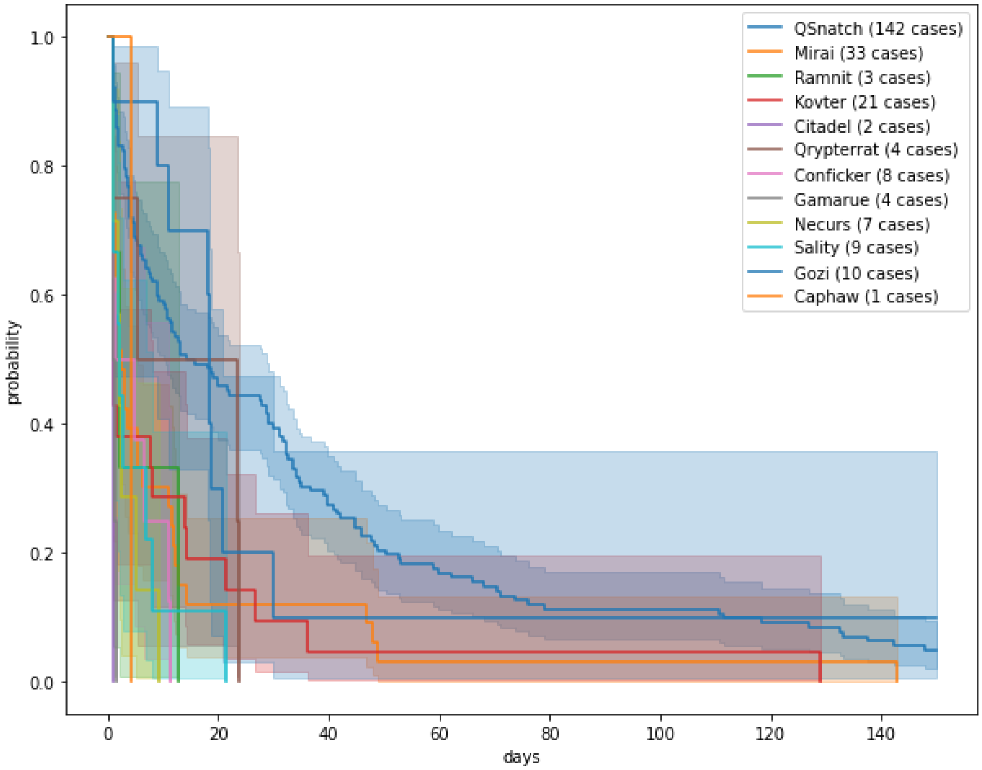
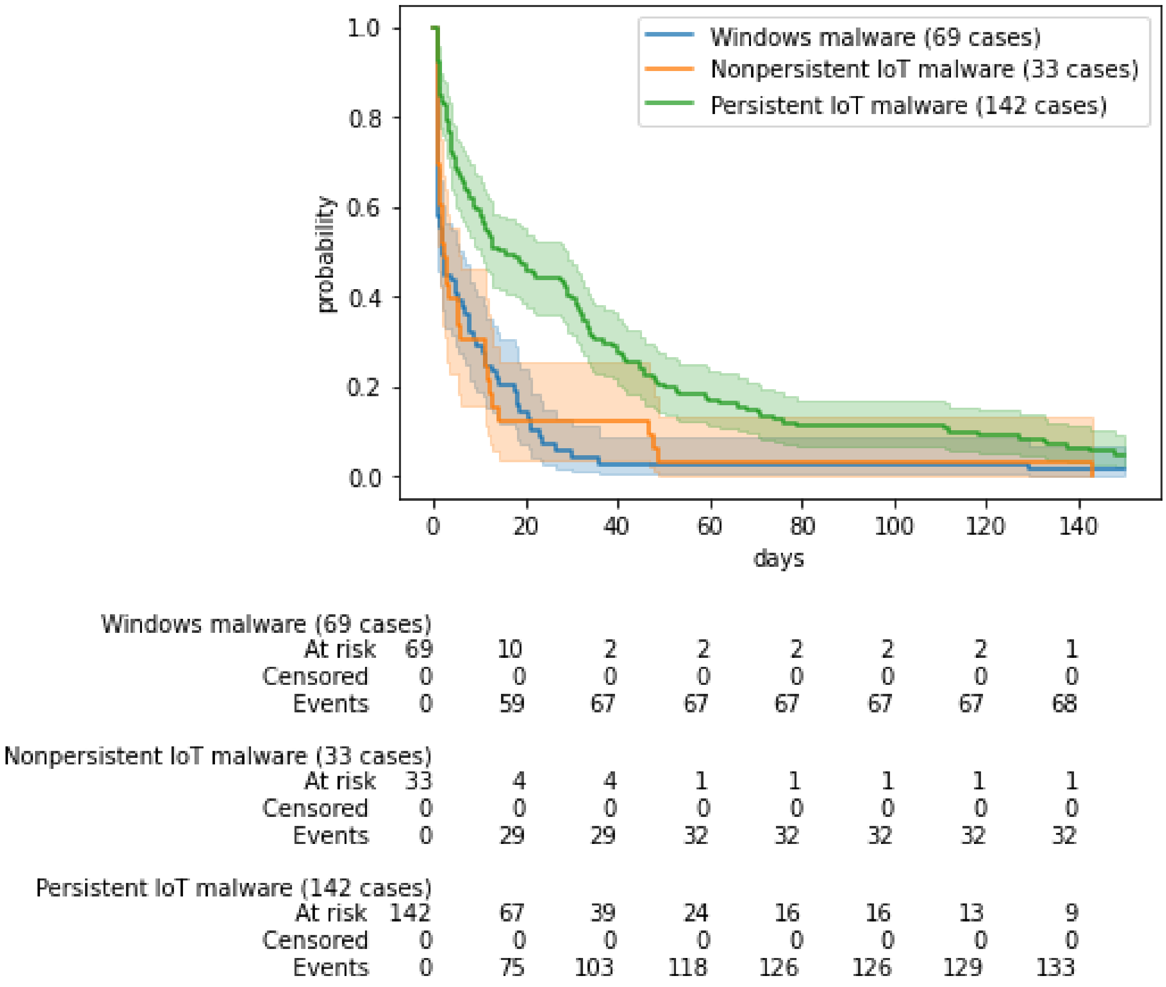


Figure 10: Kaplan-Meier curves of uncategorised malware only including notified customers infected by Mirai, QSnatch and Windows malware

When we categorise the different malware, we come to the curves presented in figure 11. In the groups of non-persistent and persistent malware, there are only respectively Mirai and QSnatch included and note again that VPNFilter is no longer included. The other malware are Windows malware. Therefore, the reader should keep in mind that we are comparing the survival probabilities of Mirai, QSnatch, and several Windows malware. Now, it becomes evident that, besides the higher life expectancy of persistent malware of the whole time period, persistent malware also seems to have the highest probability of survival.

The logrank tests on these malware also provide a statistical difference between the persistent IoT malware and the two other groups. The p-values provided by the tests were $6.726675523864104e-09$ between persistent IoT malware and windows malware, $5.817626989306576e-05$ between persistent and non-persistent IoT malware, and 0.8141171295523808. Therefore, the tests show a statistical difference between the life expectancies of the persistent malware compared to the other two types of malware. Remarkably, the p-value of the logrank test between the non-persistent IoT malware and Windows malware shows that we cannot reject the hypothesis that the life expectancies are the same. The figure also shows the cases at risk, censored and events at given moments. Based on the findings of these analyses, this research has to conclude that persistent IoT malware stays longer on the devices of customers compared to the other groups of malware, non-persistent IoT malware and windows malware.

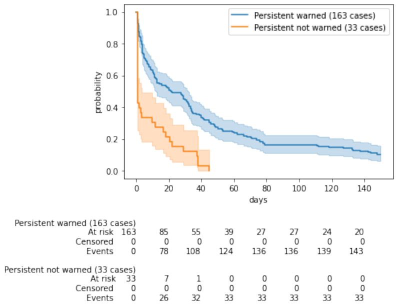


Windows malware (69 cases)								
At risk	69	10	2	2	2	2	2	1
Censored	0	0	0	0	0	0	0	0
Events	0	59	67	67	67	67	67	68
Nonpersistent IoT malware (33 cases)								
At risk	33	4	4	1	1	1	1	1
Censored	0	0	0	0	0	0	0	0
Events	0	29	29	32	32	32	32	32
Persistent IoT malware (142 cases)								
At risk	142	67	39	24	16	16	13	9
Censored	0	0	0	0	0	0	0	0
Events	0	75	103	118	126	126	129	133

Figure 11: Kaplan-Meier curves of categorised malware including notified cases

5.1.2 Notified and not-notified customers

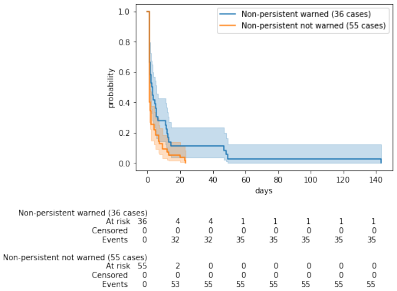
Now, that got us wondering how cases of infections perform which are not being notified. Figures 12, 13, and 14 each show the differences between the curves of the specific malware divided into customers that are notified and those who are not. Note that the cases which are not being notified perform much better in all types of malware. For persistent, non-persistent, and windows malware, the logrank tests also provide p-values of $1.2361008732332146e-08$, 0.011163472215983088, and $9.495331854198543e-21$. This means there is a statistical difference between the life expectancies of all of the different groups of malware. An essential element here is that the ISP would notify all of its customers with these types of malware at some point. So if a customer gets infected with malware, the customer would only not get notified if he remediates the issue before the ISP has the chance to notify the him. For this reason, it is not fair to compare the lines in the graphs as people that are longer infected tend to have a higher probability of being notified. After all, then there are more days on which the ISP could have enough notifications left to notify the customer. On the other hand, the graphs do indicate something intriguing. If we keep in mind that the groups of not-notified customers are in that group because the infection was remediated before the ISP had the opportunity to inform the customer, then we should take a look at the lines of the groups of customers that are not warned (the orange lines). Those lines would provide some indication of how long it takes before customers are being notified. It becomes evident that persistent malware can stay much longer on the devices than the other malware. At around 30 days, persistent malware still has a probability of a little bit less than 0.4, whereas non-persistent and windows malware reach 0 around that time. This would indicate that the persistent IoT and thus QSnatch malware is getting notified much later than the other types of malware. This would mean that the ISP has a lower priority in notifying QSnatch-infected customers compared to the other groups of malware.



Persistent warned (163 cases)								
At risk	163	85	55	39	27	27	24	20
Censored	0	0	0	0	0	0	0	0
Events	0	78	108	124	136	136	139	143

Persistent not warned (33 cases)								
At risk	33	7	1	0	0	0	0	0
Censored	0	0	0	0	0	0	0	0
Events	0	26	32	33	33	33	33	33

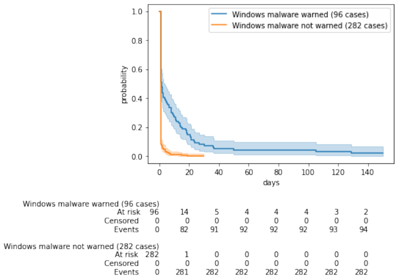
Figure 12: Kaplan-Meier curves of persistent IoT malware cases that are notified or not notified



Non-persistent warned (36 cases)								
At risk	36	4	4	1	1	1	1	1
Censored	0	0	0	0	0	0	0	0
Events	0	32	32	35	35	35	35	35

Non-persistent not warned (55 cases)								
At risk	55	2	0	0	0	0	0	0
Censored	0	0	0	0	0	0	0	0
Events	0	53	55	55	55	55	55	55

Figure 13: Kaplan-Meier curves of non-persistent IoT malware cases that are notified or not notified



Windows malware warned (96 cases)								
At risk	96	14	5	4	4	4	3	2
Censored	0	0	0	0	0	0	0	0
Events	0	82	91	92	92	92	93	94

Windows malware not warned (282 cases)								
At risk	282	1	0	0	0	0	0	0
Censored	0	0	0	0	0	0	0	0
Events	0	281	282	282	282	282	282	282

Figure 14: Kaplan-Meier curves of Windows malware cases that are notified or not notified

5.1.3 Conclusions on the malware clean-up results

In the previous subsections, we plotted the K-M curves of QSnatch, Mirai, and several Windows malware. It became clear that QSnatch has a relatively high probability of staying on the devices and dominates the curves at almost all moments in time. We then grouped the different malware into persistent IoT malware, non-persistent IoT malware, and Windows malware and analysed those groups with logrank tests. The purpose of this subsection was to answer SQ1: *'How does the QSnatch clean-up effectiveness compare to other types of malware such as non-persistent malware and non-IoT malware?'*

We found out that the persistent malware stays longer on the devices and statistically differs from the other two groups. The groups persistent IoT malware, non-persistent IoT malware and Windows malware consistent of QSnatch; Mirai; and Ramnit, Kovter, Citadel, Qrypterrat, Conficker, Gamarue, Necurs, Sality, Gozi, and Caphaw, respectively. So, compared to the clean-up results of the other malware, QSnatch stays longer on the customers' devices and has a higher probability of being on a customer's device at a certain moment in time. This confirms the hypothesis that QSnatch stays longer on the customers' devices.

A part of these results could be explained based on the prioritising of the ISP and sending notifications rather to other malware instead of QSnatch. When we plotted the curves of the different groups of malware and compared the notified and not-notified customers, we found that it was possible that QSnatch-infected customers can be not-notified for longer periods before the ISP gets the chance to notify them.

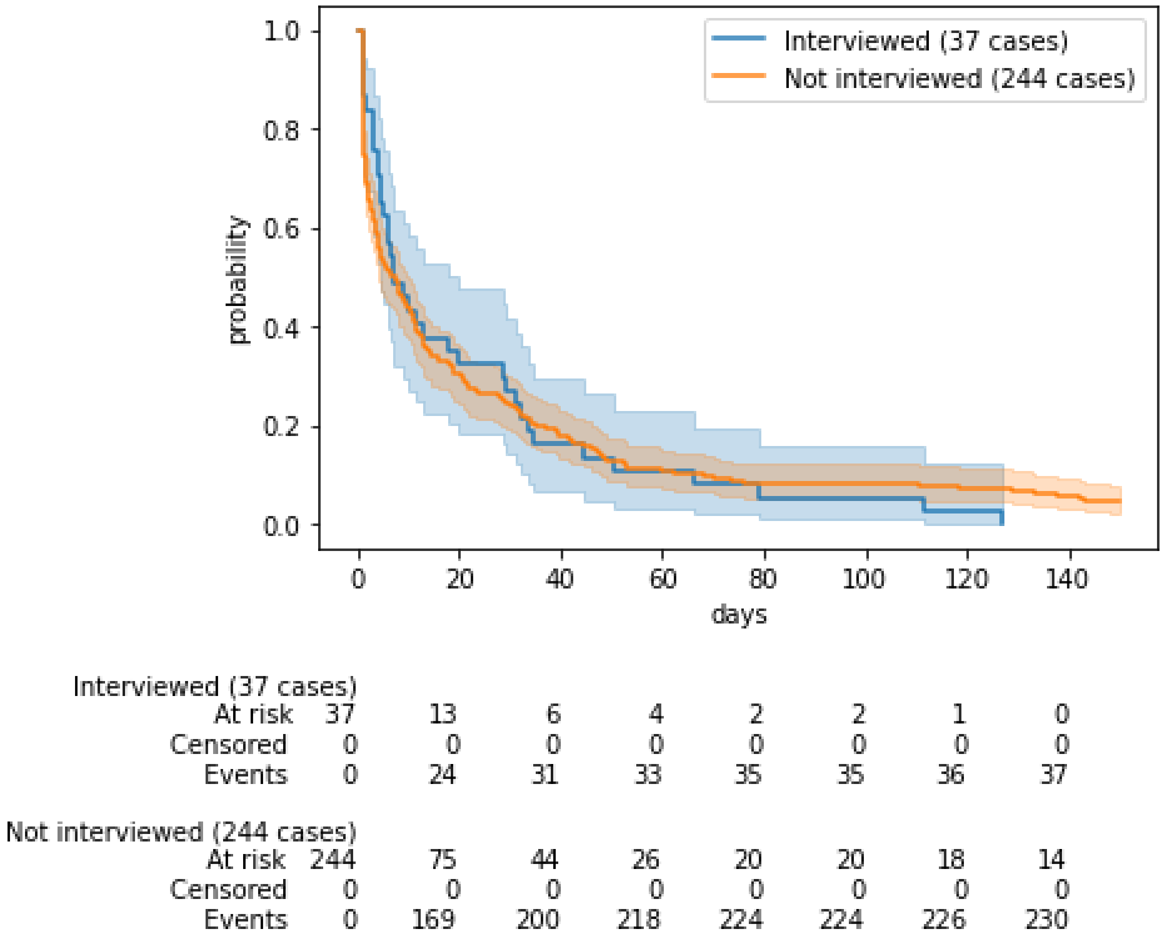
5.2 User experiences after receiving the infection notification

This chapter discusses the user experiences based on the interviews held with customers of the ISP. We will first analyse the representativeness of the selected group of customers to the wider population of the ISP's that are also infected with QSnatch but did not, for whatever reasons, participated in this part of the research. Thereafter, we will discuss the QNAP type and place of acquisition to see if there are perhaps specific model types of QNAP and places of acquisition that have relatively more trouble with QSnatch than others. We then look into how much time it took the respondents to perform the steps. Thereafter, the section discusses the results of respondents' abuse data and the questions based on the COM-B model. Finally, it will discuss the differences between the customers that use their devices for private or business purposes. This chapter uses much coding defined in section 4.3.4. If the reader feels, at any moment, disconnected from the codes in the following sections, please find the explanations to the codes and coding process in section 4.3.4.

5.2.1 The representativeness of the interviewed population compared to the wider population of the ISP's customers infected with QSnatch

To verify the group of interviewed customers do not perform better or worse in cleaning up QSnatch, this research analysed the representativeness of this group to the wider population of QSnatch-infected customers. Some respondents might back down from contributing in the form of an interview because they are afraid that their mistakes become publicly available or they could be ashamed of not being able to perform the steps. At the same time, customers that perform the steps relatively easily might be proud of their accomplishment and might be more likely to share their experiences. To cover any biases, we compared the K-M curves of both the interviewed and not-interviewed group as presented in figure 15.

From that figure, it becomes clear that the malware of the group of interviewed customers seems to have a very comparable life expectancy over the 150 days. Between 20 and 30 days the interviewed customers seem to perform a little bit less but seem to outperform a little bit after that. The interviewed group also reaches a probability of 0 at around 130 days, while the wider population still has a chance to be alive after 150 days. In other words, it is challenging to find which group performs better in cleaning up QSnatch. The LogRank test comparing the results provides a p-value of 0.9678366446769903. With an $alpha$ of 0.05 we do not reject the hypothesis that the expectancies are different. For this reason, this research argues that the group of interviewed customers are representative for the wider population of the ISP's customers that are infected with QSnatch but did not participate in this part of the research.



Interviewed (37 cases)								
At risk	37	13	6	4	2	2	1	0
Censored	0	0	0	0	0	0	0	0
Events	0	24	31	33	35	35	36	37
Not interviewed (244 cases)								
At risk	244	75	44	26	20	20	18	14
Censored	0	0	0	0	0	0	0	0
Events	0	169	200	218	224	224	226	230

Figure 15: Kaplan-Meier curves of the interviewed and not interviewed group

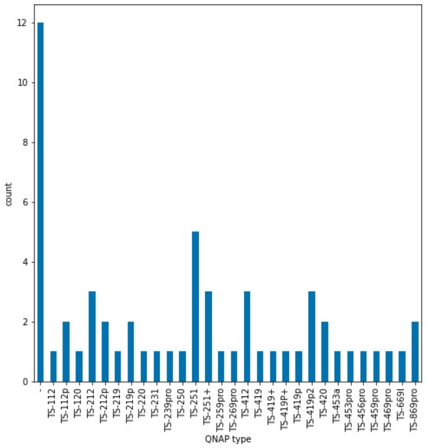


Figure 16: The QNAP types of the respondents

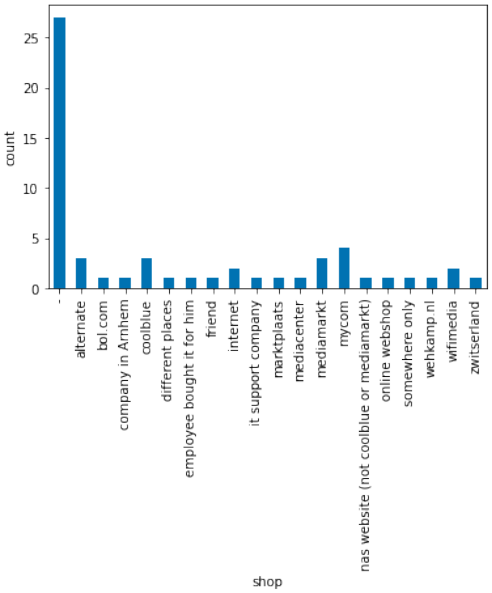


Figure 17: The place of acquisition

5.2.2 The QNAP type, place of acquisition, and the time it took the respondents to perform the steps

The type of QNAP device is something many respondents could not remember during the interview. For 12 of the respondents, this was the case. Indicated with a hyphen (first column), figure 16 shows the number of respondents that did not remember their QNAP's type. The figure also shows the types of QNAP devices of respondents that did remember which type they had. The respondents had to make some effort sometimes to look up the type number. It becomes clear that a relatively large number of respondents owned a TS-251. Out of all respondents, five of them owned such a device. We do not know what the sales numbers are for different QNAP devices. Therefore, we cannot conclude that the TS-251 is more vulnerable to QSnatch than other types of QNAP devices. QNAP provides the product support status of all of their NAS devices on its website [e]. On the website, QNAP shows the Hardware Repair or Replacement, OS and Application Updates and Maintenance, Technical Support and Security Updates.

Secondly, figure 17 shows where the respondents bought their devices. From this figure, it becomes clear that the respondents forget the place of acquisition. Some respondents would argue: *'Well, I bought the devices roughly ten years ago. I cannot remember where I bought it, to be honest.'* That for most respondents, the time between the day they bought their device and the interview was too long becomes clear from the fact that 27 respondents forgot where they bought it. On the other hand, 4 respondents could indicate buying their device from MyCom. Frankly, this company does not even exist anymore. It is interesting to see that a lot of the respondents bought their devices from an online shop or *'on the Internet'*. These are also the NAS retailers we talked about in chapter 1.1. Relatively many of the people who still knew where they had bought their device came from online stores such as Alternate, Cooblue, Bol.com, Mediamarkt, and Mycom.

On average, it took the respondents 69 minutes and 29 minutes to perform the clean-up of their devices. An overview of what the respondents indicated on how much time it took them is presented in figure 18. Out of the respondents, nine could not give a rough indication. Those cases are displayed as a -1 in the figure. Table 3 shows the results per individual respondent.

Twenty-seven respondents indicated a time less than the forty minutes it took the employees and the researcher to perform the steps. This research would assume they did not include the time to wait on running different processes such as the actual downloading, installation, and scanning by the device. That in itself is not surprising, because in principle, it is possible to do something else during that time as those processes on themselves do not require actions by the user. Intriguing are the results of the respondents that indicated that it took them 2 and 3 hours or even a day of work. Respondents 3, 10, 24, 30, 34, 35, and 48 needed 2 hours to complete the steps. It took 3 hours for respondents 25 and 52. Finally, two respondents, respondent 21 and 31, needed a day to perform the steps.

[e] https://www.qnap.com/en/product/status

duration	*respondent*	*total*
no indication	2, 4, 5, 14, 18, 26, 49, 55, 56	9
10	7	1
15	6, 20, 23, 27, 41, 44, 50, 57, 58	9
20	22, 32, 37, 42	4
30	1, 11, 12, 13, 17, 19, 28, 33, 38, 43, 45, 46, 47, 53	14
45	16, 39	2
60	9, 15, 29, 36, 40, 51, 54	7
90	8	1
120	3, 10, 24, 30, 34, 35, 48	7
180	25, 52	2
480	21, 31	2
total		*58*

Table 3: The time it took the respondents to perform the steps in minutes

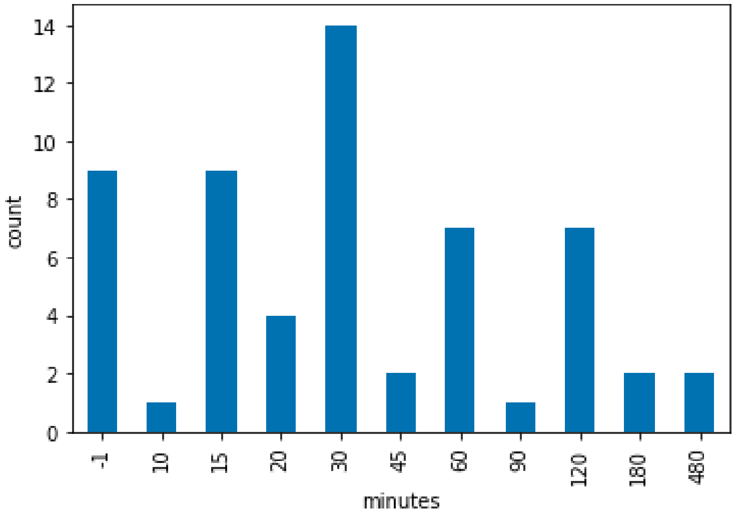


Figure 18: The time it took respondents to complete the steps

5.2.3 What happened?

Out of all respondents, 25 responded that they followed the steps of the ISP and specifically stated that they updated the firmware and installed the malware remover. Besides, in 7 cases, the respondents noted that they updated the firmware but not specifically stated that they installed the malware remover. At the same time, there were 8 respondents that indicated that they installed the malware remover tool but did not indicate that they updated the firmware. However, the respondents in the latter two groups did indicate that they performed the steps in the notification mail. Therefore, this research would assume that they performed the steps in the notification mail but forgot to mention those steps. Three respondents indicated that they followed the steps of QNAP, while 6 respondents turned off their devices. One respondent could not remember what he did anymore. This was the case for respondent 47. On October 29, 2020, he was last seen on the abuse data, while he was interviewed on April 29, 2021, more than a half year later.

Moreover, eight respondents had someone else perform the steps. In section 5.2.9, we will take a closer look at who specifically helped them to perform the steps. Noteworthy is that respondent 25 is the system administrator of respondent 4. One respondent was not able to reproduce the things he did to clean his device. If we add these numbers together, we get 58. Out of those 58 respondents, four addressed that they tried the steps but failed to remove the malware. Another 8 respondents also performed an extra step. Two respondents also discussed specific setbacks they encountered in their pursuit to remove the QSnatch malware. We will discuss the findings of those in the following sections as they act on different parts of the COM-B model. Table 4 shows the respondent's reference numbers according to their coded responses on what they have done.

Respondents were also asked whether the device was used for private or business purposes. During the research, it became evident that BM customers are not notified when a device in their network becomes infected with the

done	*respondent*	*total*
updated and installed	3, 5, 10, 11, 17, 18, 20, 21, 22, 23, 24, 25, 32, 33, 35, 39, 40, 42, 44, 45, 48, 53, 54, 56, 57	25
updated	2, 6, 19, 41, 46, 50, 52	7
installed	1, 12, 29, 30, 36, 38, 55, 58	8
steps qnap	28, 31, 43	3
turned-off	13, 14, 15, 16, 26, 49	6
someone else	4, 7, 8, 9, 27, 34, 37, 51	8
do not remember	47	1
total		*58*
tried but failed	13, 15, 34, 16	4
extra step	12, 25, 28, 30, 32, 45, 48, 53	8
setback	5, 16	2

Table 4: Results on what the respondents did

QSnatch malware. For some other malware, specific companies are notified, which is done manually by an employee. However, this does not happen for QSnatch-infected customers. At this time, the Abuse Team does not have the correct information to reach the customers. They have to make do with information that is freely accessible on the Internet. For example, an employee of the Abuse Team can find information from email addresses that start with 'info@...' and general email addresses. Therefore, no BM customers are involved in the study and it remains with customers who have a CM subscription but who have indicated that they have used the device for business purposes. From the total number of respondents, five indicated that they used the device for business purposes. Note that respondent 25 is actually the IT professional of respondent 4. So, there are actually only 4 customers that use their devices for business purposes. 6 respondents indicated that they used the device for both business and private purposes. It was not discussed with two respondents, and the rest indicated that they use it for private purposes.

use	*respondent*	*total*
private	2, 3, 5, 7, 9, 10, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 55, 56, 58	45
business	4, 8, 12, 25, 40	5
both	1, 6, 16, 27, 54, 57	6
not discussed	28, 53	2
total		*58*

Table 5: Results on the purposes of the devices

When we put the groups 'both' and 'business' together; grouped the respondents that either belonged to the groups 'updated and installed', 'updated', and 'installed'; or put the tables 4 and 5 next to each other, new insights arise. Now we have 10 cases in the business/both group. We will use this grouping for the remainder of the study. The results on the grouped respondents are shown in table 6. For each group, the shares are included by dividing the number by the total in the group. This mainly shows that devices that are also used for business purposes are relatively often cleaned by a third party. Respondents 4, 8, and 27 from the business/both group were helped by someone else. In section 5.2.9 we will discuss who helped the respondents in more detail, but we will describe them for these 3 respondents now. Respondents 4 and 8 were helped by an IT professional and 27 by an acquaintance. It, therefore, seems that respondents who also use the device for business purposes are more likely to seek help.

	updated and/or installed	*steps qnap*	*turned-off*	*someone else*	*does not remember*	*total*
business/both	7	0	1	3	0	*11*
	0.64	0.00	0.09	0.27	0.00	
private	31	3	5	5	1	*45*
	0.69	0.07	0.11	0.11	0.02	
not discussed						*2*
total						*58*

Table 6: Results on what the respondents did based on the purpose (business/both and private)

The group of respondents consists of a part of the total group of customers who are or have been infected with QSnatch. Results of these respondents are also known regarding when the customer was interviewed, when he got the notification, and at which dates malicious traffic was detected. That data can be found in appendix L in table 51 the number of infections, infection duration, time of notification, and whether the malware was successfully removed. The periods of infections can then be determined based on that data and the rule of thumb that a device is only seen as clean again after a month without having generated malicious traffic. Based on the first notification and last date of traffic generated, we can then explain the results of the clean-up effectiveness of QSnatch malware at a more precise level. This data is also in appendix L but in table 51.

Table 7 provides a summary of the information gathered from this data. Here we see the date of the first notification, the first time malicious traffic was discovered coming from the customer's network and the last time it was seen. The following columns then show the duration of the infection, the period between the first traffic and the notification, and the period between the first notification and the last malicious traffic.

Moreover, the last line shows the averages of the calculated values, and these show that the average period between the first traffic and the notification is only slightly longer than the average period between the notification and the last traffic. This may indicate that a relatively large amount of time is lost by not sending the notification to the customer for quite some time. One could even argue that the difference between the clean-up results of QSnatch compared to the other groups of malware is caused by the lower prioritisation of the ISP. We saw that, on average, the time between the infection is first discovered and the first warning is roughly the same as the time between the first warning and the last time we infected customers pop up in the data. Thus, it takes, on average, roughly half of the total duration of the infection to notify the customer in the first place. Therefore, we could assume that we could divide the duration of the QSnatch-infected customers by 2 to roughly have the duration of the QSnatch infections without the time it takes before the ISP sends a notification to the customer.

If we then assume that the other groups of malware are notified instantly (which they are not) and that we could generalise the findings of the average times between the first seen, first notification, and last seen moments to the wider population of the ISP's QSnatch-infected customers, we could again plot the K-M curves of the grouped malware. Figure 19 shows these K-M curves. It becomes evident that even with the assumption that the other malware is notified instantly, persistent malware, i.e. QSnatch is still more stubborn than the other malware. The logrank tests between persistent malware and windows malware, between persistent malware and non-persistent malware, and between windows and non-persistent malware give p-values of 0.005813754806268576, $0.03932158013057484$4, and 0.8141171295523808, respectively. With an α of 0.05, these findings mean that QSnatch still stays longer on the devices even when we take into account the time the ISP needs to notify the infected customers. On the other hand, no one would argue it would not make a big difference if the ISP warned the QSnatch infected customers earlier.

Customers for whom it took a remarkably much time to remove the malware after the notification were the respondents 1, 8, 19, 20, 24, 26, 27, and 50. Other customers have been particularly affected by a system that has given them little priority to send a notification to. Those are respondents 2, 16, 23, 48, and 51.

Table 7: Dates of invitation, interviewing, notifying, and detection traffic of the respondents

r.	*1st warning*	*1st seen*	*last seen*	*duration*	*1st seen and warning*	*1st warning and last seen*
1	23/10/2020	20/10/2020	06/03/2021	137	3	134
2	16/12/2020	07/08/2020	17/02/2021	194	131	63
3	15/02/2021	12/02/2021	15/02/2021	3	3	0
4	15/02/2021	14/02/2021		0	1	0
5	27/01/2021	26/12/2020	26/01/2021	31	32	0
6	14/01/2021	13/01/2021		0	1	0
7	11/01/2021	08/01/2021	10/01/2021	2	3	0
8	11/01/2021	09/01/2021	14/07/2021	186	2	184
9	06/01/2021	03/01/2021	05/01/2021	2	3	0
10	no warning	21/12/2020		0		
11	11/02/2021	06/02/2021	10/02/2021	4	5	0
12	data not available					
13	data not available					
14	01/12/2020	30/11/2020		0	1	0
15	30/11/2020	27/11/2020	29/11/2020	2	3	0

16	16/11/2020	18/07/2020	29/11/2020	134	121	13
17	26/11/2020	24/11/2020	26/11/2020	2	2	0
18	26/11/2020	14/09/2020	02/12/2020	79	73	6
19	31/07/2020	12/07/2020	27/11/2020	138	19	119
20	06/08/2020	17/07/2020	12/02/2021	210	20	190
21	12/10/2020	22/07/2020	29/11/2020	130	82	48
22	21/10/2020	16/10/2020	13/11/2020	28	5	23
23	12/11/2020	20/07/2020	12/11/2020	115	115	0
24	05/08/2020	13/07/2020	10/11/2020	120	23	97
25	same as 4					
26	31/07/2020	18/07/2020	06/11/2020	111	13	98
27	30/07/2020	13/07/2020	10/11/2020	120	17	103
28	no warning	03/11/2020		0		
29	13/10/2020	20/07/2020	21/10/2020	93	85	8
30	21/10/2020	17/07/2020	20/10/2020	0	96	0
31	20/10/2020	19/10/2020	20/10/2020	1	1	0
32	12/10/2020	12/07/2020	18/10/2020	98	92	6
33	19/10/2020	12/10/2020	21/10/2020	9	7	2
34	31/07/2020	19/07/2020	15/10/2020	88	12	76
35	14/10/2020	17/07/2020	14/10/2020	89	89	0
36	13/10/2020	17/07/2020	13/10/2020	88	88	0
37	12/10/2020	22/09/2020	12/10/2020	20	20	0
38	12/10/2020	21/09/2020	12/10/2020	21	21	0
39	31/07/2020	16/07/2020	23/08/2020	38	15	23
40	11/08/2020	12/07/2020	30/08/2020	49	30	19
41	21/08/2020	20/07/2020	21/08/2020	32	32	0
42	30/07/2020	13/07/2020	03/08/2020	21	17	4
43	data not available					
44	data not available					
45	15/10/2020	17/07/2020	15/10/2020	90	90	0
46	11/12/2020	09/12/2020	11/12/2020	2	2	0
47	12/10/2020	18/07/2020	29/10/2020	103	86	17
48	22/10/2020	13/07/2020	18/11/2020	128	101	27
49	01/02/2021	31/01/2021		0	1	0
50	10/10/2020	19/07/2020	28/01/2021	193	83	110
51	16/12/2020	07/08/2020	17/02/2021	194	131	63
52	16/10/2020	17/07/2020	26/12/2020	162	91	71
53	11/11/2020	10/11/2020	18/11/2020	8	1	7
54	24/12/2020	23/11/2020	05/01/2021	43	31	12
55	16/11/2020	15/11/2020		0	1	0
56	29/04/2021	17/04/2021	29/04/2021	12	12	0
57	10/05/2021	10/04/2021	18/05/2021	38	30	8
58	29/04/2021	07/04/2021	17/04/2021	10	22	0
	Average			*66.3*	*38.5*	*30.1*

If we group the customers that use their devices for business and both again, we get table 8. It becomes clear that the average infection duration of the business/both group is longer compared to the total group and that this is mainly because of the time span between the first warning and the last seen of those cases. This would indicate that the business/both group need more time to clean their devices.

5.2.4 Physical Capability

During the interviews, none of the respondents indicated that they had any physical limitations that made the steps challenging. In four cases, the physical limitations are not discussed specifically. These findings make it difficult to make meaningful statements about the influence of physical capability on the behaviour of the respondents. On the other hand, it is also not surprising that the respondents did not have any limitations that stood in the way of carrying out the steps. After all, their physical condition must meet the same requirements in order to purchase, configure and use a device. In such a case, probably someone will already be present to assist. We can therefore assume that this

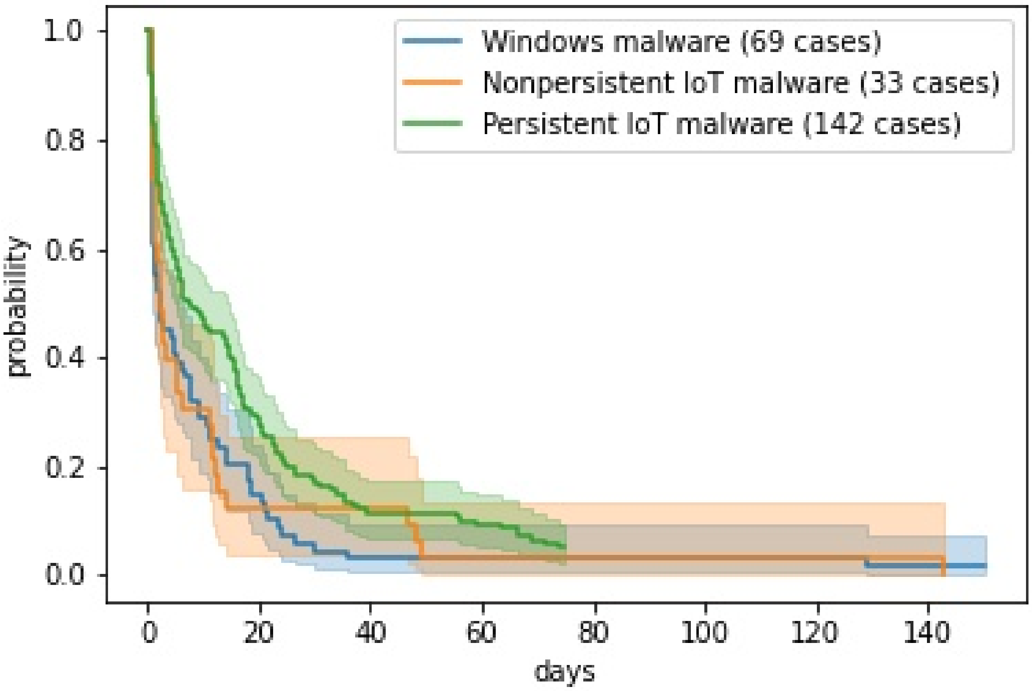


Figure 19: Kaplan-Meier curves of categorised malware including notified cases with the duration of persistent malware divided by 2

r.	*duration*	*1st seen and warning*	*1st warning and last seen*
1	137	3	134
4	0	1	0
6	0	1	0
8	186	2	184
12	data not available		
16	134	121	113
27	120	17	103
40	49	30	19
54	43	31	12
57	38	30	8
	78.6	*26.2*	*63.7*

Table 8: Results of the business/both group

factor will not have played a major role in other cases beyond the research's respondents. Table 9 shows the results of the coded answers of the respondents.

physical limitations	*respondent*	*total*
no	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58	57
yes	-	0
not discussed	26	1
total		*58*

Table 9: Results on the physical limitations of the respondents

Next, we group the respondents together according to the business/both and private groups. Table 10 presents the findings. However, we do not come to new insights based on this table as the respondents all answered that they did not have any physical limitations apart from one respondent that was not asked about it.

	no	*yes*	*total*
business/both	11	0	*11*
	1.00	0.00	
private	45	0	*45*
	1.00	0.00	
not discussed			*2*
total			*58*

Table 10: Results on the physical limitations of the respondents (business/both and private)

5.2.5 Psychological Capability

The respondents showed much psychological capability to perform the steps. No fewer than 48 of the respondents said that their understood the steps and were able to explain what was expected from them. One of the respondents indicated that he did not take the time to understand the steps and just turned off the device. Two respondents indicated that they were not sure if they understood the steps but thought they did as the problem appeared to be resolved. Respondent 8 explained: *'Well that [the steps] is going too far for me so I won't. The risk that if things don't go well, even though they are not that difficult, is too big.'* For that reason the respondent consulted an IT professional to perform the steps. Table 11 shows the results on the coded responses of the respondents. Five respondents indicated they did not remember whether they understood the steps. Those are respondents 1, 27, 38, 49, and 51. Of the respondents 27 and 38 it is quite understandable that they no longer know whether they understood the steps when they performed them. The dates that they were last seen in the data and their interview were respectively November 6, 2020 and April 26, 2021; October 12, 2020 and April 28, 2021. For both cases, the time between the last time they appear in the data and the interview is approximately six months. The cases 49 and 51 forgot whether they understood the step in the periods of January 31, 2021 and May 4, 2021 and February 17, 2021, and May 4, 2021, respectively approximately 3 and 4 months. Intriguingly, respondent 1 was not able to indicate whether he understood the steps after the period from March 6, 2021, to April, 2021, roughly 1 month.

understand	*respondent*	*total*
yes	3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 28, 29, 30, 31, 32, 33, 35, 36, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 50, 52, 53, 54, 55, 56, 57, 58	48
to a certain extent	34	1
no	8	1
not remember	1, 27, 38, 49, 51	5
did not look	2, 14	2
not discussed	26	1
total		*58*

Table 11: Results on whether the respondents understood the steps

Table 12 shows the business/both and private grouped respondents and their responses to whether they understood the steps. In the business/both group, it seems as relatively more respondents forgot whether they understood the steps. This could be a result of that more of those respondents got help in cleaning their devices.

	yes	*to a certain extent*	*no*	*not remember*	*did not look*	*not discussed*	*total*
business/both	7	0	1	2	0	0	*11*
	0.63	0.00	0.09	0.18	0.00	0.00	
private	39	1	0	3	2	1	*45*
	0.85	0.02	0.00	0.07	0.04	0.02	
not discussed							*2*
total							*58*

Table 12: Results on whether the respondents understood the steps (business/both and private)

Also, 47 respondents indicated that they did not find the steps challenging. Those respondents explained that the steps provided enough information on how to tackle the issue. Their argumentation was quite similar and mainly boiled down to the fact that it was a step-by-step explanation and that the steps were clearly described in their eyes. Although most of the respondents found the steps adequate, some respondents found the steps challenging. For example, respondent 27 indicated that people have different levels of background knowledge. Therefore, he advised making the steps simpler. Respondent 15 argued: *'I found it frustrating because they did not work.'*. The reasoning of respondent 19 was *'because I'm actually not that often in the QNAP environment. So, I had to put in some effort to get back in there.'* Other respondents indicated that they found the steps challenging because they did not work. Respondent 35 said: *'Simply installing the virus scanner did not work. I think that a less experienced user would then stop and give up.'* Respondent 40 could well imagine that there are other people who do not know what to do with the steps. It was challenging to uncover what those respondents specifically found challenging of the steps because they did not go into that. Table 13 summarises the findings. One respondent, respondent 49 indicated that he did not remember he found the steps challenging. He was last seen in the data on January 31, 2021 and interviewed on May 4, 2021. So, the interview was roughly 3 months after he was last seen in data.

challenge	*respondent*	*total*
no	1, 2, 3, 4, 5, 6, 7, 10, 12, 13, 16, 17, 18, 20, 21, 22, 23, 24, 25, 28, 29, 30, 31, 32, 33, 34, 36, 37, 39, 41, 42, 43, 44, 45, 46, 48, 50, 52, 54, 57, 58	41
yes	8, 9, 11, 15, 19, 27, 35, 38, 40, 47, 51, 55, 56	13
did not see	14	1
not discussed	26, 53	2
not remember	49	1
total		*58*

Table 13: Results on whether the respondent found the steps challenging

If we then group the respondents according to the business/both and private purposes of the devices, we get table 14. When looking at those results, it becomes clear that the business/both group answered in roughly the same way the private group did. So, there is not much difference between the groups to be found on this question.

	no	*yes*	*did not see*	*not discussed*	*not remember*	*total*
business/both	8	3	0	0	0	*11*
	0.7	0.3	0.0	0.0	0.0	
private	32	10	1	1	1	*45*
	0.72	0.22	0.02	0.02	0.02	
not discussed						*2*
total						*58*

Table 14: Results on whether the respondent found the steps challenging (business/both and private)

Malware is a term known to most of the respondents. In 49 cases, the respondents indicated that they knew what malware is and were able to explain it correctly. Two respondents indicated that they did know what it was but were not able to explain the term correctly and then admitted that they did not know after all. For example, respondent 16 said he knew what malware was, but when he was asked to explain it, he confessed: *'Yes, that is quite difficult'*. Respondent 41 gave a similar response, saying: *'Some kind of spam, a bit of virus. I can't explain that, no.'* The remaining 8 respondents indicated they did not know what malware is or at least that they could not explain it. A summary of the findings is in table 15.

malware	*respondent*	*total*
yes	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 40, 42, 43, 44, 45, 47, 48, 49, 51, 52, 53, 54, 55, 57, 58,	49
no	16, 27, 32, 39, 41, 46, 50, 56	8
not discussed	26	1
total		*58*

Table 15: Results on the respondents’ knowledge about malware

Grouped into the business/both and private groups, the respondents per group seem to perform the same. Table 16 shows these findings. There does not seem to be a difference between the groups based on their knowledge about malware.

	yes	*no*	*not discussed*	*total*
business/both	9	2	0	11
	0.82	0.18	0.00	
private	38	6	1	*45*
	0.85	0.13	0.02	
not discussed				2
total				*58*

Table 16: Results on the respondents’ knowledge about malware (business/both and private)

On the other hand, almost none of the respondents knew what persistent malware would mean or the difference between persistent and non-persistent malware. Only 7 respondents were able to indicate the difference. Some respondents were not able to indicate the term correctly but had a suspicion that it would be more difficult to remove such malware. For example, respondent 20 said: *’I can guess. I have not heard of those terms per se but I take persistent as if you try to remove it from your device, it still makes it spread and sticky, and keeps working. And non-persistent would be that you can probably remove it more easily and then it will be gone.’* The results show that the respondents were willing to try to get the answer right. As with the question about malware, some are still trying to give the right answer. Based on these results summarised in table 17, it does not seem to be required knowledge in order to remove the malware. Most people were able to remove the virus, while only a small percentage knew what persistent malware is. Even more, the terms persistent and non-persistent malware is not mentioned in the steps. It is therefore not surprising that customers could perform the steps without knowing what the difference between these concepts is.

persistent	*respondent*	*total*
no	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 58, 59	50
yes	12, 13, 14, 21, 30, 45, 57	7
not discussed	26	1
total		*58*

Table 17: Results on the respondents’ knowledge about the different between persistent and non-persistent malware

When grouping the respondents again in the business/both and private groups, it becomes evident that there is not much difference in the responses between the groups as presented by table 18. In both cases, most of the respondents did not know what persistent malware was or could explain the difference with non-persistent malware, and both perform at about the same rate. Therefore, there is not much difference to be found between the groups based on the knowledge about persistent malware.

On the question of whether the respondents thought they were able to perform the steps, forty-five respondents responded they could. Most of them already indicated during this question that they had experience in IT. One of the respondents answered: *’Installing and uninstalling software are tasks that I have done many times before.’* This summarises well

	no	*yes*	*not discussed*	*total*
business/both	9	2	0	*11*
	0.82	0.18	0.0	
private	39	5	1	*45*
	0.87	0.11	0.02	
not discussed				*2*
total				*58*

Table 18: Results on the respondents' knowledge about the different between persistent and non-persistent malware (business/both and private)

what many respondents answered. Many respondents do not let them be scared away when they have to perform steps like these, especially if there are explained to them in a step-by-step manner. On the other hand, four respondents indicated they did not think they could perform the steps.

Six respondents could not remember whether they thought they could perform the steps or did not see the steps. Respondent 38 said: *'I can no longer reproduce this. So I do not know. I know I struggled with them, but I do not even know if I performed them correctly.'* This is a problem that more customers encounter. The ISP can send a notification that the device is infected with the QSnatch malware, but can then not confirm that the virus has been successfully removed. As a result, customers are sometimes left in despair because they do not know whether they are still at risk or whether they can safely use the device again. Nevertheless, this research can state that, in general, customers are confident enough to perform or let someone else perform the steps. Table 19 shows the results and corresponding respondents.

self confidence	*respondent*	*total*
yes	1, 3, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 39, 40, 41, 42, 43, 45, 46, 47, 48, 50, 52, 53, 56, 58	45
no	4, 8, 9, 51	4
other	2, 37, 38, 44, 54, 55	6
not discussed	26, 49, 57	3
total		*58*

Table 19: Results on the respondents belief they could perform the steps

Then, we grouped the respondents according to the business/both and private purpose groups and presented it in table 20. The business/both group show that they are less likely to believe they could perform the steps. That is a difference between the two groups that becomes clear. Logically, it follows that they are more likely to hire someone to help them, as we saw earlier in the research.

	yes	*no*	*other*	*not discussed*	*total*
business/both	7	2	1	1	*11*
	0.64	0.18	0.09	0.09	
private	36	2	5	2	*45*
	0.80	0.04	0.11	0.04	
not discussed					*2*
total					*58*

Table 20: Results on the respondents belief they could perform the steps (business/both and private)

Most of the respondents had experience in IT. Of all respondents, thirty had work experience. For example, one respondent said: *'I have over 30 years of working experience.'* Besides work experience, five respondents also indicated to have some form of education. Nineteen respondents expressly indicated not to have experience. However, a lot of those did find IT fun and approach it enthusiastically. For example, one respondent told *'No, I do not have IT experience, but I do have a server running for 15 years.'* Respondent 28 indicated he did not have work experience but sees IT as a hobby. A somewhat cryptic answer came from respondent 11. He said: *'I have all kinds of equipment here and have to give them all an address, so giving an address in your network. [...]. That is my networking experience.'* This respondent probably would have been a hobbyist, but he has been given his own label just to be sure. The results on the experience of the respondents is summarised in table 21.

It is becoming evident that the group of respondents are not representative for average customers of the ISP. Most Dutch people will have no experience with IT and in this research, most of them even had work experience or often an affinity with the subject. It could very well be, as we reasoned in section 1.1.3, that people with a NAS device in general often have experience with IT. This could mean that the steps that are now offered by the respondents of this research can be carried out easily, but that other users would have more difficulty with them. This can have severe consequences if NAS devices will be used in the future by users with less experience or if other IoT devices also become susceptible to persistent malware. Although the former does not seem likely since people with little experience will not soon start buying NAS devices, the NAS market is growing, as discussed in section 1.1.5. Therefore, it is likely that a wider group of customers will buy the devices. And perhaps, that group has less experience than the relatively selected group that is currently working with them.

experience	*respondent*	*total*
work	2, 5, 6, 12, 13, 14, 15, 16, 18, 20, 22, 25, 29, 30, 31, 33, 35, 36, 40, 41, 42, 43, 45, 46, 48, 50, 52, 53, 56, 57	30
no	3, 4, 7, 8, 9, 17, 19, 23, 24, 27, 32, 38, 39, 47, 49, 51, 54, 55, 58	19
work and education	1, 10, 21, 34, 44	5
hobby	28	1
education	37	1
networking	11	1
not discussed	26	1
total		*58*

Table 21: Results on the respondents IT-related experience

If we group the responses of the customers according to the business/both and private groups, there are not many differences to find. These results are presented in table 22. So, it seems that both groups have about the same IT-related experience.

	work	*no*	*work and education*	*hobby*	*education*	*networking*	*not discussed*	*total*
business/both	6	4	1	0	0	0	0	*11*
	0.55	0.36	0.09	0.00	0.00	0.00	0.00	
private	23	15	4	1	0	1	1	*45*
	0.52	0.33	0.09	0.02	0.00	0.02	0.02	
not discussed								*2*
total								*58*

Table 22: Results on the respondents IT-related experience (business/both and private)

5.2.6 Reflective Processes

In total, 50 respondents indicated that they found the steps helpful. Besides providing information on how to solve the issue, the notification mail also offered the insight that there was something wrong with the device in the first place. One respondent explained: *'Despite my experience, I did not know my device was infected and the steps helped to remove the malware.'* Some respondents also indicated that without the notification, they surely would not have noticed the infection at all. Respondent 20 also indicated that the steps helped him a lot. He said: *'The advantage of the steps is that you then have a good overview, and you do not have to figure out what to do yourself. You don't have to pioneer, but you can simply follow this step-by-step plan.'* The steps helped respondent 21 well on his way. This respondent explained: *'At least it gave me a starting point to see where to look for a solution to the problem.'*

A relatively limited number could not remember it anymore or did not found the steps helpful. A big note to make here is that the steps were found to be useful unless it turned out they did not remove the malware. In those cases, it was especially frustrating for the respondents that the steps were not sufficient. For example, respondent 28 explained that the steps were not useful because *'they did not work'*. And respondent 41 said: *'No, not really, because in this case, it was not the solution to the problem.'* Nevertheless, the majority of respondents see the usefulness of the steps. The codes and respondents who made the statements are listed in table 23.

useful	*respondent*	*total*
yes	1, 3, 4, 5, 7, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40, 42, 43, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 57, 58	49
no	28, 41	2
limited	45	1
not sure	2, 6, 8, 38	4
not discussed	15, 26	2
total		*58*

Table 23: Results on whether the respondents found the steps useful

When grouped according to the business/both and private groups, it becomes clear that about the same rate of the groups found the steps useful. This is shown in table 24. It is true that in the business/both group, there are more respondents who were not sure whether the steps were useful, but both groups seem to find the steps useful to about the same degree. It is, therefore, difficult to find a difference between the two groups in this respect.

	yes	*no*	*limited*	*not sure*	*not discussed*	*total*
business/both	9	0	0	2	0	*11*
	0.818	0.000	0.000	0.182	0.000	
private	39	1	1	2	2	*45*
	0.867	0.022	0.022	0.044	0.044	
not discussed						*2*
total						*58*

Table 24: Results on whether the respondents found the steps useful (business/both and private)

Except for 8 persons, the respondents knew there was some kind of risk as long as they did not get rid of the malware. Some of them said that their systems could be hacked, which is ironic as that has already happened to some extent. Others also talked about the possibility of data loss or theft. Even ransomware attacks were touched upon by the respondents. Some respondents indicated that the malware would stay on their devices and infect other devices. All in all, 30 respondents indicated that some form of harm could arise towards the user of the device. Three respondents were especially afraid that their network would be shut down, and one specifically indicated that the ISP would get in trouble. Finally, one respondent indicated that staying infected could also lead to infections with other malware.

Only 4 respondents could not indicate what could happen. It could be that they had an idea, but they did not know specifically what the malware was doing. For example, respondent 35 said: 'Yes, that is difficult. I didn't experience anything like that myself.' Respondent 41 was the only one that argued that nothing would happen. The steps did not work for him, so his answer to the question of what would happen was: *'Nothing will happen because the steps were not the solution.'* Despite these few cases, there is still enough reason to assume that the respondents were aware of what the consequences would be if the steps were not performed. Table 25 summarises the findings on what the respondent thought would happen if someone does not follow the steps.

not follow	*respondent*	*total*
harm to the user	1, 2, 3, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17, 19, 20, 22, 25, 28, 29, 32, 36, 39, 44, 46, 49, 50, 51, 54, 55, 58	30
malware stays	14, 24, 33, 38, 45, 53	6
harm to the user and network	23, 30, 31, 37, 43, 57	6
harm to the user and spread	10, 18, 47	3
shut-off	40, 48, 52	3
spread	15, 42	2
nothing	41	1
does not know	4, 21, 27, 35	4
not discussed	26, 34, 56	3
total		*58*

Table 25: Results on the respondents' thoughts on what would happen if someone does not follow the steps

After grouping the respondents and their answers according to the business/both and private groups as done in table 26, it becomes evident that both groups mainly indicate that something terrible is going to happen if they do not follow the steps. A striking difference between the groups is that the business/both group does not know exactly what will happen if someone does not perform the steps.

	harm to the user	*malware stays*	*harm to the user and network*	*harm to the user and spread*	*shut-off*	*spread*	*nothing*	*does not know*	*not discussed*	*total*
business/both	7 0.636	0 0.000	1 0.091	0 0.000	1 0.091	0 0.000	0 0.000	2 0.182	0 0.000	*11*
private	21 0.467	6 0.133	5 0.111	3 0.067	2 0.044	2 0.044	1 0.022	2 0.044	3 0.067	*45*
not discussed										*2*
total										*58*

Table 26: Results on the respondents' thoughts on what would happen if someone does not follow the steps (business/both and private)

The respondents generally felt responsible for carrying out the steps. Three respondents did not belong to that group and did not feel they were responsible. To illustrate, respondent 43 stated: *'because it happens so secretly and under the radar. With this malware, I also saw that it had been active for almost a month or two before it was found. If professionals have so much trouble dealing with it, what can you expect from an ordinary citizen?'* And there is something to be said for that. It is incredibly difficult to find out that you are infected with malware. And it is also true that even professionals have difficulty detecting malware.

However, a person wondered if someone is only responsible when he has been informed of the situation. Three respondents also indicated that they think they are responsible for executing the steps when they are aware of the infection. Respondent 42 argued: *'Well, I think, at that point if I've been warned about it. Look, I didn't know. I am going to assume that. And certainly not about the QNAP server [...] that it could become infected as a result, because I actually run everything via my laptop.'* And as such, respondent 54 gave a comparable answer: *'No, but I do think it is my responsibility when I am pointed out that action is being taken.'*

The other 51 respondents indicated that they were or felt responsible. One of them explained: *'Well, it is my device.'* Another argued: *'I am the owner of the device. Who else is going to do it?'* The rest of the answers are comparable to those just mentioned. The answers were mostly along the lines of whether it was their device, network or problem or respondents wondering who else would do it. So overall, respondents felt responsible for carrying out the steps. Table 27 contains the respondents' responses regarding their responsibility.

responsible	*respondent*	*total*
yes	1, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 55, 56, 57, 58	51
yes, if you know	10, 42, 54	3
no	2, 41, 43	3
not discussed	26	1
total		*58*

Table 27: Results on the respondents' thoughts on responsibility

When grouped according to the business/both and private groups as in table 28, the responses of the interviewed customers do not seem to give us intriguing insights. Both groups answered at roughly the same rates, and differences between the groups cannot be found based on the presented data. The overall group clearly states that they themselves are responsible for carrying out the steps.

Finally, the respondents also asked a variety of questions during and at the end of the interview. These questions came from the respondents themselves and are therefore not specifically based on one of the subcomponents of the COM-B model. Because the respondents were shown to their questions that they reasoned about the research and the clean-up,

	yes	*yes. if you know*	*no*	*not discussed*	*total*
business/both	10	1	0	0	*11*
	0.91	0.09	0.00	0.00	
private	40	2	3	1	*45*
	0.89	0.04	0.06	0.02	
not discussed					*2*
total					*58*

Table 28: Results on the respondents' thoughts on responsibility (business/both and private)

this research considers the questions as expressions of reflective processes. It shows that the respondents wanted to know more and were curious. The study concludes from this that the respondents were motivated.

Several respondents were interested in how the ISP knew they were infected and how they determined it. For example, respondent 1: *'How do you do it and how do you get that information?'* and respondent 18: *'How do you actually see that I am running malware? Is that immediately visible to [name ISP]?* And respondent 12 was worried about whether the steps were enough to secure his device. He asked: *'Are these very vulnerable devices or do you say to be really safe you should put a full server between them?'*

Respondents 16 and 23, on the other hand, were more interested in the malware itself. They also asked almost the same questions: *'Are there many problems with this virus? Within the QNAPs?'* and *'Is it a big problem, the virus?'* Similarly, questions were asked about how the respondent could be infected, how the respondent could reach the device remotely, the significance of malware, whether it was safe to turn on the device, and whether other types of malware are also detected.

A group of 7 respondents also had questions about the research itself. For example, respondent 12 wondered what the role of TU Delft was in the study and respondent 23 wanted to know whether the interviewer was from the university. Respondent asked the question: *'Because this is a study from [name ISP] or something? And what happens to it then?'* All of the above questions were answered by the interviewer. An overview of all labelled questions is shown in table 29.

question	*respondent*	*total*
determination	1, 18, 23, 40, 44, 58	6
extra measures	12	1
how infected	14	1
malware	16	1
operate from distance	19	1
other malware	25	1
persistent	34, 37, 40	3
qsnatch	16, 23, 30, 37, 48	5
remediated	47	1
research	12, 22, 23, 37, 45, 48, 58	7

Table 29: Results on the respondents' questions

5.2.7 Automatic Processes

Despite the fact that 23 respondents did not feel anything in particular, the feelings among the other respondents were distinctive. Five respondents had feelings of fear when they got the notification, were anxious or frightened or felt unsafe. Seven other respondents were primarily surprised or startled that they were infected with malware. Another group of 5 respondents were happy or relieved when they removed the malware from their device. One indicated even to feel euphoric when his device was clean again. Some respondents even felt angry, frustrated, annoyed, fed up, or tense. Some did not trust the notification, so they had their doubts and found it suspicious and untrustworthy. Table 30 shows the summary of these findings.

feelings	*respondent*	*total*
nothing	6, 13, 21, 23, 25, 29, 31, 34, 35, 39, 40, 41, 44, 45, 46, 48, 49, 50 52, 53, 54, 56, 58	23
angry	42	1
anxious	55	1
challenged	28	1
disappointed	33	1
distrust	14, 27	2
euphoric	18	1
excitement	3, 17, 24, 32, 38	5
fear	2, 8, 33, 36, 38	5
frightened	51	1
frustration	3, 10, 15, 28, 36, 43	6
glad	12	1
happy	5, 30, 47, 51	4
insecurity	19	1
pride	33	1
relief	2, 17, 20	3
shitty	37	1
startled	9, 37	2
surprised	1, 4, 7, 11, 12, 16, 30, 38, 57	9
suspicion	27	1
urgency	4	1
worried	22	1
not discussed	26	1

Table 30: Results on the respondents' feelings

When we group the feelings of the respondents into either if a feeling helped them perform the steps or not, i.e. they felt something or nothing and according to the business/both and private groups, we come to table 31. It becomes evident that there is not much difference between the results based on these groups. Both groups answered this question roughly the same.

	nothing	*something*	*not discussed*	*total*
business/both	4	7	0	*11*
	0.36	0.64	0.00	
private	18	26	1	*45*
	0.40	0.58	0.02	
not discussed				*2*
total				*58*

Table 31: Results on the respondents' feelings (business/both and private)

An impulse did not help respondents perform the steps in 20 cases. On the other hand, nineteen respondents indicated that the notification mail acted as an impulse for them to perform the steps. It is understandable that the respondents answered this way. According to Merriam-Webster [168], an is either *'a sudden spontaneous or incitement to some usually unpremeditated action'* or *'a propensity or natural tendency other than rational decision'*. In the context of automatic processes, the mail address is not considered a factor here but rather something to look out for in the physical opportunity section. Others said that they knew that they had to fix it quickly. For another, things like *'get rid of that thing'* went through his mind. Another respondent said: *'I wanted to help my client as good as I can'*. The latter seems to be more like a motivation than an impulse. As far as it has been successful so far, it is challenging to distil actual impulses from the respondents' reactions. Table 32 includes a summary of these findings.

impulse	*respondent*	*total*
no	1, 2, 7, 13, 15, 16, 17, 25, 29, 30, 33, 34, 37, 38, 39, 41, 43, 48, 51, 58	20
yes	4, 5, 6, 11, 14, 19, 22, 23, 28, 32, 44, 46, 53	13
email	3, 8, 9, 10, 12, 18, 20, 21, 31, 35, 40, 42, 45, 47, 49, 50, 52, 54, 55	19
not sure	24, 27, 57	3
not discussed	26, 36, 56	3
total		*58*

Table 32: Results on the respondents' thoughts on impulses

Table 33 shows the results when we group the responses of the customers according to the business/both and private groups. The business/both are less sure about whether an impulse helped them perform the steps. Beyond that, there are no fundamental differences between the two groups based on the findings on this question.

	no	*yes*	*email*	*not sure*	*not discussed*	*total*
business/both	3	2	4	2	0	*11*
	0.2727	0.1818	0.3637	0.1818	0.0000	
private	17	9	15	1	3	*45*
	0.38	0.20	0.33	0.02	0.07	
not discussed						*2*
total						*58*

Table 33: Results on the respondents' thoughts on impulses (business/both and private)

5.2.8 Physical Opportunity

Out of all respondents, only 3 did not receive the notification mail. Both of them followed the steps from QNAP and were also able to remove the malware from their devices. The rest received the notification and had the opportunity to follow the steps in the notification mail. Table 34 shows these findings.

received	*respondent*	*total*
yes	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 50, 51, 52, 53, 54, 55, 56, 57, 58	54
no	28, 43, 49	3
not discussed	26	1
total		*58*

Table 34: Results on whether the respondents received the notification email

When we group the responses of the customers according to the business/both and private groups, we get table 35. In the business/both group, there were no respondents who did not receive the notification mail. This gives a slight difference with the results of the group that use their devices for private purposes. However, the difference is minimal and mainly due to some respondents that assumed they did not receive the notification. Based on this information, it cannot be concluded that there is a difference between the two groups.

	yes	*no*	*not discussed*	*total*
business/both	11	0	0	*11*
	1.0	0.0	0.0	
private	42	2	1	*45*
	0.933	0.045	0.022	
not discussed				*2*
total				*58*

Table 35: Results on whether the respondents received the notification email (business/both and private)

Besides the provided steps, the respondents also used other tools to remove the malware. Nine of them indicated that they used a laptop or computer, or a keyboard. Those are tools that are necessary for executing the steps. This research does not take those into account as specific tools that someone need to have beyond what is already expected. The reason they are still mentioned is to show that the respondents indicated as such. However, it is not relevant information for this research. Others used Google or 'the Internet' to search the malware remover and the email address of the ISP and check them for authenticity. Moreover, one of the respondents used other virus scanners to make sure QSnatch was removed. Here, we see that the respondents did not trust the notification at first and had to do more research before taking action. One respondent found a Secure Shell (SSH) script to remove the malware. The respondent indicated that he had found the script somewhere online at a forum and had used it before. According to this respondent, the script removed the malware. Appendix M provides this script. The rest, 33 respondents, indicated they did not use tools besides the notification mail. Table 36 shows the different respondents and how their statements are labelled. Note that if one would add up the number of respondents for each code, then he would have a total of 61 respondents. Respondents 22, 28, and 57 did indicate to have used two tools. We coded the tools individually, but if we would only count one tool per respondent, we would come to 58 respondents in total.

tools	*respondent*	*total*
no	1, 2, 4, 6, 7, 8, 9, 11, 13, 14, 15, 17, 18, 20, 27, 28, 29, 30, 32, 35, 36, 37, 38, 39, 41, 48, 49, 50, 51, 53, 54, 56	33
Google	5, 16, 19, 25, 34, 40, 46	7
Internet	21, 31, 43, 47, 57	4
keyboard	58	1
laptop	3, 23, 33, 42, 44, 45, 55, 57	8
malwarebytes	10	1
other scanners	12	1
PC	22	1
script	24, 28	2
steps QNAP	22, 52	2
not discussed	26	1

Table 36: Results on whether the respondents used tools

	no	*some*	*total*
business/both	6	5	*11*
	0.55	0.45	
private	27	18	*45*
	0.60	0.40	
not discussed			*2*
total			*58*

Table 37: Results on whether the respondents used tools (business/both and private)

Fifty respondents had enough time to perform the steps. Six of them specified that in case of a malware infection, one should make time to fix the issue. Another argued: *'It is a matter of setting priorities.'* In general, respondents had enough time to complete the steps. It was, therefore, not a limiting factor. There was, however, one respondent who indicated that he did not have enough time. Respondent 34 said: *'Well, just like you, time is always a limited thing. So no, I would rather do something else.'* Apart from the fact that this is the only respondent who has indicated that he has no time, it is, of course, also debatable whether this respondent really did not have enough time. In table 38, there is a summary of the findings.

time	*respondent*	*total*
yes	1, 2, 3, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 57, 58	52
no	34	1
not sure	4, 16, 43	3
not discussed	10, 26	2
total		*58*

Table 38: Results on whether the respondents had enough time

Table 37 shows the results of the interviewed customers' responses when grouped according to the business/both and private groups. Also, here, we find that there is not much difference between the two groups and we, therefore, assume the respondents per group had at roughly the same rate enough time to perform the steps.

	yes	*no*	*not sure*	*not discussed*	*total*
business/both	10	0	1	0	*11*
	0.91	0.00	0.09	0.00	
private	40	1	2	2	*45*
	0.90	0.02	0.04	0.04	
not discussed					*2*
total					*58*

Table 39: Results on whether the respondents had enough time (business/both and private)

In most cases, the location was not an issue to access the NAS device. After all, the whole purpose of such a device is that one can access it from anywhere through an Internet connection. However, it became a problem in two cases. For one respondent, it was not possible to access the device as he was initially abroad when he got the first notification. Strangely, it was also not possible for him to access the device via the Internet. Eventually, he was able to perform the steps when he got home. Another respondent indicated that he was asked to remove the malware, but the device was at the office. Due to COVID, he was not able to pick up the device, and someone else had to do it. Table 40 shows the findings on whether the location was an issue for the respondents.

location	*respondent*	*total*
no	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 27, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 54, 55, 56, 57, 58	52
yes	19, 25, 29, 52	4
not discussed	26, 53	2
total		*58*

Table 40: Results on whether the location of the device or any of the used tools was an issue for performing the steps

As shown in table 41, it becomes clear that there is not much difference between the business/both and private groups according to the answers of the interviewed customers. This research would, therefore, assume that there is no difference between these groups when it comes to whether the location or any of the tools used was a problem to perform the steps.

	no	*yes*	*not discussed*	*total*
business/both	10	1	0	*11*
	0.91	0.09	0.0	
private	40	4	1	*45*
	0.89	0.09	0.02	
not discussed				*2*
total				*58*

Table 41: Results on whether the location of the device or any of the used tools was an issue for performing the steps (business/both and private)

5.2.9 Social Opportunity

To the question of whether the respondents had any help while executing the steps, most of them replied that they did it themselves. On 9 occasions, the respondents received help from others. For example, 3 people contacted a friend with experience for help. This was the case for respondents 9 and 27. Respondents 24, 34, and 37 got in touch with the QNAP help desk. They scheduled an appointment only, and the employees of QNAP were able to help in those cases. According to one of the respondents, the QNAP employees remotely logged in to their devices and removed the malware with some sort of program. Unfortunately, the customers could not provide more specific details about this help. In the other cases, customers got help from some IT professional. They got help from a system operator, IT support company, or an 'IT guy' that came to the rescue. This was the case for respectively respondent 4, 8, and 51. In those cases, the IT professional performed the steps for them. An overview of the results can be found in table 42.

help	*respondent*	*total*
no	1, 2, 3, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 28, 29, 30, 31, 32, 33, 35, 36, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 52, 53, 54, 55, 56, 57, 58	49
acquaintance	9, 27	2
IT	4, 8, 51	3
help desk qnap	24, 34, 37	3
not discussed	26	1
total		*58*

Table 42: Results on whether the respondents got help

When comparing the results on this question while grouping the responses according to the business/both and private groups, it becomes clear that the business/both group was more likely to ask for help from an IT professional compared to the results of the private group. Therefore, this research will assume that the interviewed customers that use their devices for business/both purposes are more likely to get help from an IT professional.

	no	*acquaintance*	*IT*	*help desk qnap*	*not discussed*	*total*
business/both	8	1	2	0	0	*11*
	0.73	0.09	0.18	0.00	0.00	
private	39	1	1	3	1	*45*
	0.87	0.02	0.02	0.07	0.02	
not discussed						*2*
total						*58*

Table 43: Results on whether the respondents got help (business/both and private)

In general, acquaintances, friends, and family members of the respondents did not have a strong opinion about executing the steps. It is essential here to note that a lot of respondents also indicated that they did not really talk about it with anyone. One respondent said: *'I do not know. They do not know I was infected with the malware.'* Probably, only a few of the respondents did actually talk about being infected by QSnatch with others. An example of someone you talked about the infection said: *'My wife, and she thought it was good that I had followed the steps'*. Another respondent had some concerns. He argued: *'I think some people will resist following the steps as they would not be able to perform them.'* Table 44 summarises the findings.

opinion	*respondent*	*total*
no	1, 2, 6, 7, 10, 12, 13, 14, 16, 17, 18, 19, 23, 25, 27, 28, 29, 30, 31, 32, 34, 35, 40, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 53, 55, 56, 57, 58	38
not talk	3, 5, 9, 15, 20, 21, 37, 38, 52	9
yes	4, 11, 22, 24, 33, 36, 54	7
not sure	8, 39, 43	3
not discussed	26	1
total		*58*

Table 44: Results on whether the respondents know people who have a strong opinion on performing the steps

Table 45 shows the results when we group the responses of the interviewed customers according to the business/both and private groups. The group of business/both customers seem to be more likely to have indicated that people they know do not have a strong opinion about performing the steps. This difference is mainly due to that respondents from the private group have indicated much more often that they did not talk to other people about the infection or the steps. In both cases, we can assume that the influence of the social environment of the respondents is not that prominent.

5.2.10 Expressions and suggestions

At the end of each interview, the respondents had the opportunity to add something they would like to mention, which was not discussed yet but, from their perspective, essential for the issue. The respondents had different answers to it. The answers to these and statements made during the interview are discussed in this section.

	no	*not talk*	*yes*	*not sure*	*not discussed*	*total*
business/both	8	0	2	1	0	*11*
	0.73	0	0.18	0.09	0	
private	28	9	5	2	1	*45*
	0.63	0.20	0.11	0.04	0.02	
not discussed						*2*
total						*58*

Table 45: Results on whether the respondents know people who have a strong opinion on performing the steps (business/both and private)

Eleven respondents expressed some distrust towards the communication of the ISP. They were not sure whether the information they got was fake and spam. As such, respondent 6 noted: *’I was surprised by your message, and I checked if your message was real.’* Respondent 12 put it like this: *’And what I always check when an email arrives: does it come from a known address, or an address that is trusted. So I had to be able to go back to something like: Gosh, is this recognisable, is this familiar? And I don’t remember exactly how that was in that email because usually when there is such an alarm soon: Gosh, you have to do this and press that button, I always expand it by definition. So before I take action, the mail must, in any case, be drafted in such a way that I also trust it and that I can also check it. To blindly follow the directions of an email from an email address that I have never received, you have to cross a few thresholds first.’* Those respondents are not the only ones because respondent 14 indicated: *’At first I thought, is this a real email?’* and respondent 47 said: ’You send an email out of the blue and that can come across as spam. [...]. That was also my first reaction.’ Respondent 17 also checked the email address first before taking any steps. Respondents 27 and 33 also indicated that they were reluctant to assume that the email really came from the ISP. They noted that this reluctance stemmed from media reports to be wary of emails claiming to have found malware on your systems.

Even more, respondent 15 noted about his contribution to the interview: *’I trust it, but also a little bit not. So I’m not going to release passwords and stuff like that, but I do want to participate. You understand my reluctance in that.’* This specific case also shows the motivation of respondent 15 to participate in the study despite the possible dangers of which he is aware. But the way of communication can also influence the clean-up results, as can be seen from the wording of respondent 50: *’I also omitted the email twice at first because I was like: Is this official [name ISP] or not? Nowadays, you get so many emails that you have doubts, but after I had received it repeatedly [...] then I was like: this is serious.’* It thus appears from the interviews with the respondents that a significant proportion is wary. And that is a fact that sounds like music to the ears of a researcher with a heart for cybersecurity, but in this case, it can hinder the clean-up. In addition, there is also the possibility that customers will get the idea that it is normal for the ISP to send emails with links to websites and will probably be less wary if real rogue emails are sent their way on behalf of the ISP. All in all a problematic situation.

Other respondents are particularly happy about receiving the notification. Respondent 1 said: *’Well, I think it is really good that you are doing this, and I think you are the only one [the only ISP]. I have never gotten a signal like this before. So, I really like what you’re doing.’* even went a step further and indicated: *’Well, I think it’s a good service from [name ISP] to handle it like this. I don’t think we would have found out otherwise.* This statement also makes sense since, for many, the notification email was the reason to take action. Sending the notification can, in many cases, also be seen as the essential link to get the customer started. Respondent 20 even though it was such an important aspect that he said: *’I thought it was fantastic that this was done proactively.’* To emphasise later: *’I think I reported that I think it is great that they are proactively approached when something like this is noticed on the server-side at [name of ISP]. I would like to emphasise that again.’* All these statements reinforce the insight that these customers understand the importance of the notification process. We can also conclude that these customers are motivated to get started with the steps because this shows that they do not underestimate the value of an infection notification.

Out of these respondents, four also indicated that they liked the steps. Respondent 1 indicated: *’I thought there were very good links in the mail where you could find certain programs. That was fantastic.’* Respondent 11 was also praising the steps: *’Very short and powerful by the way, good. For a layman also understandable.’* The employee of the ISP who drafted the steps gets compliments from respondents 12 and 33. They said: *’So whoever drafted the email did a good job.’* and *’All compliments to [name ISP], because A) I really liked to be notified and B) the steps were very clear.’*

Unfortunately, some customers were also less happy with the notification, specifically with the fact that if they didn’t remove the virus, their Internet would be shut down. For example, respondent 3, who was very enthusiastic about the notification itself, said: *’However, the immediate pressure that you have to solve it was unnecessary.’* Later in the interview he explained: *’I just found it very unpleasant that I immediately received an email with a threat in it stating: within so many days or we will shut you down. If you do that, you must either give an explanation of the severity of the risk, but I would rather say try to solve it within so many hours and then you can always send an email afterwards* *[with the sanction stated].'* A respondent said about it: *'Back then we were not so dependent on the Internet, but if it happened now, it is worse. We all work from home of course. It was not that bad then, but it took a long time [before we had Internet again]. [...]. But that's probably just the process.'* Respondent 42 could even imagine a doomsday scenario in which a company is affected. He said about this: *'For me it was still private, but if you are then locked out for business then...'* This customer does not know that business customers do not get notifications when it comes to QSnatch. Table 46 includes the findings.

expression	*respondent*	*total*
distrust	6, 12, 14, 15, 19, 27, 32, 33, 46, 47, 50, 58	12
happy with the notification	1, 3, 4, 5, 6, 8, 11, 12, 14, 20, 24, 33, 39, 52, 54	15
nice steps	3, 11, 12, 33	4
sanction	3, 10, 42	3

Table 46: Results on the expressions of the respondents

Respondent 2 argued that if he could not understand the steps, someone else without knowledge about IT would most certainly not see the steps or understand them. Therefore, he reasoned: *'So, maybe it is more convenient if the communication is going to take place differently.'* Respondent 25 agrees with the reasoning of some notifications not getting to the destined readers and indicated: *'I also think that some emails simply end up in some sort of spam mailbox and that is never checked. I would always advise that if there is no response to call or send a letter.'* We probably do not need to clarify further that sending letters is not efficient and fast enough to implement it as an improvement, but we can extract an interesting insight from his view. The ISP has several channels to reach customers. For example, the ISP has a telephone number to call or send SMS; an application where customers can now view their usage, subscriptions, and the like; and an email address to send emails. Respondent 33, for example, stated the following about informing the customers about the interview: *'The email is the first thing and [name of ISP] has my mobile number, or that of my wife. For example, you could [...] send a text message to indicate that I will be called by this and that person.'* Respondent 46 also gave the same tip to first inform customers with a text message. Respondent 50 also had the suggestion to refer in the email to the ISP's official application and to explain the steps there. He said: *'I find that very hard to say, but maybe by referring to the [ISP name] portal. And I think that if an email is sent from: check in your account at [name ISP] without links and see a step-by-step plan. [...] Log in to prevent certain complications and that [the customer] will then receive information there.'*

In line with his previous statements about the vehemence of immediately threatening as an ISP that the Internet will be shut down, respondent 3 suggested building in an intermediate step. If it were up to him, the ISP would first send an email to indicate that something is wrong and only indicate in a subsequent email that there are consequences for not performing the steps. Although there were not many customers who expressed a negative attitude towards the sanctions announced in the notification, it is important for the ISP that the customer experiences the notification as a service and not as harassment. Ultimately, it is a welcome service that the ISP informs customers about possible infections, but then the ISP has to package it as something positive. Customers who would be less aware of the possibility of infection or the dangers of a compromised device will be a lot less happy if they are shut down at a moment's notice. That response can only be strengthened if the customer himself hardly notices the infection, since the performance of the device does not always have to decrease. In fact, more often than not, the performance will remain the same.

Another suggestion came from respondent 12. He said: *'It is also nice to know that something is being checked. Now I was actually only confronted with it when something was wrong. That is nice in itself, because then you don't get any unnecessary ballast, but you could say: We monitor the traffic at that address. And that can be done on an annual basis if not much happens every month. If not much happens, not much needs to happen, but then you are already familiar. It came as a bit of a surprise to me now and it was a surprise in a positive sense, because I also know: if I hadn't heard anything, that thing would have just kept going.'* As mentioned earlier, the ISP is not actively scanning the network for QSnatch-infected devices. Moreover, an important limitation is that the ISP cannot determine that a device is infected with malware. Therefore, it is not wise to inform customers that the ISP is searching for this type of malware. After all, there are no guarantees that the ISP will get through all infected devices. In addition, if the ISP sends notifications that they have not found malware on the customer's network, the customer could get the impression that there is no malware on their devices when there could be. That could then lead to customers who are less wary, even though that should be necessary.

There are customers who do not trust the notification email. Can the ISP ensure that the trust is there by sending out a message well in advance that the email address used is valid and that notifications will be sent from there when malware is detected? Then, the message will no longer have to be: Beware you have malware on your device and especially click on this link to fix it. But then, the first message could be something like: We cannot and will not detect all malware, but if we find something, you will hear from us this way. Then, customers will not first wait for a number of reports before taking action as respondent 50 did, but immediately understand the seriousness of the situation in order to take action.

Respondent 28, who figured out to use an SSH script to remove the malware after it turned out that the Malware Remover was not able to do it, suggested an addition to the steps: *'It would be great if there was a very good manual that always works. That I don't have to search for that script again first.'* Naturally, it would be best to provide the customers with all possible information, but at the same time, the ISP does not want to overload the customer with information. The chances are that the customer would be overwhelmed with information and figures that he no longer has the confidence to be able to solve the problem and, therefore, gives up. Another addition to the steps came from respondents 37 and 40. They advised involving QNAP in the process. Respondent 40 argued: *'Maybe it would be useful if you also say in that article that there is a knowledge base at QNAP and that they can also look there.'* Respondent 37 also indicated: *'[ISP name] could also have said: You can contact the QNAP helpdesk. That is the easiest for everyone, and then [the customer] will be helped the next day. Only yes, if they say so, maybe [name of ISP] should consult with QNAP.'* It is not at all a strange thought to involve QNAP more closely in the cleaning process. After all, they are also the party that initially supplies devices with weaknesses. In addition, they are also the party that has the most knowledge of the devices. In any case, more knowledge than the ISP that has to focus on a wide range of malware in addition to QSnatch.

According to respondent 33, the style of the mail could still be improved. He indicated: *'Well, I found him to have very little of [ISP name] in terms of style. That could be an improvement. And another, [...] would be to give a slightly different introduction to how [name ISP] gets the information. That might also make the message a bit more believable. This can be realised by more links to their own website where the team in question or the service is further explained. That would improve credibility.'* In line with this reasoning, respondent 46 argued: *'really clearly show the [name ISP] logo everywhere, [...] but that that is clear in any case.'*. Although these tips do not guarantee that the email is valid, any hacker can get this information or logos because it is freely available online, if it makes customers more inclined to take it seriously, then it is worth considering. Hackers, of course, also take all the time to imitate legitimate emails as well as possible. They would not if it would not be effective in tricking unsuspecting victims. Therefore, it makes sense to imitate the emails that customers are used to be receiving. Table 47 shows the findings on the respondents' suggestions.

suggestion	*respondent*	*total*
suggestion	2, 3, 12, 25, 28, 33, 36, 37, 40, 43, 46, 47, 50, 54	14

Table 47: Results on the suggestions of the respondents

5.2.11 Conclusions on the user experiences after receiving the infection notification

In this subsection, we analysed the user experiences according to the responses of 58 respondents and their abuse data. It became evident that the clean-up results of the interviewed customers were comparable to the wider population of the ISP's QSnatch-infected customers. We can therefore assume that the results that we find as the user experiences are representable for the experiences of the wider population of QSnatch-infected customers of the ISP. We will first discuss the experiences of the respondents without any grouping. With this, we address SQ2 of this research: *What are the customer's experiences of the QSnatch clean-up after they received the infection notification?* The answer to this question cannot be summed up in a single sentence. That is why the following paragraphs are devoted to it.

We have seen that most respondents were able to complete the steps without too many problems. That was quite inconsistent with the clean-up results of the subsection before it. It became clear that part of the difference in the clean-up rate could be explained by the time the ISP takes to notify customers. About half of the total infection time is the time between the first moment that the customers appear in the abuse data until the notification. It should be noted that even if we deduct this time from the total infection time, there is still a statistical difference between the clean-up results of the groups of malware. However, it remains an intriguing finding.

The analysis showed that the respondents had no physical limitations that made the steps challenging. From this, we conclude that they had sufficient physical capability to perform the steps and were not hindered by this aspect. The psychological capability was omnipresent. To illustrate, most respondents indicated that they understood steps and that they did not find the steps challenging. In addition, most respondents knew what malware is. In contrast, few respondents knew the difference between persistent and non-persistent malware. Since many have succeeded in performing the steps with desired results, this suggests that it is not essential to know what persistent malware is to eradicate it.

Most respondents were confident that they could perform the steps, but a striking observation is the amount of IT experience the respondents had. More than half of the respondents indicated that they had work experience, and a few more were at least IT enthusiasts. It could be that a NAS user, in general, has more experience, but we can say for sure that these are not your average customers of the ISP. Given the expansion of the NAS market, it could be that in the future, people with less IT affinity will come into contact with the QNAP devices. That could have serious consequences for the results of the QSnatch clean-up. This was not an issue during this research, and it can be said that the respondents had a high degree of psychological capability. Since both physical and psychological capability were widely present, this research concludes that the respondents had sufficient capability to perform the steps.

The motivation is made up of reflective and automatic processes. In the reflective processes, we looked at whether the respondents found the steps useful. There, it became clear that the respondents indeed understood the usefulness of the steps. The respondents were also able to indicate that harm in various forms could occur to the QNAP user and the network. The respondents also agreed that they were responsible for carrying out the steps.

The automatic processes were measured based on the feelings and impulses that the respondents experienced while executing the steps. In terms of feelings, most respondents were quite indifferent and said they did not feel that much in particular. As hypothesised, some respondents were surprised or shocked by the infection. Others experienced frustration mainly because the steps seemed to work with difficulty. Impulses seem to have had no influence in a third of the cases. Another one-third indicated that the email, in particular, had acted as an impulse. Although the email is not officially part of automatic processes, it is an intriguing insight from this part of the research that is discussed in the physical opportunity part.

The insights on the reflective and automatic processes suggest that the respondents were motivated to carry out the steps. The suggestion that the respondents were motivated was reinforced by the questions they asked at the end and during the interview. Questions about the interview, research, persistent malware, their security, other steps they could take to become even more secure indicate that even after a 10 to 15-minute interview, some were eager to learn more about the topic.

The respondents received the notification email in almost all cases. Only three respondents did not get the notification. In those cases, this did not lead to the respondent's failure to carry out the steps. As for the tools, most respondents did not use any in particular that should have had an impact on the clean-up. To illustrate, Google and the internet were used to determine the authenticity of the notification email and malware, and it goes without saying that a laptop or PC was used to perform the steps. Only the script that respondent 24 found online was probably decisive in that case to remove the malware. In addition, time and the location of the device or the tools used were no problem. In addition, time and the location of the device or the used tools were not a problem during the clean-up.

In most cases, the respondents did not get help to perform the steps. Some respondents did receive help from, for example, acquaintances, QNAP employees or an IT professional. In addition, many respondents indicated that they do not know people with strong opinions about performing the steps. Some indicated that they had not spoken to people about it, and only a few indicated that someone did have a strong opinion. All in all, we cannot say that social opportunity was present among the respondents. One could argue that the respondents did not need help to perform the steps. However, some respondents who only managed to remove the virus with great difficulty or not at all did not get in touch with someone to help them out. That in itself is not surprising because the notification email does not suggest calling in help if you are unable to perform the steps yourself. Moreover, it might be a good thing if there were disgrace about people who do not secure their devices properly and leave malware on them. One might argue that a little peer pressure would help users keep their devices more organised.

Fascinating insights regarding opportunity came mainly from the expressions and suggestions of the respondents. As such, the respondents showed that they did not trust the email notification in many cases. For some, that had also led to them not taking the notification seriously and performing the steps only after multiple notifications. In addition, the steps did not seem to work in some cases. Only after a few tries, with the help of QNAP, or a separate script, those respondents were able to remove the malware successfully.

Then, we also grouped the respondents according to whether they used their devices for private or business purposes. In this way, we were able to provide the answer to the SQ3: *What are the differences in user experiences of the QSnatch clean-up of customers that use their devices for private and business purposes after receiving the infection notification?*

The group that uses their devices for business purposes seem to be more likely to take more time to perform the steps. The time between the first notification and the moment they are last seen in the data is longer than the group that use their devices for private purposes. In addition, the business users seem more inclined to call in help from a third party instead of carrying out the steps themselves.

When taking a look at the capability of the two groups, we see that both groups score relatively the same. On the questions about physical limitations, whether they found the steps challenging, whether they knew what malware and persistent malware is, and what their IT experience was, both groups provided roughly the same answers. However, the customers who used their devices for business purposes often could not remember whether they understood the steps. The main difference was that the customers who used their devices for business purposes more often did not feel confident being able to perform the steps themselves.

Both groups also scored about the same in terms of motivation. They responded to questions about whether they thought they were responsible and whether feelings had helped them perform the steps to about the same degree. Remarkably, the customers who used their devices for business purposes were more likely not to know whether the steps were useful, not sure what would happen if someone did not follow the steps, and were not sure whether an impulse had helped them complete the steps.

As for opportunity, both groups provided fairly the same answers. We were unable to find any differences regarding the physical opportunity with questions about whether they received the notification email, whether they used tools when completing the steps, whether they had enough time to complete the steps, and whether the location of the device or one of the tools they used was a problem to perform the steps. With regard to the social opportunity, we did see that the customers who used their devices more often called in the help of an IT professional compared to the group who used the devices for private purposes. In addition, the customers who use their devices for business purposes more often indicated that there were no people in their environment who had a strong opinion about performing the steps. Now that we have discussed the results of the study and answered the subquestions, it is now time to discuss the conclusions of the study. We will do this in the next chapter.

6 Conclusion

In this chapter, we will draw conclusions from the findings of the previous section. We have performed several analyses to answer the main research question: *'How does the QSnatch clean-up effectiveness compares to other types of malware, what are the customers' experiences with the clean-up procedure, and to which extent do they succeed?'* To answer this question, we have divided it into three sub-questions. In the following sections, we will conclude per subquestion on the findings found in the previous section.

6.1 RQ1: How does the QSnatch clean-up effectiveness compare to other types of malware such as non-persistent malware and non-IoT malware?

This research was interested in how the results of the QSnatch clean-up compares to other malware such as non-persistent malware and Windows malware. That is why the research consisted of plotting Kaplan-Meier curves and calculating logrank tests to determine whether the probability that a malware is still present on a device at a certain time is higher with the QSnatch malware compared to other malware. When the different malware were divided into the groups of persistent malware, non-persistent malware and Windows malware, it became evident that QSnatch does indeed remain longer on the devices. With the logrank tests, we were also able to statistically prove that there was a difference between the results of the groups of malware.This result is bad news for both the customers and the ISP. This shows that QSnatch-infect customer remain longer infected and vulnerable for data theft and other malicious intent. Moreover, this also means that the devices could be potentially longer used for malicious intent towards the network of the ISP. In the future, it might be possible for the hackers to use the infected devices for sending spam and DDoS attacks.

6.2 RQ2: What are the customer's experiences of the QSnatch clean-up after they received the infection notification?

To better understand what exactly happens in the process on the customer side, this research focused on their experience of the clean-up after they received the infection notification. In this pursuit, this research conducted interviews with infected or once infected customers of the ISP with questions based on the COM-B model. There was a paradox in the results that most respondents were able to complete the steps without too many problems while the clean-up results showed worse results for the QSnatch clean-up. A part of the difference in the clean-up rate could be explained by the time the ISP takes to notify customers.

The analysis showed that the respondents had no physical limitations that made the steps challenging. The respondents had much psychological capability. Strikingly, most respondents had experience in IT. The average customer of the ISP does not have that much experience and it is frightening to think want would happen if less-experienced customers get QNAP devices. In any case, for the respondents this research concludes that the respondents had sufficient capability to perform the steps.

As far as motivation is concerned, the respondents benefited a lot from their reflective processes to perform the steps, while little seemed to come from the automatic processes. This research would conclude that the the respondents were motivated to carry out the steps. This was reinforced by the questions many respondents asked at the end and during the interview.

In most cases, the respondents did receive the notification mail and did not need additional tools to perform the steps as well as the location of the device and time were no problem hindering the execution of the steps. This research would, therefore, assume the respondents had enough physical opportunity. However, the respondents also showed that they did not trust the email notification resulting in the respondents not taking the notification seriously in many cases. The respondents even indicated that they only took action after receiving the notification a few times. According to some respondents, the steps did not seem to remove the malware from their devices. This suggests that the firmware update and malware remover are not the solution for removing the malware in all cases. If the steps to remove the malware do not work, then it is next to impossible for a user to remove the malware. In this sense, the success of malware removal depends on whether the software provided by QNAP works. Therefore, the clean-up process depends on the software provided by QNAP.

Moreover, the respondents did not get help to perform the steps in most cases and indicated that they do not know people with strong opinions about performing the steps. In many cases, social opportunity could have been more present than it is now. In this way, people who have sought help came out of the steps where they could not do it themselves. That could have been an option for others as well. Overall, the results show that opportunity is an essential component in getting the customer to carry out the steps.

6.3 RQ3: What are the differences in user experiences of the QSnatch clean-up of customers that use their devices for private and business purposes after receiving the infection notification?

The research divided the interviewed customers up into two groups, where the first group used the devices for private purposes and the other for at least business purposes, i.e. for business purposes or both private and business purposes. The group that uses their devices for business purposes seem to take more time to perform the steps. The time between the first notification and the moment they are last seen in the data is longer than the group that use their devices for private purposes. In addition, the business users seem more inclined to call in help from a third party instead of carrying out the steps themselves. In terms of capability, both groups answered about the same. Only the customers who used their devices for business purposes were less sure that they could perform the steps themselves. Both groups also answered roughly the same to questions concerning motivation. There were some slight differences as the customers who used their devices for business purposes were more likely not to know whether the steps were useful, not sure what would happen if someone did not follow the steps, and were not sure whether an impulse had helped them complete the steps. As for opportunity, the respondents that use their devices for business purposes were more likely to get help from an IT professional and more often indicated that they did not know people from their environment who had a strong opinion about performing the steps.

So we were able to find minor differences between the customers who use their devices for private or business purposes. It is important to note that these differences have not been statistically proven. Nevertheless, there are minor differences between the two groups that need not differ much from each other in other respects. It is no coincidence that customers who used their devices for business purposes have opted for a CM subscription with the ISP. The devices used by the customers interviewed were mostly used in small companies employing only a few or just one person. One could argue that the small differences between the results in this research would increase as the devices are used by larger companies. It will be interesting to send the ISP's BM customers notifications and see what their experiences will be.

Therefore, the whole process is also dependent on QNAP. And now we find a critical difference between the clean-up of persistent and non-persistent malware. While the steps to remove non-persistent malware depend only on the actions of the infected customer, the steps to remove persistent malware also depend on QNAP's software. As this research has uncovered, this dependency can drastically hinder the clean-up process. This is a painful discovery as we already saw that QNAP does not always have the right incentives to ensure good security. In fact, the negligence of this company has put the customer in this situation in the first place.

7 Recommendations

At this point, the ISP has taken on the role of the bearer of bad news. They have become entangled in a process where they themselves have a poor view on the infections because they depend on other parties to determine them. They cannot do much more themselves because they are bound by privacy legislation. At the same time, there are NAS manufacturers who market vulnerable devices without security-by-design and customers who, in many cases, connect the devices without considering the security settings. Finally, they do get all the trouble if the steps they provide to their customers do not seem to work. 'Do not shoot the messenger' should probably be above every email that the Abuse Team sends to the infected customers. The only reason the ISP is involved in this process is that it can make the link between an IP address and a specific customer; Shadowserver determines the infections, QNAP provides the steps and software, and the customer must perform the steps. Not even to mention that the whole system only comes into action once a device has already been compromised. This should also be possible otherwise. Somehow, the ISP should spin the process in such a way that it forms a team with its customers against the hackers and NAS manufacturers. The study shows that the respondents themselves were able to come up with sound improvements and this chapter will be guided by those.

7.1 More capacity or more efficiency

As we have seen in the results, it appears that customers have to wait a long time for their notification from the ISP. On average, it amounts to half the total time a customer is infected with QSnatch until the customer is notified. Much progress could be made if there were more people to answer customer questions. In this way, more notifications can be sent to customers per day. Naturally, it is challenging to hire more employees and therefore, this research will also give several recommendations to increase the efficiency of the current team and process. We provide the following three suggestions specifically to increase the efficiency of the notification process. It is essential to note that these three recommendations do not specifically follow from the research, but logically they can provide an increase in efficiency.

7.1.1 Multiple information channels

A significant improvement is to have the notification take place over different channels. In this way, it can be prevented that people who do not read the given email box or who have entered an incorrect email address still receive the notification. But there are even more advantages because this way, the ISP can also increase the customer's trust. One could even argue it is insane that the ISP sends out emails informing its customers that a device in their network is infected with malware and providing a link to fix it. As the respondents also pointed out, everything looks like the email is spam or phishing. In addition, anyone who has received the email and has determined that it is legitimate might be less wary when he gets a non-legitimate email with a link. The benefits of this measure, therefore, come threefold.

The ISP currently has three channels to inform its customers. For example, the customer can be reached by sending SMS to the customer's telephone number, messages via the dedicated application of the ISP or the well-known notification emails. Via each channel, reference can be made to the other media to increase reliability. In addition, the ISP can choose only to show the steps and links in the dedicated application. Then, the SMS and email will only contain a reference to the application and an incentive to the customer to check it as soon as possible.

7.1.2 Notify before the infection

One respondent indicated that he would have liked to get a notification before the infection took place. He especially wanted to know that the ISP is monitoring, but for that reason, it would not be wise to send such a notification. On the contrary, the ISP would like the customers to be attentive, and if the customer feels that the ISP is monitoring, they may become laxer and blame the ISP if something goes wrong. However, someone with a crisis management background would encourage this notification before the infection. It is wise to test the lines of communication before a crisis occurs. That way, the ISP can check if they are working. So, whether the emails come in and are being read by the customer.

A pre-notification should state that the ISP is not actively monitoring the network, but in some cases, is notified of possible infections. In addition, the ISP must state that customers will receive a notification via this email address the moment an infection is detected on the customer's network. To check the authenticity of the mail, the customer can check at that time whether the mail really comes from the ISP. Then, they do not have to do that when there is malware on the device. In addition, the first contact does not have to consist of: 'Your device is infected with QSnatch malware', which raises many suspicions. The ISP could send this email when a new customer signs up and message current customers in batches.

7.1.3 Intermediate step

Several respondents were shocked by the sanction for not taking the steps. By being fierce from the start, the ISP does not create a sense of solving the problem together and working as a team with its customer. Ultimately, it is crucial that the customer remains satisfied with the service of the ISP and in this way, the ISP especially displeases them. On the other hand, this study will in no way question the effectiveness of the measure. Well-founded previous research has already substantiated the usefulness of putting infected customers in quarantine. However, in this way, the ISP draws all the negative attention to itself, and that could lead to angry customers who might even consider switching to another ISP. Therefore, the ISP would do well to consider only naming the sanction after a second or third notification. The ISP could even choose to clearly state that the customers will then be quarantined for their own safety and that of the network. This study showed that many respondents were fully aware that the ISP was only trying to help them yet reacted angrily because they had been cornered right away. Hopefully, the ISP can create more understanding in that way.

7.2 Involve QNAP

The role of QNAP remains in the background during the current process. Customers have also indicated that they would have liked references to QNAP. The ISP's notification email may also refer to the steps provided by QNAP. These are a little more comprehensive than the ISP's steps. In addition, it is possible for the customers to contact the QNAP help desk. In some cases, QNAP staff have helped the customers, and it can be decisive for those customers where the steps are not sufficient. The study has shown that few customers are currently taking advantage of the help that QNAP could provide. The ISP can refer to QNAP and forward contact details in the email. It is advisable to contact QNAP first to make the forwarding process as smooth as possible.

Moreover, we discovered that the process is dependent on the software provided by QNAP. If it is not able to remove the malware, as some respondents have indicated, then the steps are pointless. Closer communication between the ISP's customers but also the employees could improve the process. In some cases, if QNAP's software does not seem to remove the malware, the ISP can quickly sound the alarm and QNAP will sooner know that there is a version of QSnatch on the loose that is resistant to the firmware update and malware remover. Likewise, the ISP could inform infected customers if QNAP is not able to provide the software to cover vulnerabilities to QSnatch.

7.3 Notify Business Market Customers

Although BM customers are informed in other cases of malware, in the current system of the ISP, they are not informed of any infections. The research did find that some QNAP devices were used for business purposes, but these parties, being small businesses, had a CM subscription. The larger companies do not even get a notification that they have an infected device in their network. When an attempt was made to contact these parties during the study, it turned out that the right contact person of the ISP could not even be found. In an emergency situation, the Abuse Team will have to contact the company using information freely available on the Internet. Not only will that be too slow, it will also come across as unprofessional. If a malicious person decides to lock away or encrypt files, it will be much more disastrous for a company than for an individual. A step forward would be to include BM customers in the system.

7.4 Involve the retailers

The respondents who remembered where they bought their QNAP device found that they generally bought it from a number of retailers, such as Alternate, Bol.com, and Mediamarkt. These retailers are in a unique position and have direct contact with a user the moment he decides to buy a QNAP device. The ISP could decide to also involve these parties or some of these parties in the process by convincing them of the importance of informing users in advance about the desired steps and settings to keep the devices safe. The retailer can inform the customer about the steps during the purchase, explain the importance and share a link to the steps. Hopefully, that will prevent users from getting infected in the first place.

8 Discussion

This chapter addresses the discussion on the research. Subsequently, it considers the research's validity, interprets the results, discusses the limitations and implications, and provides suggestions for future research. In this way, we will reflect on the research and pave the way for future scholars.

8.1 Validity

For the first part, the study used Kaplan-Meier curves. This is a well-known technique for mapping the life expectancy of certain things on the basis of entities that are alive and entities that have had a certain lifespan [160, 161]. Based on that, the research has made determinations of how QSnatch virus or persistent malware, in general, compares to non-persistent and Windows malware. Unfortunately, the study had no specific indication of when the malware was no longer present on a device. The determination was made on the basis of whether or not traffic was generated and after 30 days of inactivity, the study assumed that the malware was no longer present on the device. In addition, it could have been the case that customers in the data set have already been infected before, but no definitive answer is given in the data set itself. Finally, it was possible that the customer had removed the malware during the period that the data set consisted of, but that it was later infected again. This is also not included in the calculation of the curves. On the other hand, it was the case for all types of malware. Therefore, it could have affected the results if there were relatively more re-infections for specific malware.

With regard to determining whether the malware has been removed, the same can be said about the interviews. On the other hand, it was possible here to ask the respondents what had happened. That way it could be determined that some customers had not removed the malware, but had simply disabled the device. This gave us more focus on those cases. The questions of the interview are based on the COM-B model. This model has also made a reputation within the academic world. Previous studies have already used the model to construct interviews [143, 144, 145, 146, 147, 148].

It is also fair to mention that the research did not directly ask about matters such as motivation. The question was, therefore, not: 'Were you motivated to carry out the steps?' But the determination was made on the basis of several factors. So, the respondent found the steps helpful, the respondent could estimate what the consequences would be if someone did not perform the steps and the respondent felt responsible for carrying out the steps. Interviews were then held until no new insights emerged from the answers from the respondents.

On the other hand, the findings from the interviews are difficult to generalise to a larger population. For example, the respondents had much experience with IT and were mainly male. If we compare these insights with the population of the ISP, the first thing we see is that the ratio between men and women is much more evenly distributed. In addition, with the assumption that the Dutch population generally does not have years of work experience in IT, there is reason to assume that the respondents were not representative of a larger group of customers. It does raise the question of whether the respondents are representative of a larger group of NAS users. As described earlier, it could very well be that NAS users generally have experience in IT and are mainly male.

8.2 Interpretation

We will now discuss the results and how they relate to expectations and hypotheses mentioned earlier in this study. At the start of the study, persistent malware was expected to last longer on the devices compared to other types of malware, as more complex steps are presented by customers to remove it. Unfortunately, the research showed that persistent malware does indeed remain active longer on the devices. This could also mean that the steps are more difficult for the customers to perform.

Then, we made some hypotheses about the components of the COM-B model. For most components, the previously indicated expectations correspond to the results of the study. As expected, the component physical capability had little to no impact on the execution of the steps. Whether the psychological capability had much influence on the execution of the steps remains somewhat unclear. It was apparent that the respondents were highly knowledgeable. That will have had an impact on their performance. We did see that people who called in help generally had less knowledge and were less confident. It also turned out that it was unnecessary to know what persistent malware is to perform the steps.

From the notification email and tools, it quickly became apparent that it was, in some cases, also possible to perform the steps without the notification mail. However, it is challenging to give a definite answer about it, because almost all respondents had received it. Several respondents did note that they probably would not have a clue that there was malware on their device if the notification email had not arrived. What was very striking was that in many cases, the respondents were suspicious of the email notification. This went against the expectations of the study. In the beginning, it was expected that the respondents would trust the email, and that turned out to be wrong. In many cases, the respondents only trust the email after a while or after research by the respondent. The most important insight, however, was that the steps sometimes could not remove the malware. The respondents depended on the software

QNAP offers. The physical opportunity is, therefore, an essential component in the behaviour to perform the steps. In addition, respondents generally did not need any additional tools to perform the steps.

As discussed in a previous chapter, the work of Cetin et al. [36] that quarantining infected customers is very effective. The finding of this research that customers do not take the notification mails seriously could explain why the quarantining work so well. After all, the customers simply do not believe the emails are actually coming from the ISP until they are quarantined and actively convinced that it was really the ISP that sent the notification emails. From that moment on, the customers quickly take action and carry out the steps.

8.3 Limitations

Despite all the effort and passion that a researcher can put into research, no study will come without flaws. It is therefore vital to point out the limitations of this study. The limitations of this research can be divided into not-remediated customers in the interviews, the business customers, and the measurement of the automatic processes.

The main flaw of this study is the determination of whether the malware has been removed. The investigation assumed that the malware was removed after 30 days of inactivity. However, previously infected devices may no longer generate malware traffic for many reasons. For example, the device can be turned off, removed from the internet, and broken.

Another limitation of this research was that the datasets did not contain information about the same dates. During the analysis, one could decide to limit the information to the intersection of the datasets. Then, this research should have taken out some of the respondents that were not yet present in the datasets that were used for the K-M curves and vice versa.

Moreover, a limitation of the interviews is that few not remediated customers were involved. During the pilot study, many customers were enthusiastically contacted who had not yet removed the malware. It later turned out that the group that had not yet removed the malware was smaller than initially expected. That was because some of the customers who had not yet removed QSnatch were customers of the ISP's wholesale partners. Customers of these wholesale partners could not be approached and were therefore excluded as potential respondents. This inevitably disturbed the results of the study. Fortunately, respondents who succeeded in taking the steps did indicate what setbacks they encountered. This made it possible to draw conclusions from the results to a certain extent.

The third limitation of this study is the measuring of automatic processes. The results show that in few cases, these were mentioned by the respondents. Many respondents indicated that feelings and impulses had no influence on the process. But by definition, respondents should not be able to distinguish between them. After all, they are automatic processes. To ask them this after, in some cases, months is a bit too positive thought. It would have been better to be closer to the respondent after or even while performing the steps.

Finally, the BM customers could not be approached. As discussed, respondents were involved who used their device for business purposes but who had subscribed to a CM subscription. This may indicate that their devices are used by very small businesses. And in those kinds of settings, professionalism can sometimes differ from a larger company.

8.4 Future Research

After drawing up the limitations of this research, the suggestions for future research come as a logical consequence. Future researchers can conduct similar research and specifically avoid the limitations of this research. In this way, it can be examined whether, in a setup in which the researcher is also examining the device and the steps taken by the respondent, it is possible to determine in other ways when the device becomes infected and until when it remains infected. As discussed, there are many possible scenarios where infected devices stop generating malicious traffic. If a researcher can analyse the devices more closely, he can determine with more certainty whether the virus has been removed. In addition, this study has a bias towards customers who manage to remove the virus. Future researchers can engage more customers who no longer appear to be infected to overcome this issue.

In the same way, researchers could also pay more attention to involving BM customers. For this, adjustments must first be made in the process of the ISP so that BM customers are also informed of QSnatch infections. Future researchers could also address the limitation that this research tried to analyse the automatic processes of the customers by simply asking them. By definition, the respondent should be barely aware of these processes happening, let alone that they can still remember that after some time. Future research could partly overcome this by sitting closer to the customer as they perform the steps. A situation where the researcher can monitor the customer's processes while standing almost next to him would be more suitable to make determinations of this. In this way, the researcher could examine the customer's immediate reactions.

Beyond fixing the limitations of this research, there is an essential follow up to be made to scientifically find other improvements to those of this research. We already discussed ENISA's report [137] and its view on the COM-B model and the Fogg's model [141] in section 2. According to the ENISA, the COM-B model should be used to identify why the desired behaviour may be performed and Fogg's model to come to possible interventions. In line with this reasoning, this research would urge future researchers to use this study's insights on why desired behaviour may be performed and analyse interventions with Fogg's model. Moreover, future studies should also aim to substantiate the findings using an experimental setup comparing the current notification process to the improved version.

The research suggests that researchers should keep a close eye on the developments of persistent IoT malware. Based on the findings, the respondents of the interviews were very skilled and had much affinity with IT. Should other persistent IoT malware arise targeting devices of which the users are less educated, then nasty consequences can follow. Moreover, future research should address different types of persistent malware. In this research, we focused on IoT and Windows malware, while it could also be intriguing what the clean-up results would be of malware targeting, for example, mobile phones.

From this research, it becomes clear that the ISP's customers have a negative attitude towards being put in quarantine. When the ISP puts a customer in quarantine, the customer is likely to respond with a very hostile attitude. Even customers who clearly stated the importance of the quarantine and that they saw that it was in their best interest were often angry that they would be shut down. A researcher can examine the impact on customers' attitudes after they have been quarantined and see what consequences they have on the ISP's clientele.

8.5 Reflection

This section reflects on the Management of Technology (MOT) curriculum and this master thesis' researching process. But let's first look into the reason for me to study MOT in the first place. After completing the BSc Technische Bestuurskunde (TB), I really wanted to investigate complex problems from a business perspective instead of the governmental perspective that I used during the bachelor. MOT after TB is not the usual path to take and it was only possible to continue with MOT if a student did a technical minor. With a minor in Computer Science, I was able to pursue an MOT degree. And I am happy with this decision to this day. The business perspective was immediately reflected in the first block in subjects such as Financial Management and Leadership and Technology Management. Later that first year, this was only reinforced with courses such as High-Tech Marketing and Emerging and Breakthrough Technologies. In short, these courses taught me how to deal with finance, leadership, marketing and innovations. In my opinion, this made it easier for me to understand the different interests of the actors in the NAS market. I was also able to view the problem through the eyes of an ISP and with the knowledge I gained in the MOT courses.

Perhaps even more important were the insights I gained during the Digital Business Process Management course. There, I learned how to gain insight into the various processes of the ISP. Finally, I am very happy that I have chosen the Cyber Security specialisation. I was able to apply the knowledge from those courses on many occasions during my thesis. Economics of Cyber Security, Cyber Risk Management, Network Security, Governance of Cybersecurity, and Cyber Crime Science. As a result, I was able to weigh up various economic interests, map out risks, include technical network security, and identify criminal motives. Although certain insights did not make it through this research, they have laid the foundation for me to understand the QSnatch problem.

But I also look back on a challenging period. Two years ago, I decided to combine two master degrees and squeeze a study workload of 3 years in just 2. This has meant that there has been little time for other activities besides studying in the past two years. After an exam week of MOT, there was often still much work to do for the other program and vice versa. But the pinnacle of all this hustle was during the writing of this thesis. I found it difficult during the third period to also take three courses (one of which was no longer for the credits) in addition to the thesis. I would advise future students not to plan anything besides the thesis.

In addition, I would have liked to carry out experiments with customers by looking at whether better results are achieved based on the findings of the research. Some results have a stronger foundation than others, but none of them have now been tested in an experiment. Unfortunately, there was no time for that during this research. It will, in any case, be beneficial for KPN to be able to implement specific suggestions and to see for themselves based on the results whether the recommendations lead to faster clean-ups. However, it would have been even better to substantiate the suggestions scientifically even further in experiments.

Finally, I would like to say something about the supervision and guidance I have had during this project. KPN has a very warm and pleasant working atmosphere and, from the first moment, I felt very welcome. Two very nice KPN employees were always ready to help me with everything. We had a weekly meeting to discuss progress and they kept a close track of where I was in my research. Whenever I had questions, they were there for me to help me. Without their assistance, this study probably would not have been possible. That also applies to the supervision I received from Delft. Sometimes it was pretty impressive that so much research had already been done before my research into subjects that touch on the subject of this research. Based on stories from fellow students, I realise all too well how lucky I have been with the guidance I have received. I have always received a quick response to questions I had and the feedback I

received during feedback moments was very extensive. Without all this guidance, I probably would not have been able to complete this research.

References

[1] S. Michie, M. M. Van Stralen, and R. West, “The behaviour change wheel: a new method for characterising and designing behaviour change interventions,” *Implementation science*, vol. 6, no. 1, pp. 1–12, 2011.

[2] Deloitte, “Internet of things (iot) the rise of the connected world,” 2000.

[3] B. Dorsemaine, J. Gaulier, J. Wary, N. Kheir, and P. Urien, “Internet of things: A definition taxonomy,” in *2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies*, pp. 72–77, 2015.

[4] V. Albino, U. Berardi, and R. M. Dangelico, “Smart cities: Definitions, dimensions, performance, and initiatives,” *Journal of Urban Technology*, vol. 22, no. 1, pp. 3–21, 2015.

[5] F. Dahlqvist, M. Patel, A. Rajko, and J. Shulman, “Growing opportunities in the internet of things,” 2019.

[6] G. Omale, “Gartner identifies top 10 strategic iot technologies and trends,” 2018.

[7] M. Silverio-Fernández, S. Renukappa, and S. Suresh, “What is a smart device? - a conceptualisation within the paradigm of the internet of things,” *Visualization in Engineering*, vol. 6, 2018.

[8] D. Bianchini and I. Avila, “Smart cities and their smart decisions: Ethical considerations,” *IEEE Technology and Society Magazine*, vol. 33, no. 1, pp. 34–40, 2014.

[9] O. Freiman, “Towards the epistemology of the internet of things techno-epistemology and ethical considerations through the prism of trust,” *International Review of Information Ethics*, vol. 22, pp. 6–22, 2014.

[10] G. Baldini, M. Botterman, R. Neisse, and M. Tallacchini, “Towards the epistemology of the internet of things techno-epistemology and ethical considerations through the prism of trust,” *Sci Eng Ethics*, vol. 24, pp. 905–925, 2018.

[11] X. Caron, R. Bosua, S. B. Maynard, and A. Ahmad, “The internet of things (iot) and its impact on individual privacy: An australian perspective,” *Computer Law & Security Review*, vol. 32, no. 1, pp. 4 – 15, 2016.

[12] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, “Privacy in the internet of things: threats and challenges,” *Security and Communication Networks*, vol. 7, no. 12, pp. 2728–2742, 2014.

[13] F. Allhoff and A. Henschke, “The internet of things: Foundational ethical issues,” *Internet of Things*, vol. 1-2, pp. 55 – 66, 2018.

[14] A. Chaudhuri, “Philosophical dimensions of information and ethics in the internet of things (iot) technology,” *EDPACS*, vol. 56, no. 4, pp. 7–18, 2017.

[15] A. Alcaide, E. Palomar, J. Montero-Castillo, and A. Ribagorda, “Anonymous authentication for privacy-preserving iot target-driven applications,” *Computers & Security*, vol. 37, pp. 111 – 123, 2013.

[16] N. Kshetri, “Cyber infrastructure protection volume iii,” tech. rep., Strategic Studies Institute, US Army War College, 2017.

[17] A. AboBakr and M. A. Azer, “Iot ethics challenges and legal issues,” in *2017 12th International Conference on Computer Engineering and Systems (ICCES)*, pp. 233–237, 2017.

[18] S. Tzafestas, “Ethics and law in the internet of things world,” *Smart Cities*, vol. 1, pp. 98 – 120, 2018.

[19] H. Atlam and G. Wills, “Iot security, privacy, safety and ethics,” in *Digital Twin Technologies and Smart Cities. Internet of Things (Technology, Communications and Computing)* (M. Farsi, A. Daneshkhah, A. Hosseinian-Far, and H. Jahankhani, eds.), pp. 266–290, Cham: Springer, 2019.

[20] S. Sholla, R. N. Mir, and M. A. Chishti, “Towards the design of ethics aware systems for the internet of things,” *China Communications*, vol. 17, no. 2, pp. 239–252, 2020.

[21] S. Eresheim, R. Luh, and S. Schrittwieser, “On the impact of kernel code vulnerabilities in iot devices,” in *2017 International Conference on Software Security and Assurance (ICSSA)*, pp. 1–5, 2017.

[22] S. Dange and M. Chatterjee, “Iot botnet: The largest threat to the iot network,” in *Data Communication and Networks* (L. Jain, G. Tsihrintzis, V. Balas, and D. Sharma, eds.), pp. 137–157, Springer, 2020.

[23] G. Kambourakis, C. Kolias, and A. Stavrou, “The mirai botnet and the iot zombie armies,” in *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, pp. 267–272, 2017.

[24] C. Brierley, J. Pont, B. Arief, D. J. Barnes, and J. C. Hernandez-Castro, “Persistence in linux-based iot malware,” 2020.

[25] C. Cimpanu, “Thousands of qnap nas devices have been infected with the qsnatch malware,” 2019.

[26] L. Tung, “Ransomware crooks hit synology nas devices with brute-force password attacks,” 2019.

[27] D. Palmer, “This new ransomware is targeting network attached storage devices,” 2019.

[28] C. Cimpanu, “White-hat hacks muhstik ransomware gang and releases decryption keys,” 2019.

[29] C. Cimpanu, “Qnap tells nas users to update firmware to avoid new type of ransomware,” 2020.

[30] C. Cimpanu, “Qnap warns users of a new crypto-miner named dovecat infecting their devices,” 2021.

[31] Vegelien, “Synology waarschuwt voor malware die nas-apparaten infecteert,” 2021.

[32] Synology, “Synology® investigates ongoing brute-force attacks from botnet,” 2021.

[33] C. Cimpanu, “Cisa says 62,000 qnap nas devices have been infected with the qsnatch malware,” 2020.

[34] S. Gatlan, “Qnap urges users to update malware remover after qsnatch alert,” 2020.

[35] O. Çetin, C. Ganán, L. Altena, T. Kasama, D. Inoue, K. Tamiya, Y. Tie, K. Yoshioka, and M. van Eeten, “Cleaning up the internet of evil things: Real-world evidence on isp and consumer efforts to remove mirai.,” in *NDSS*, 2019.

[36] O. Cetin, C. Ganán, L. Altena, S. Tajalizadehkhoob, and M. van Eeten, “Let me out! evaluating the effectiveness of quarantining compromised users in walled gardens,” in *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*, pp. 251–263, 2018.

[37] Council of European Union, “Council regulation (EU) no 1211/2009,” 2009. https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32009R1211.

[38] BEREC, “Berec guidelines on the implementation by national regulators of european net neutrality rules,” 2016. https://berec.europa.eu/eng/document_register/subject_matter/berec/regulatory_best_practices/guidelines/6160-berec-guidelines-on-the-implementation-by-national-regulators-of-european-net-neutrality-rules.

[39] BEREC, “What is berec?,” n.d.

[40] A. C. en Markt, “Over ons,” n.d.

[41] A. Telecom, “Over agentschap telecom.”

[42] Staatscourant, “Samenwerkingsprotocol autoriteit consument en markt & agentschap telecom,” 2015.

[43] Telecommunicatiewet, 2020. https://wetten.overheid.nl/BWBR0009950/2020-12-21.

[44] Wet telecommunicatievoorzieningen BES, 2019. https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32009R1211.

[45] A. C. en Markt, “Wetten en regels voor de telecommarkt,” n.d.

[46] E. Tjong Thjin Tai, “Aansprakelijkheid voor robots en algoritmes,” 2017.

[47] E. Tjong Thjin Tai and B.-J. Koops, “Zorgplichten tegen cybercrime,” 2015.

[48] Burgerlijk Wetboek, 2021. https://wetten.overheid.nl/BWBR0002656/2021-01-01.

[49] ACM, “Vervolg invulling ’zorgplicht internetveiligheid’ (artikel 11.3 telecommunicatiewet),” 2007.

[50] KPN, “Algemene voorwaarden voor vaste en mobiele telecommunicatiediensten,” 2020.

[51] ACM, “Een klacht over mijn telecomaanbieder,” n.d. https://www.consuwijzer.nl/telefoon-internet-en-televisie/klacht-over-telecomaanbieder/geen-of-slecht-bereik.

[52] ACM, “Uw recht halen,” n.d. https://www.consuwijzer.nl/uw-recht-halen.

[53] ACM, “Wilt u een slim apparaat kopen? zorg dat u voldoende informatie krijgt,” 2019-11-26. https://www.consuwijzer.nl/nieuws/wilt-u-een-slim-apparaat-kopen-zorg-dat-u-voldoende-informatie-krijgt.

[54] ACM, “Over acm consuwijzer,” n.d. https://www.consuwijzer.nl/over-acm-consuwijzer.

[55] F. B. Insights, “Network-attached storage market,” 2020.

[56] M. R. Future, “Global network-attached storage (nas) market research report,” 2019.

[57] G. V. Research, “Consumer nas market size worth $8.20 billion by 2025 | cagr: 15.8%,” 2017.

[58] PCM, “Nas kopen: Waar moet je op letten?,” 2018.

[59] C. Totaal, “Wat kun je precies met een nas?,” 2021.

[60] Computest, “Onze visie op de beveiliging van iot,” n.d.

[61] CSR, “Cybersecurity guide ’every business has duties of care in the field of cyber security’,” 2017.

[62] Qbit, “Essential requirements for securing iot consumer devices,” 2020.

[63] NIST, “Digital identity guidelines: Authentication and lifecycle management,” 2017.

[64] Shadowserver, “Who we are,” n.d.

[65] J. Leonard, S. Xu, and R. Sandhu, “A first step towards characterizing stealthy botnets,” in *2009 International Conference on Availability, Reliability and Security*, pp. 106–113, 2009.

[66] G. Vormayr, T. Zseby, and J. Fabini, “Botnet communication patterns,” *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2768–2796, 2017.

[67] H. Pieterse and M. S. Olivier, “Android botnets on the rise: Trends and characteristics,” in *2012 Information Security for South Africa*, pp. 1–5, 2012.

[68] S. Khattak, N. R. Ramay, K. R. Khan, A. A. Syed, and S. A. Khayam, “A taxonomy of botnet behavior, detection, and defense,” *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 898–924, 2014.

[69] J. Kohout and T. Pevný, “Unsupervised detection of malware in persistent web traffic,” in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1757–1761, 2015.

[70] F. Casino, N. Lykousas, I. Homoliak, C. Patsakis, and J. Hernandez-Castro, “Intercepting hail hydra: Real-time detection of algorithmically generated domains,” *arXiv preprint arXiv:2008.02507*, 2020.

[71] S. Upadhyay and A. Ghorbani, “Feature extraction approach to unearth domain generating algorithms (dgas),” in *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, pp. 399–405, 2020.

[72] S. Ray, S. Bhunia, Y. Jin, and M. Tehranipoor, “Security validation in iot space,” in *2016 IEEE 34th VLSI Test Symposium (VTS)*, pp. 1–1, 2016.

[73] S. Singh and N. Singh, “Internet of things (iot): Security challenges, business opportunities reference architecture for e-commerce,” in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, pp. 1577–1581, 2015.

[74] C. Kolias, A. Stavrou, and J. Voas, “Securely making "things" right,” *Computer*, vol. 48, no. 9, pp. 84–88, 2015.

[75] W. Zhou, Y. Jia, A. Peng, Y. Zhang, and P. Liu, “The effect of iot new features on security and privacy: New threats, existing solutions, and challenges yet to be solved,” *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1606–1616, 2019.

[76] R. Perdisci, T. Papastergiou, O. Alrawi, and M. Antonakakis, “Iotfinder: Efficient large-scale identification of iot devices via passive dns traffic analysis,” in *2020 IEEE European Symposium on Security and Privacy (EuroS P)*, pp. 474–489, 2020.

[77] J. Choi, A. Anwar, H. Alasmary, J. Spaulding, D. Nyang, and A. Mohaisen, “Iot malware ecosystem in the wild: a glimpse into analysis and exposures,” in *Proceedings of the 4th ACM/IEEE Symposium on Edge Computing*, pp. 413–418, 2019.

[78] D. Yin, L. Zhang, and K. Yang, “A ddos attack detection and mitigation with software-defined internet of things framework,” *IEEE Access*, vol. 6, pp. 24694–24705, 2018.

[79] E. Anthi, S. Ahmad, O. Rana, G. Theodorakopoulos, and P. Burnap, “Eclipseiot: A secure and adaptive hub for the internet of things,” *Computers & Security*, vol. 78, pp. 477 – 490, 2018.

[80] F. A. Alaba, M. Othman, I. A. T. Hashem, and F. Alotaibi, “Internet of things security: A survey,” *Journal of Network and Computer Applications*, vol. 88, pp. 10 – 28, 2017.

[81] S. S. Chawathe, “Monitoring iot networks for botnet activity,” in *2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)*, pp. 1–8, 2018.

[82] Z. Zhang, M. C. Y. Cho, C. Wang, C. Hsu, C. Chen, and S. Shieh, "Iot security: Ongoing challenges and research opportunities," in *2014 IEEE 7th International Conference on Service-Oriented Computing and Applications*, pp. 230–234, 2014.

[83] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "Ddos in the iot: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.

[84] A. Johnson, "Iot devices being increasingly used for ddos attacks," 2016.

[85] D. Lewis, "The ddos attack against dyn one year later," 2017.

[86] N. Woolf, "Ddos attack that disrupted internet was largest of its kind in history, experts say," 2016.

[87] A. Kashaf, V. Sekar, and Y. Agarwal, "Analyzing third party service dependencies in modern web services: Have we learned from the mirai-dyn incident?," in *Proceedings of the ACM Internet Measurement Conference*, IMC '20, (New York, NY, USA), p. 634–647, Association for Computing Machinery, 2020.

[88] N. Choucri and G. Agarwal, "Analytics for smart grid cybersecurity," in *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, pp. 1–3, IEEE, 2017.

[89] A. Wang, W. Chang, S. Chen, and A. Mohaisen, "Delving into internet ddos attacks by botnets: Characterization and analysis," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2843–2855, 2018.

[90] B. Nagpal, P. Sharma, N. Chauhan, and A. Panesar, "Ddos tools: Classification, analysis and comparison," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 342–346, 2015.

[91] S. Samonas and D. Coss, "The cia strikes back: Redefining confidentiality, integrity and availability in security.," *Journal of Information System Security*, vol. 10, no. 3, 2014.

[92] S. Newman, "Under the radar: the danger of stealthy ddos attacks," *Network Security*, vol. 2019, no. 2, pp. 18 – 19, 2019.

[93] M. J. G. van Eeten and J. M. Bauer, "Economics of malware," *OECD Science, Technology and Industry Working Papers*, 2008.

[94] S. J. Saidi, A. M. Mandalari, R. Kolcun, H. Haddadi, D. J. Dubois, D. Choffnes, G. Smaragdakis, and A. Feldmann, "A haystack full of needles: Scalable detection of iot devices in the wild," in *Proceedings of the ACM Internet Measurement Conference*, IMC '20, (New York, NY, USA), p. 87–100, Association for Computing Machinery, 2020.

[95] P. Wainwright and H. Kettani, "An analysis of botnet models," in *Proceedings of the 2019 3rd International Conference on Compute and Data Analysis*, ICCDA 2019, (New York, NY, USA), p. 116–121, Association for Computing Machinery, 2019.

[96] M. M. Salim, S. Rathore, and J. H. Park, "Distributed denial of service attacks and its defenses in iot: a survey," *The Journal of Supercomputing*, pp. 1–44, 2019.

[97] T. Mahjabin, Y. Xiao, T. Li, and C. L. P. Chen, "Load distributed and benign-bot mitigation methods for iot dns flood attacks," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 986–1000, 2020.

[98] S. Kirk, "The darlloz linux worm diversifies to mine cryptocurrencies," 2014.

[99] Q.-D. Ngo, H.-T. Nguyen, V.-H. Le, and D.-H. Nguyen, "A survey of iot malware and detection methods based on static features," *ICT Express*, vol. 6, no. 4, pp. 280 – 286, 2020.

[100] E. Cozzi, M. Graziano, Y. Fratantonio, and D. Balzarotti, "Understanding linux malware," in *2018 IEEE Symposium on Security and Privacy (SP)*, pp. 161–175, 2018.

[101] W. Chen, X. Helu, C. Jin, M. Zhang, H. Lu, Y. Sun, and Z. Tian, "Advanced persistent threat organization identification based on software gene of malware," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 12, p. e3884, 2020.

[102] J. C. Sapalo Sicato, P. K. Sharma, V. Loia, and J. H. Park, "Vpnfilter malware analysis on cyber threat in smart home network," *Applied Sciences*, vol. 9, no. 13, p. 2763, 2019.

[103] C. Brierley, J. Pont, B. Arief, D. J. Barnes, and J. Hernandez-Castro, "Paperw8: an iot bricking ransomware proof of concept," in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, pp. 1–10, 2020.

[104] Z. Gittins and M. Soltys, "Malware persistence mechanisms," *Procedia Computer Science*, vol. 176, pp. 88–97, 2020.

[105] M. F. Botacin, P. L. de Geus, and A. R. A. Grégio, "The other guys: automated analysis of marginalized malware," *Journal of Computer Virology and Hacking Techniques*, vol. 14, no. 1, pp. 87–98, 2018.

[106] B. Al-Duwairi, W. Al-Kahla, M. A. AlRefai, Y. Abdelqader, A. Rawash, and R. Fahmawi, "Siem-based detection and mitigation of iot-botnet ddos attacks.," *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 10, 2020.

[107] W. Li, S. Tug, W. Meng, and Y. Wang, "Designing collaborative blockchained signature-based intrusion detection in iot environments," *Future Generation Computer Systems*, vol. 96, pp. 481–489, 2019.

[108] O. P. Dwyer, A. K. Marnerides, V. Giotsas, and T. Mursch, "Profiling iot-based botnet traffic using dns," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2019.

[109] Y. Xu, H. Koide, D. V. Vargas, and K. Sakurai, "Tracing mirai malware in networked system," in *2018 Sixth International Symposium on Computing and Networking Workshops (CANDARW)*, pp. 534–538, 2018.

[110] Kurniabudi, B. Purnama, Sharipuddin, D. Stiawan, Darmawijoyo, and R. Budiarto, "Preprocessing and framework for unsupervised anomaly detection in iot: Work on progress," in *2018 International Conference on Electrical Engineering and Computer Science (ICECOS)*, pp. 345–350, 2018.

[111] E. Bertino and N. Islam, "Botnets and internet of things security," *Computer*, vol. 50, no. 2, pp. 76–79, 2017.

[112] I. Ko, D. Chambers, and E. Barrett, "Feature dynamic deep learning approach for ddos mitigation within the isp domain," *International Journal of Information Security*, vol. 19, no. 1, pp. 53–70, 2020.

[113] R. W. Rogers, "A protection motivation theory of fear appeals and attitude change1," *The journal of psychology*, vol. 91, no. 1, pp. 93–114, 1975.

[114] P. van Schaik, D. Jeske, J. Onibokun, L. Coventry, J. Jansen, and P. Kusev, "Risk perceptions of cyber-security and precautionary behaviour," *Computers in Human Behavior*, vol. 75, pp. 547–559, 2017.

[115] J. G. Adams, "Risk homeostasis and the purpose of safety regulation," *Ergonomics*, vol. 31, no. 4, pp. 407–428, 1988.

[116] G. Gigerenzer and P. M. Todd, *Simple heuristics that make us smart*. Oxford University Press, USA, 1999.

[117] D. Kahneman, *Thinking, fast and slow*. Macmillan, 2011.

[118] M. L. Finucane, A. Alhakami, P. Slovic, and S. M. Johnson, "The affect heuristic in judgments of risks and benefits," *Journal of behavioral decision making*, vol. 13, no. 1, pp. 1–17, 2000.

[119] C. Starr, "Social benefit versus technological risk," *Science*, pp. 1232–1238, 1969.

[120] P. Slovic, "Perception of risk," *Science*, vol. 236, no. 4799, pp. 280–285, 1987.

[121] P. M. Gollwitzer, "Implementation intentions: strong effects of simple plans.," *American psychologist*, vol. 54, no. 7, p. 493, 1999.

[122] L. Hadlington, "Employees attitudes towards cyber security and risky online behaviours: an empirical assessment in the united kingdom," *International Journal of Cyber Criminology*, 2018.

[123] J. D. Thompson, G. L. Herman, T. Scheponik, L. Oliva, A. Sherman, E. Golaszewski, D. Phatak, and K. Patsourakos, "Student misconceptions about cybersecurity concepts: Analysis of think-aloud interviews," *Journal of Cybersecurity Education, Research and Practice*, vol. 2018, no. 1, p. 5, 2018.

[124] S. Milgram, "Behavioral study of obedience.," *The Journal of abnormal and social psychology*, vol. 67, no. 4, p. 371, 1963.

[125] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *science*, vol. 185, no. 4157, pp. 1124–1131, 1974.

[126] S. L. Pfleeger and D. D. Caputo, "Leveraging behavioral science to mitigate cyber security risk," *Computers & security*, vol. 31, no. 4, pp. 597–611, 2012.

[127] D. G. Mayo and R. D. Hollander, *Acceptable evidence: Science and values in risk management*. Oxford University Press on Demand, 1991.

[128] P. E. Slovic, *The perception of risk*. Earthscan publications, 2000.

[129] M. Fishbein and I. Ajzen, “Belief, attitude, intention, and behavior: An introduction to theory and research,” 1977.

[130] I. Ajzen, “From intentions to actions: A theory of planned behavior,” in *Action control*, pp. 11–39, Springer, 1985.

[131] H. Yildirim and A. M. Ali-Eldin, “A model for predicting user intention to use wearable iot devices at the workplace,” *Journal of King Saud University-Computer and Information Sciences*, vol. 31, no. 4, pp. 497–505, 2019.

[132] F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS quarterly*, pp. 319–340, 1989.

[133] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, “User acceptance of information technology: Toward a unified view,” *MIS quarterly*, pp. 425–478, 2003.

[134] V. Venkatesh and F. D. Davis, “A theoretical extension of the technology acceptance model: Four longitudinal field studies,” *Management science*, vol. 46, no. 2, pp. 186–204, 2000.

[135] E. M. Redmiles, N. Warford, A. Jayanti, A. Koneru, S. Kross, M. Morales, R. Stevens, and M. L. Mazurek, “A comprehensive quality evaluation of security and privacy advice on the web,” in *29th USENIX Security Symposium (USENIX Security 20)*, pp. 89–108, USENIX Association, Aug. 2020.

[136] V. C. Conzola and M. S. Wogalter, “A communication–human information processing (c–hip) approach to warning effectiveness in the workplace,” *Journal of Risk Research*, vol. 4, no. 4, pp. 309–322, 2001.

[137] Enisa, “Cybersecurity culture guidelines: Behavioural aspects of cybersecurity,” tech. rep., Enisa, 2000.

[138] A. Bandura, “Self-efficacy: toward a unifying theory of behavioral change.,” *Psychological review*, vol. 84, no. 2, p. 191, 1977.

[139] F. Karlsson, M. Karlsson, and J. Åström, “Measuring employees’ compliance–the importance of value pluralism,” *Information & Computer Security*, 2017.

[140] S. Michie, L. Atkins, R. West, *et al.*, “The behaviour change wheel,” *A guide to designing interventions. 1st ed. Great Britain: Silverback Publishing*, pp. 1003–1010, 2014.

[141] B. J. Fogg, “A behavior model for persuasive design,” in *Proceedings of the 4th international Conference on Persuasive Technology*, pp. 1–7, 2009.

[142] C. Jackson, L. Eliasson, N. Barber, and J. Weinman, “Applying com-b to medication adherence: a suggested framework for research and interventions,” *European Health Psychologist*, vol. 16, no. 1, pp. 7–17, 2014.

[143] F. Barker, L. Atkins, and S. de Lusignan, “Applying the com-b behaviour model and behaviour change wheel to develop an intervention to improve hearing-aid use in adult auditory rehabilitation,” *International Journal of Audiology*, vol. 55, no. sup3, pp. S90–S98, 2016.

[144] K. Ekberg, S. Schuetz, B. Timmer, and L. Hickson, “Identifying barriers and facilitators to implementing family-centred care in adult audiology practices: a com-b interview study exploring staff perspectives,” *International Journal of Audiology*, vol. 59, no. 6, pp. 464–474, 2020.

[145] S. O. Ojo, D. P. Bailey, D. J. Hewson, and A. M. Chater, “Perceived barriers and facilitators to breaking up sitting time among desk-based office workers: a qualitative investigation using the tdf and com-b,” *International journal of environmental research and public health*, vol. 16, no. 16, p. 2903, 2019.

[146] E. K. Wakida, C. Obua, G. Z. Rukundo, S. Maling, Z. M. Talib, and E. S. Okello, “Barriers and facilitators to the integration of mental health services into primary healthcare: a qualitative study among ugandan primary care providers using the com-b framework,” *BMC health services research*, vol. 18, no. 1, pp. 1–12, 2018.

[147] C. Flannery, S. McHugh, A. E. Anaba, E. Clifford, M. O’Riordan, L. C. Kenny, F. M. McAuliffe, P. M. Kearney, and M. Byrne, “Enablers and barriers to physical activity in overweight and obese pregnant women: an analysis informed by the theoretical domains framework and com-b model,” *BMC pregnancy and childbirth*, vol. 18, no. 1, pp. 1–13, 2018.

[148] M. Courtenay, S. Rowbotham, R. Lim, S. Peters, K. Yates, and A. Chater, “Examining influences on antibiotic prescribing by nurse and pharmacist prescribers: a qualitative study using the theoretical domains framework and com-b,” *BMJ open*, vol. 9, no. 6, p. e029177, 2019.

[149] K. E. Alexander, B. Brijnath, and D. Mazza, “Barriers and enablers to delivery of the healthy kids check: an analysis informed by the theoretical domains framework and com-b model,” *Implementation Science*, vol. 9, no. 1, pp. 1–14, 2014.

[150] E. Kim, J. Yoon, J. Kwon, T. Liaw, and A. M. Agogino, “From innocent irene to parental patrick: Framing user characteristics and personas to design for cybersecurity,” in *Proceedings of the Design Society: International Conference on Engineering Design*, vol. 1, pp. 1773–1782, Cambridge University Press, 2019.

[151] G. Kabanda, “Performance of machine learning and other artificial intelligence paradigms in cybersecurity,” *Oriental journal of Computer Science and Technology*, vol. 13, no. 1, pp. 1–21, 2020.

[152] M. J. Dupuis and M. Weiss, “Veterans and their inherent cybersecurity preparedness: Myth or reality?,” in *2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 1841–1845, 2019.

[153] M. Denscombe, “Communities of practice: A research paradigm for the mixed methods approach,” *Journal of mixed methods research*, vol. 2, no. 3, pp. 270–283, 2008.

[154] D. T. Campbell and D. W. Fiske, “Convergent and discriminant validation by the multitrait-multimethod matrix.,” *Psychological bulletin*, vol. 56, no. 2, p. 81, 1959.

[155] J. Eugene, D. T. Campbell, R. D. Schwartz, and L. Sechrest, *Unobtrusive measures: Nonreactive research in the social sciences*. Rand McNally, 1966.

[156] N. Denzin, “Strategies of multiple triangulation,” *The research act in sociology: A theoretical introduction to sociological method*, vol. 297, no. 1970, p. 313, 1970.

[157] T. D. Jick, “Mixing qualitative and quantitative methods: Triangulation in action,” *Administrative science quarterly*, vol. 24, no. 4, pp. 602–611, 1979.

[158] M. L. Small, “How to conduct a mixed methods study: Recent trends in a rapidly growing literature,” *Annual review of sociology*, vol. 37, 2011.

[159] D. Cramer, *Advanced quantitative data analysis*. McGraw-Hill Education (UK), 2003.

[160] J. T. Rich, J. G. Neely, R. C. Paniello, C. C. Voelker, B. Nussenbaum, and E. W. Wang, “A practical guide to understanding kaplan-meier curves,” *Otolaryngology—Head and Neck Surgery*, vol. 143, no. 3, pp. 331–336, 2010.

[161] J. Ranstam and J. Cook, “Kaplan–meier curve,” *British Journal of Surgery*, vol. 104, no. 4, pp. 442–442, 2017.

[162] S. Kvale, *Doing interviews*. Sage, 2008.

[163] A. Galletta, *Mastering the semi-structured interview and beyond: From research design to analysis and publication*, vol. 18. NYU press, 2013.

[164] S. E. Baker and R. Edwards, “How many qualitative interviews is enough,” 2012.

[165] Z. A. Hassan, P. Schattner, and D. Mazza, “Doing a pilot study: why is it essential?,” *Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia*, vol. 1, no. 2-3, p. 70, 2006.

[166] M. Arain, M. J. Campbell, C. L. Cooper, and G. A. Lancaster, “What is a pilot or feasibility study? a review of current practice and editorial policy,” *BMC medical research methodology*, vol. 10, no. 1, pp. 1–7, 2010.

[167] M. Cope, “Transcripts: Coding and analysis,” *International encyclopedia of geography: People, the earth, environment and technology: people, the earth, environment and technology*, pp. 1–7, 2016.

[168] Merriam-Webster, “Impulse,” in *Merriam-Webster.com dictionary*.

[169] J. Jesson, L. Matheson, and F. Lacey, *Doing Your Literature Review: Traditional and Systematic Techniques*. SAGE Publications, 2011.

[170] J. Webster and R. T. Watson, “Analyzing the past to prepare for the future: Writing a literature review,” *MIS Quarterly*, vol. 26, no. 2, pp. xiii–xxiii, 2002.

[171] Y. Levy and T. J. Ellis, “A systems approach to conduct an effective literature review in support of information systems research.,” *Informing Science*, vol. 9, 2006.

[172] B. V. Wee and D. Banister, “How to write a literature review paper?,” *Transport Reviews*, vol. 36, no. 2, pp. 278–288, 2016.

[173] P. D. Leedy and J. E. Ormrod, *Practical research.* Pearson Custom, 2005.

Appendices

Appendix A QSnatch notification mail

The following is the exact text in the QSnatch notification mail of the ISP. Subsequently, this section provides the translation to English below. Finally, it includes the recommendated steps from QNAP.

A.1 The ISP's QSnatch notification in Dutch

Wat is er aan de hand en hoe kan ik dit oplossen? Een op uw internetverbinding aangesloten NAS van de leverancier Qnap is besmet met de QSnatch malware. Deze besmetting vormt een groot risico voor de veiligheid van uw bestanden op het apparaat. Het is belangrijk dat u handmatig het besturingssysteem van uw NAS en de malware remover app bijwerkt. Maak hierbij gebruik van de onderstaande stappen:

Besturingssysteem:

- Ga naar de website: qnap.com/nl-nl/download > Naar de website
- Selecteer onder '1 – Producttype' de optie 'NAS / Uitbreiding' en selecteer rechts het aantal aanwezige sleuven.
- Selecteer onder '3-Model' het type NAS dat u gebruikt.
- Kies onder het tabblad 'Besturingssysteem' voor de meest recente versie en download deze via de knop 'Europe'.
- Open de NAS op uw pc of Mac en kies voor ->Firmware-update -> Handmatige update.
- Blader naar het gedownloade bestand en update de firmware/besturingssysteem

Malware Remover app:

- Ga naar APP Center en kies voor 'Malware Remover' en download deze op uw pc of Mac.
- Klik op 'handmatige update' in App center, blader naar het downloadbestand en update de Malware Remover.
- Voer een scan uit met de Malware remover

Wat gebeurt er als ik niets doe?

Het beveiligingsprobleem op uw internetverbinding vormt een groot gevaar. Als u de stappen niet of niet goed uitvoert, bestaat de mogelijkheid dat wij uw internetverbinding in onze veilige omgeving (quarantaine) plaatsen. U kunt dan tijdelijk beperkt gebruikmaken van uw internetverbinding. Door dit te doen beschermen wij ook uw persoonlijke bestanden en gegevens.

Hebt u nog vragen?
Dan kunt u deze stellen in een antwoord op deze e-mail.

A.2 The ISP's QSnatch notification translated to English

What's going on and how can I fix it? A NAS from the supplier Qnap connected to your internet connection is infected with the QSnatch malware. This contamination poses a major risk to the safety of your files on the device. It is important to manually update your NAS operating system and malware remover app. Use the steps below:

Operating system:

- Go to the website: qnap.com/en-en/download > To the website
- Under "1 - Product type", select the option "NAS / Expansion" and select the number of slots present on the right.
- Under "3-Model", select the type of NAS you are using.
- Under the "Operating System" tab, select the most recent version and download it via the "Europe" button.
- Open the NAS on your PC or Mac and choose -> Firmware update -> Manual update.

- Browse to the downloaded file and update the firmware / operating system

Malware Remover app:

- Go to APP Center and choose "Malware Remover" and download it on your PC or Mac.
- Click on "manual update" in App center, browse to the download file and update the Malware Remover.
- Run a scan with the Malware remover

What happens if I don’t do anything?

The security problem on your Internet connection is a major threat. If you do not perform the steps or do not perform them correctly, we may place your internet connection in our secure environment (quarantine). You can then temporarily make limited use of your internet connection. By doing this we also protect your personal files and data.

Do you have any questions?
Then you can ask this in a reply to this e-mail.

A.3 QNAP’s QSnatch Recommendation

Recommendation

To prevent malware infections, we strongly recommend the following steps:

- Update QTS to the latest available version.
- Install and update Malware Remover to the latest version.
- Install and update Security Counselor to the latest version.
- Update your installed QTS applications to the latest versions if available in the App Center.
- Configure the following settings to enhance system security.

Important: QSnatch collects confidential information from infected devices, such as login credentials and system configuration. Due to these data breach concerns, QNAP devices that had been infected may still be vulnerable to reinfection after removing the malware. We strongly recommend applying these settings to further secure your system and to prevent reinfection.

- Change the admin password.
- Change other user passwords.
- Change QNAP ID password.
- Use a stronger database root password
- Remove unknown or suspicious accounts.
- Enable IP and account access protection to prevent brute force attacks.
- Disable SSH and Telnet connections if you are not using these services.
- Disable Web Server, SQL server or phpMyAdmin app if you are not using these applications.
- Remove malfunctioning, unknown, or suspicious apps
- Avoid using default port numbers, such as 22, 443, 80, 8080 and 8081.
- Disable Auto Router Configuration and Publish Services and restrict Access Control in myQNAPcloud.
- Subscribe to QNAP security newsletters.

Note:

- Malware Remover (supported by QTS 4.2 and later) and Security Counselor (supported by QTS 4.3.5 and later) may not be available on older QNAP NAS models. You can check the product support status of your NAS model.
- Installing Security Counselor helps further enhance the security of your NAS. Nevertheless, you can still protect your device from the QSnatch malware following other steps without Security Counselor.

—

Then, QNAP explains the steps in more detail. These have not been added in this appendix to save space and paper but can be found via the link: https://www.qnap.com/th-th/security-advisory/nas-201911-01

Appendix B Traditional Literature Review

This section consists of a literature review to gain a broad understanding and description of the field of the mitigation of IoT botnets. This pursuit is in line with a traditional review based on the description of Jesson et al. [169]. This methodology allows flexibility as this form of literature review does not require a defined path [169]. Another type of literature review is a systematic review. This review differs from the traditional review as it is the scientifically prescribed model to complement the traditional review. A traditional review's value is that it provides insights researchers can neglect or pass over in the steps towards exclusion and quality control for the systematic review. As this literature review intends to support a systematic literature review and future research, this section focused on the traditional review.

The traditional review also has five different approaches [169]:

- the critical approach assessing theories or hypotheses by critically examining the methods and results
- the conceptual approach aiming to combine conceptual knowledge and the understanding of issues
- the state-of-the-art review underlining the most recent knowledge on a subject
- the expert review which is written by an expert
- the scoping review which inspires future research

For each question, this literature review focused on a different approach. For each part of the review, it was primarily a conceptual approach to better understand the issues concerning IoT botnets. After that, the review also considered the methods and especially the results of the consulted researches.

Moreover, this review wanted a broad overview of the issues. Therefore, the analysis only included research from journals and proceedings and sorted searches on relevance. Thus, in this way, this research gave priority to studies with more relevance.

Inspired by the approach of Webster and Watson [170], this review focused first on contributions from leading journals. After that, it moved backwards by reviewing the citations for the articles from and went forward by using Google Scholar to identify articles that mention those contributions in leading journals. The researchers Levy and Ellis [171] called those three steps the keywords search, the backward search, and the forward search. Another research identified the last two steps as backward and forward snowballing, but those come down to the same things [172]. As proposed by Leedy and Ormrod [173], the end of this research review reached its completion when new articles only introduce familiar arguments, studies, authors, methodologies, and findings.

As the research of Webster et al., Levy et al., and Leedy et al. [170, 171, 173] all take a systematic approach and this research aimed to provide a traditional review, this review did not follow the steps as strict as the other researchers did. Therefore, this research did not include a rigorous and comprehensive search for all studies, predetermined criteria for including and excluding studies, and a tabular format and summary that typically belong to the systematic review. The next section will illustrates the keyword search. However, this research kept the backward and forward search to a minimum to keep the aim at gaining a broad understanding and grasping the bigger picture of the IoT malware issue. In doing so, this research hoped to prevent it from lingering too much in a series of like-minded studies.

Based on the top-ranked journals for management on information systems provided by Levy et al. [171], this review analysed documents from the literature vendors Elsevier (Scopus), IEEE (Xplore), and JSTOR. This review used different search combinations to find articles providing answers for the various aspects of IoT botnets. The search terms were: 1) botnet AND (concepts OR operat* OR workings OR characteristics), 2) iot OR 'internet of things' AND security and iot AND botnet AND security, 3), 4) IoT botnet AND detection AND (methods OR techniques), 5) IoT botnet AND (mitigation OR remediation), and 6) 'IoT botnet' AND mitigation AND (ISP OR 'internet service provider'). Whenever the search terms resulted in papers applicable to other questions, that was not a reason to exclude them.

Appendix C Systematic Literature Review

Like the traditional literature review, the following review is inspired by the methodologies of Webster et al., Levy et al., and Leedy et al. [170, 171, 173]. The difference is that this time, it has a systematic approach. For that part of the literature review, the analysis needed to be more in-depth and systematic in other to dive deep into those parts because those are the main elements of this research. Therefore, the keyword search was more extensive.

For the subsection about QSnatch and persistent malware, this research used 6 different search terms: 1) 'qsnatch', 2) 'qnap AND malware', 3) 'qnap', 4) 'persistent AND iot AND malware', 5) "persistent iot malware", and 6) "persistent malware". The review analysed the results of these terms on the sources Scopus, Xplote, and Google Scholar. The part of the review concerning behavioural theories started with an analysis on the literature provided as required readings for the course 'Behavioural Change Approaches to Cybersecurity' as part of the MSc Crisis and Security Management program of University Leiden. This resulted in 7 relevant documents. Then, it continued by used in the search terms 7) '(behaviour OR behavior) AND user AND iot AND device' and 8) '(abuse OR vulnerability) AND (notification OR notifying) AND security AND cyber)' in Scopus. Table 48 illustrates the results of the searches. It first shows the number of hits per search combination from different sources. Per search combination a maximum of 50 hits were studied. The reviewer then analysed the documents on their relevance and checked whether the paper was new to the research. Finally, the table shows the date on which the search was performed.

Source	*Keyword combination*	*Hits*	*Studied*	*Relevant*	*New*	*Date*
Scopus	qsnatch	0	0	0	0	15/02/2021
	qnap AND malware	0	0	0	0	15/02/2021
	qnap	19	19	0	0	15/02/2021
	persistent AND iot AND malware	22	4	4	4	16/02/2021
	"persistent iot malware"	0	0	0	0	16/02/2021
	"persistent malware"	9	9	0	0	16/02/2021
	(behaviour OR behavior) AND user AND iot AND device	606	50	4	3	18/02/2021
	(abuse OR vulnerability) AND ("notification" OR "notifying") AND security AND cyber)	9	9	0	0	21/02/2021
Xplore	qsnatch	0	0	0	0	15/02/2021
	qnap AND malware	0	0	0	0	15/02/2021
	qnap	12	12	0	0	15/02/2021
	persistent AND iot AND malware	9	9	2	0	16/02/2021
	"persistent iot malware"	0	0	0	0	16/02/2021
	"persistent malware"	1	1	1	1	16/02/2021
Google Scholar	qsnatch	6	6	2	2	16/02/2021
	qnap AND malware	7220	50	2	0	22/02/2021
	qnap	308	50	0	0	22/02/2021
	persistent AND iot AND malware	4750	50	4	3	22/02/2021
	"persistent iot malware"	5	5	2	1	23/03/2021
	"persistent malware"	240	50	3	3	25/03/2021

Table 48: Keyword combinations of the Systematic Literature review

Appendix D Mirai notification mail

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

Geachte heer, mevrouw,

Een veilig internet is in ieders belang. Wij maken ons als KPN sterk om uw (vertrouwelijke) informatie te beschermen.

Wij hebben een beveiligingsprobleem waargenomen op uw internetaansluiting. Meestal merkt u hier zelf niets van, omdat het om processen gaat die op de achtergrond draaien.

Wat is er aan de hand en hoe kunt u dit oplossen? Een of meer apparaten die zijn aangesloten op uw internetverbinding zijn geïnfecteerd met het Mirai virus. We kunnen niet met zekerheid zeggen welk apparaat geïnfecteerd is. Waarschijnlijk is het een digitale video recorder (DVR), beveiligingscamera of printer die op het internet is aangesloten en dus geen computer, laptop, tablet of mobiele telefoon.

Hoe kunt u het Mirai virus verwijderen en een infectie in de toekomst voorkomen? Volg onderstaande stappen. Mocht het niet lukken een stap uit te voeren, ga dan verder naar de volgende.

1. Bepaal welke apparaten zijn aangesloten op uw internetverbinding. Herinnering: Het Mirai virus infecteert met name op het internet aangesloten apparaten zoals een DVR, beveiligingscamera of printer.

2. Verander het wachtwoord van de op het internet aangesloten apparaten. Kies een wachtwoord dat moeilijk te raden is. Als u het huidige wachtwoord niet weet, raadpleeg dan de handleiding. Door het uitvoeren van deze stappen heeft u toekomstige infecties voorkomen.

3. Herstart de op het internet aangesloten apparaten door deze uit en opnieuw aan te zetten. Hierna is het Mirai virus verwijderd uit het geheugen van de apparaten.

Nu uw op het internet aangesloten apparaten veilig zijn, zijn de laatste stappen om uw router/modem te beschermen.

4. Reset uw modem/router naar de fabrieksinstellingen. Op https://kpn.com/reset-kpn-experiabox is beschreven hoe u dit kunt doen voor een Experia Box.

5. Stel het wachtwoord van uw modem/router in. Op https://www.kpn.com/faq/16176 is beschreven hoe u dit kunt doen voor een Experia Box.

Let op: Als toegang op afstand voor een apparaat absoluut noodzakelijk is, stel dan handmatig port forwards in op uw router voor het betreffende apparaat. Op https://forum.kpn.com/internet-9/port-forwarding-upnp-wat-waarom-en-hoe-322560 is beschreven hoe u dit kunt doen voor een Experia Box.

Wij vragen u de bovenstaande stappen binnen een dag uit te voeren en te reageren op dit bericht.

Ook aanvullende vragen kunt u stellen in een antwoord op deze mail.

LET OP: Het is belangrijk dat u zo spoedig mogelijk een reactie stuurt op deze waarschuwing.

Veilige omgeving

Indien blijkt dat de stappen binnen deze termijn niet (of onvoldoende) zijn uitgevoerd bestaat de mogelijkheid dat wij uw internetaansluiting in onze veilige omgeving (quarantaine) plaatsen. U kunt dan tijdelijk beperkt gebruik maken van uw internetaansluiting. Een dergelijke maatregel nemen wij ook om uw vertrouwelijke gegevens en bestanden te beschermen.

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

Dear Sir/Madam,

A safe internet is in everyone's interest. We, KPN, strongly care about protecting your (confidential) information.

We have observed a security issue on your internet connection. You probably have not noticed anything, because it's about processes that run in the background.

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

What should you do to remove the Mirai virus and prevent future infections? Please follow the steps below. If you cannot complete a step, please proceed to the next one. 1. Determine which devices are connected to your Internet connection. Reminder: The Mirai virus mainly infects Internet connected devices such as a DVR, security camera or printer connected to the Internet.

2. Change the password of the Internet connected devices. Choose a password that is hard to guess. If you do not know the current password, please refer to the manual. By following these steps, you have prevented future infections.

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

Now that your Internet connected devices are safe, the last steps are to protect your router/modem. 4. Reset your modem/router to the factory settings. On https://forum.kpn.com/internet-9/reset-de-kpn-experia-box-modem-97446#M8199 it is described how you do this for an Experia Box. 5. Set the password of your modem/router. On https://www.kpn.com/faq/16176 it is described how you do this for an Experia Box.

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

We ask you to take above steps within a day and to respond to this message. You can also ask additional questions in a reply to this email.

Appendix E Windows notification mail

The following text is exactly what customers receive when they are infected with one of the Windows malware. The section after that provides the translation to English.

E.1 Windows notification in Dutch

Geachte heer, mevrouw,

Een veilig internet is in ieders belang. Wij maken ons als KPN sterk om internetverbindingen veilig te houden. Hiervoor vragen wij uw hulp. Wij verzoeken u onderstaande stappen vandaag nog uit te voeren en ons hierover een bericht te sturen.

Waarom is mijn hulp nodig? Wij hebben een beveiligingsprobleem aangetroffen op uw internetverbinding. Hier merkt u zelf meestal niets van. Toch is het belangrijk om hier samen zo snel mogelijk iets aan te doen.

Wat is er aan de hand? Een of meerdere computers of laptops die gebruik maken van uw internetverbinding zijn besmet met een virus. Door deze besmetting lopen andere internetgebruikers en uw vertrouwelijke gegevens en bestanden gevaar.

Hoe kan ik dit oplossen? Doorloop de volgende stappen op al uw Windows computers/laptops: A. Ga naar onze website: kpn.com/virusverwijderen B. Klik op de knop ‘Malwarebytes’ C. Volg de instructies op het scherm

Wij kunnen niet zien om welke specifieke computer of laptop het gaat. Wanneer er meerdere zijn aangesloten, is het belangrijk dat deze allemaal worden onderzocht. Telefoons en tablets hoeft u niet te scannen.

Toekomstige problemen voorkomen? - Gebruik een up-to-date virusscanner. - Houd computersoftware, zoals uw besturingssysteem, up-to-date.

Wat gebeurt er als ik niets doe? Het beveiligingsprobleem op uw internetverbinding vormt een groot gevaar. Als u de stappen niet of niet goed uitvoert, bestaat de mogelijkheid dat wij uw internetverbinding in onze veilige omgeving (quarantaine) plaatsen. U kunt dan tijdelijk beperkt gebruikmaken van uw internetverbinding. Door dit te doen beschermen wij ook uw persoonlijke bestanden en gegevens.

Hebt u nog vragen? Dan kunt u deze stellen in een antwoord op deze e-mail.

E.2 Windows notification in English

Dear Sir / Madam,

A safe internet is in everyone’s interest. As KPN, we are committed to keeping internet connections safe. We ask for your help for this. We request that you complete the steps below today and send us a message.

Why is my help needed? We have found a security issue on your internet connection. You will usually not notice this yourself. Nevertheless, it is important to do something about this together as soon as possible.

What is going on? One or more computers or laptops that use your internet connection are infected with a virus. This infection puts other internet users and your confidential data and files at risk.

How can I solve this? Go through the following steps on all your Windows computers / laptops: A. Go to our website: kpn.com/virusremove B. Click the "Malwarebytes" button C. Follow the instructions on the screen

We cannot see which specific computer or laptop it is. When several are connected, it is important that they are all examined. You don’t have to scan phones and tablets.

Prevent future problems? - Use an up-to-date virus scanner. - Keep computer software, such as your operating system, up to date.

What happens if I don’t do anything? The security problem on your internet connection is a major danger. If you do not follow the steps or do not follow them properly, it is possible that we will place your internet connection in our safe environment (quarantine). You can then temporarily make limited use of your internet connection. By doing this we also protect your personal files and data.

Do you have any questions? Then you can ask this in a reply to this e-mail.

Appendix F Interview email

F.1 Mail remediated customers

Geachte heer/mevrouw [naam],
In het verleden hebben we een veiligheidsprobleem op uw internetverbinding ontdekt. Hierover hebben wij u destijds per e-mail geïnformeerd. Daarna kwamen uw gegevens niet langer voor in onze overzichten. Wij gaan er dan ook vanuit dat het probleem door u is opgelost. Het lastige is dat we niet zeker weten welke veiligheidsmaatregelen u precies heeft genomen. Graag zouden wij dit alsnog met u willen bespreken.
Om welk veiligheidsprobleem ging het ook alweer?
In het verleden is geconstateerd dat een apparaat in uw netwerk was geïnfecteerd met het [naam virus] virus. U heeft daarvan notificaties gekregen en het probleem lijkt daarna te zijn opgelost.
Graag bellen wij u op [beldag/datum]
Onze collega Max Fukkink belt u [beldag/datum] op om met u te bespreken welke maatregelen u heeft genomen. Wij doen dit om controleren of u ook echt veilig bent voor het virus en om ons proces voor andere klanten te verbeteren. Deze actie maakt deel uit van een onderzoek dat wij op dit moment samen met TU Delft verrichten. Met uw hulp zijn en blijven wij het veiligste netwerk van Nederland!
Heeft u nog vragen of bent u verhinderd?
Reageer dan op deze e-mail. Of stel uw vragen tijdens het telefoongesprek.

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

Dear Sir / Madam [name],
In the past we have discovered a security issue on your internet connection. We informed you about this by e-mail at the time. After that, your data no longer appeared in our overviews. We therefore assume that the problem has been solved by you. The tricky part is that we are not sure which security measures you have taken exactly. We would like to discuss this with you.
What security issue was this again?
In the past, it was found that a device on your network was infected with the [virus name] virus. You have been notified and the problem seems to have been resolved afterwards.
We would like to call you on [call day / date]
Our colleague Max Fukkink will call you [call day / date] to discuss with you what measures you have taken. We do this to check whether you are really safe from the virus and to improve our process for other customers. This promotion is part of a study that we are currently conducting together with TU Delft. With your help, we are and will remain the safest network in the Netherlands!
Do you have any questions or are you unable to attend?
Then respond to this email. Or ask your questions during the telephone conversation.

Sincerely,
KPN Abuse Team
abuse@kpn.com
The KPN department handles security incidents for KPN. More information about the department Abuse can be found at: https: //www.kpn.com/abuse

F.2 Mail not remediated customers

Geachte heer/mevrouw [naam],
In het verleden hebben we een veiligheidsprobleem op uw internetverbinding ontdekt. Hierover hebben wij u destijds per e-mail geïnformeerd. Het probleem lijkt nog steeds aanwezig aangezien uw gegevens nog steeds voorkomen in onze overzichten. Het lastige is dat we niet zeker weten welke veiligheidsmaatregelen u wellicht al heeft genomen. Graag zouden wij dit met u willen bespreken.
Om welk veiligheidsprobleem ging het ook alweer?
In het verleden is geconstateerd dat een apparaat in uw netwerk was geïnfecteerd met het [naam virus] virus. U heeft daarvan notificaties gekregen. Het probleem lijkt echter nog niet te zijn opgelost.

Graag bellen wij u op [beldag/datum]
Onze collega Max Fukkink belt u [beldag/datum] op om met u te bespreken welke maatregelen u heeft genomen. Wij doen dit om controleren of u ook echt veilig bent voor het virus en om ons proces voor andere klanten te verbeteren. Deze actie maakt deel uit van een onderzoek dat wij op dit moment samen met TU Delft verrichten. Met uw hulp zijn en blijven wij het veiligste netwerk van Nederland!
Heeft u nog vragen of bent u verhinderd?
Reageer dan op deze e-mail. Of stel uw vragen tijdens het telefoongesprek.
Met vriendelijke groet,
KPN Abuse Team
abuse@kpn.com
De afdeling van KPN handelt veiligheidsincidenten af voor KPN. Meer informatie over de afdeling Abuse vindt u op: https://www.kpn.com/abuse

Dear Sir / Madam [name],
In the past we have discovered a security issue on your internet connection. We informed you about this by e-mail at the time. The problem still seems to be there as your data is still in our overviews. The tricky part is that we are not sure what security measures you may have already taken. We would like to discuss this with you.
What security issue was this again?
In the past, it was found that a device on your network was infected with the [virus name] virus. You have been notified of this. However, the problem does not seem to be solved yet.
We would like to call you on [call day / date]
Our colleague Max Fukkink will call you [call day / date] to discuss with you what measures you have taken. We do this to check whether you are really safe from the virus and to improve our process for other customers. This promotion is part of a study that we are currently conducting together with TU Delft. With your help, we are and will remain the safest network in the Netherlands!
Do you have any questions or are you unable to attend?
Then respond to this email. Or ask your questions during the telephone conversation.
Sincerely,
KPN Abuse Team
abuse@kpn.com
The KPN department handles security incidents for KPN. More information about the department Abuse can be found at: https: //www.kpn.com/abuse

Appendix G Informed consent form

You will be interviewed as part of a research project on persistent IoT malware by KPN. This research aims to get a better understanding of user comprehension about the clean up procedure. With this research, ISPs could inform their customers more adequate. With your participation, this study might deliver more security to IoT users like yourself.

What you agree with:
The interview will (upon agreement will not) be recorded and a transcript will be produced.
Upon agreement the transcript can be send for correction of any factual errors.
Access to the interview transcript will be limited to Max Fukkink and academic colleagues and researchers from the TU Delft with whom he/she might collaborate as part of the research process.
I also understand that my words may be directly quoted and the results will be shared in an academic publication for a master thesis, journal, or conference. Direct quotes will not include names.

Interviewees are participating in this research voluntarily and are free to answer questions or not. At any time they may decide to end participation. This conversation is anonymous and the research team will take care that participants and organizations cannot be identified.

Verbal or written permission will be sought from each participant.

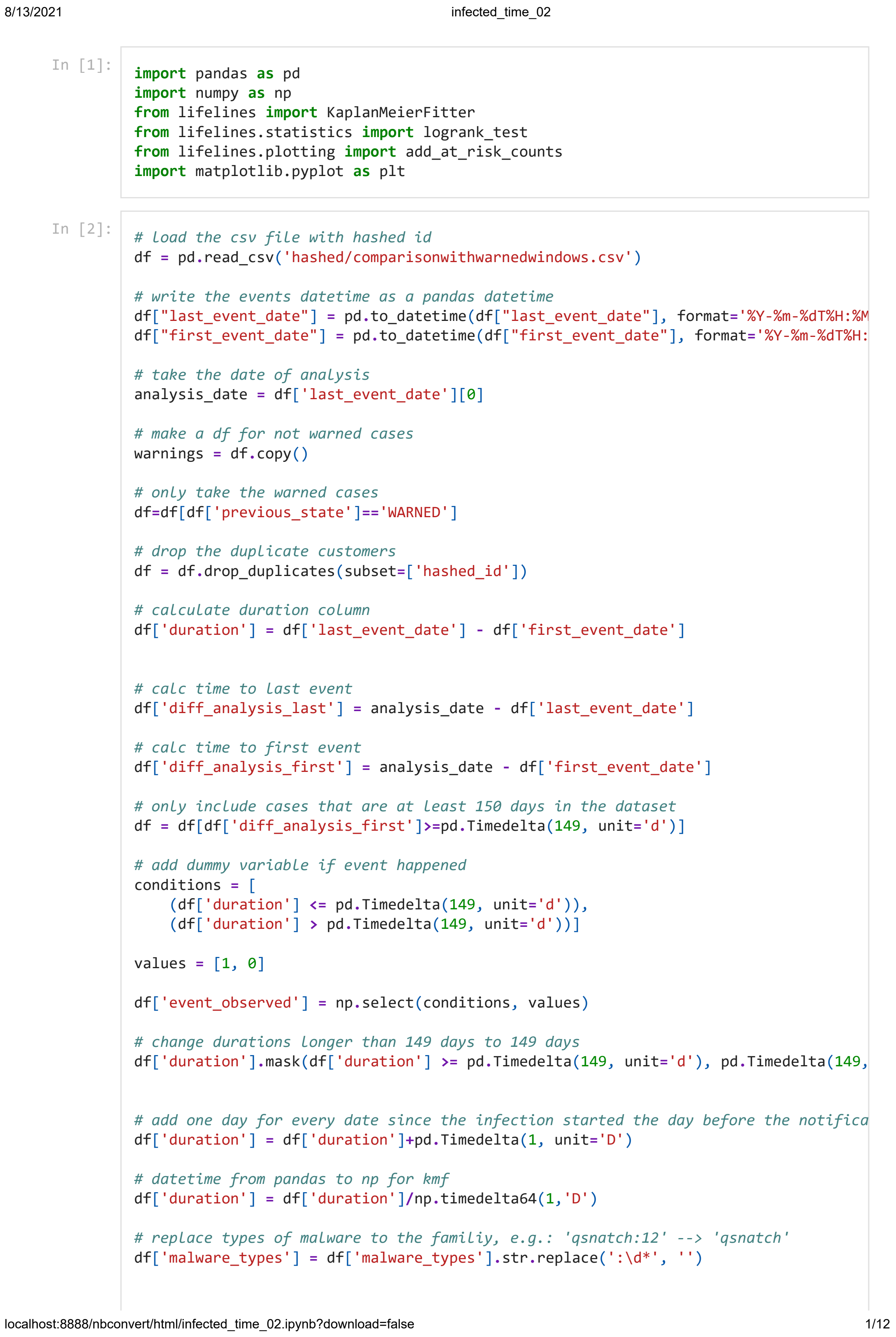
Please place a cross in the box that is applicable consent to the use of the information collected about me for this research project

Date

Name

Signature

Appendix H The Python program used for plotting the K-M curves and calculating the LogRank tests



8/13/2021 infected_time_02

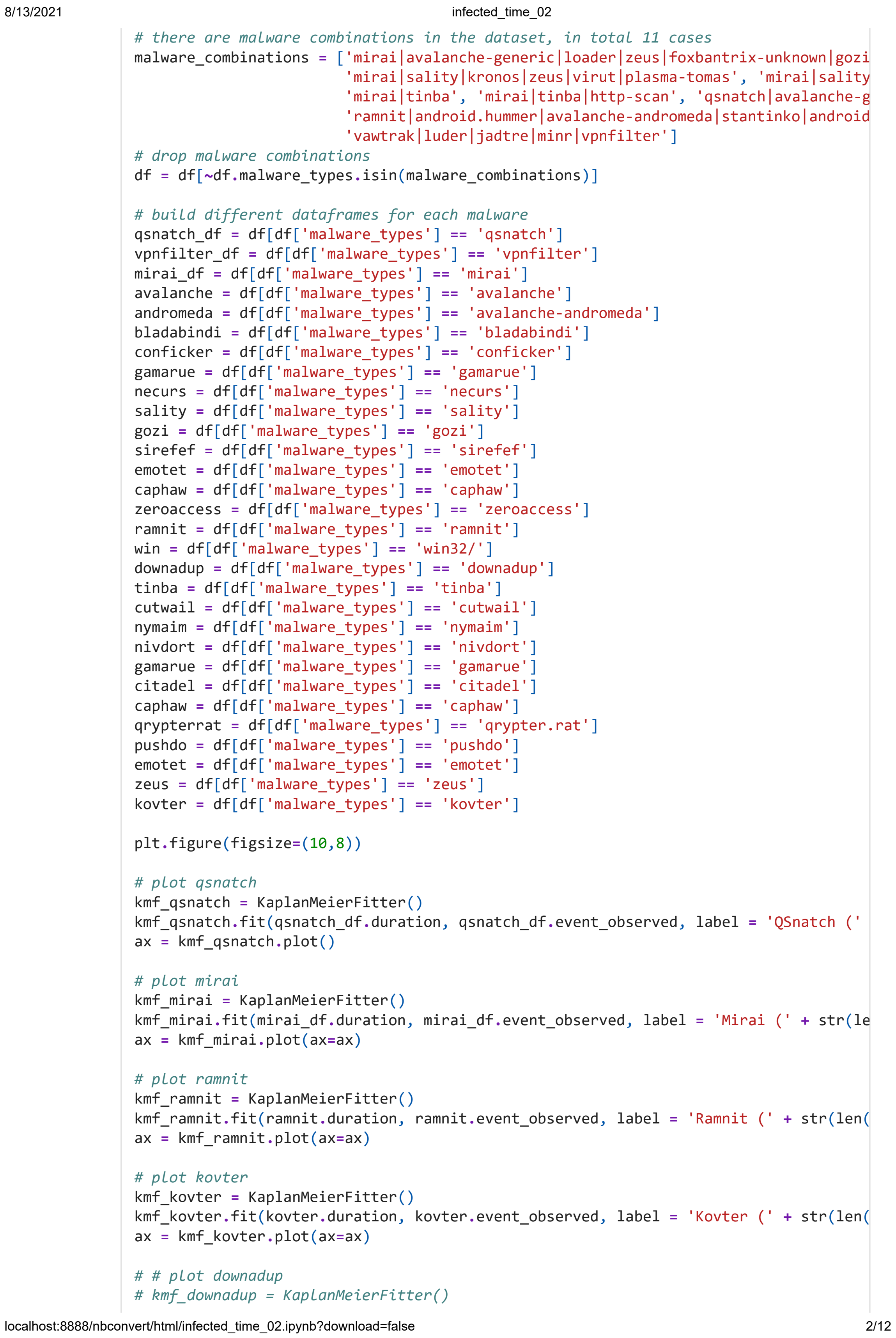
In [1]:

```
import pandas as pd
import numpy as np
from lifelines import KaplanMeierFitter
from lifelines.statistics import logrank_test
from lifelines.plotting import add_at_risk_counts
import matplotlib.pyplot as plt
```

In [2]:

```
# load the csv file with hashed id
df = pd.read_csv('hashed/comparisonwithwarnedwindows.csv')

# write the events datetime as a pandas datetime
df["last_event_date"] = pd.to_datetime(df["last_event_date"], format='%Y-%m-%dT%H:%M
df["first_event_date"] = pd.to_datetime(df["first_event_date"], format='%Y-%m-%dT%H:

# take the date of analysis
analysis_date = df['last_event_date'][0]

# make a df for not warned cases
warnings = df.copy()

# only take the warned cases
df=df[df['previous_state']=='WARNED']

# drop the duplicate customers
df = df.drop_duplicates(subset=['hashed_id'])

# calculate duration column
df['duration'] = df['last_event_date'] - df['first_event_date']


# calc time to last event
df['diff_analysis_last'] = analysis_date - df['last_event_date']

# calc time to first event
df['diff_analysis_first'] = analysis_date - df['first_event_date']

# only include cases that are at least 150 days in the dataset
df = df[df['diff_analysis_first']>=pd.Timedelta(149, unit='d')]

# add dummy variable if event happened
conditions = [
    (df['duration'] <= pd.Timedelta(149, unit='d')),
    (df['duration'] > pd.Timedelta(149, unit='d'))]

values = [1, 0]

df['event_observed'] = np.select(conditions, values)

# change durations longer than 149 days to 149 days
df['duration'].mask(df['duration'] >= pd.Timedelta(149, unit='d'), pd.Timedelta(149,


# add one day for every date since the infection started the day before the notifica
df['duration'] = df['duration']+pd.Timedelta(1, unit='D')

# datetime from pandas to np for kmf
df['duration'] = df['duration']/np.timedelta64(1,'D')

# replace types of malware to the familiy, e.g.: 'qsnatch:12' --> 'qsnatch'
df['malware_types'] = df['malware_types'].str.replace(':\d*', '')
```

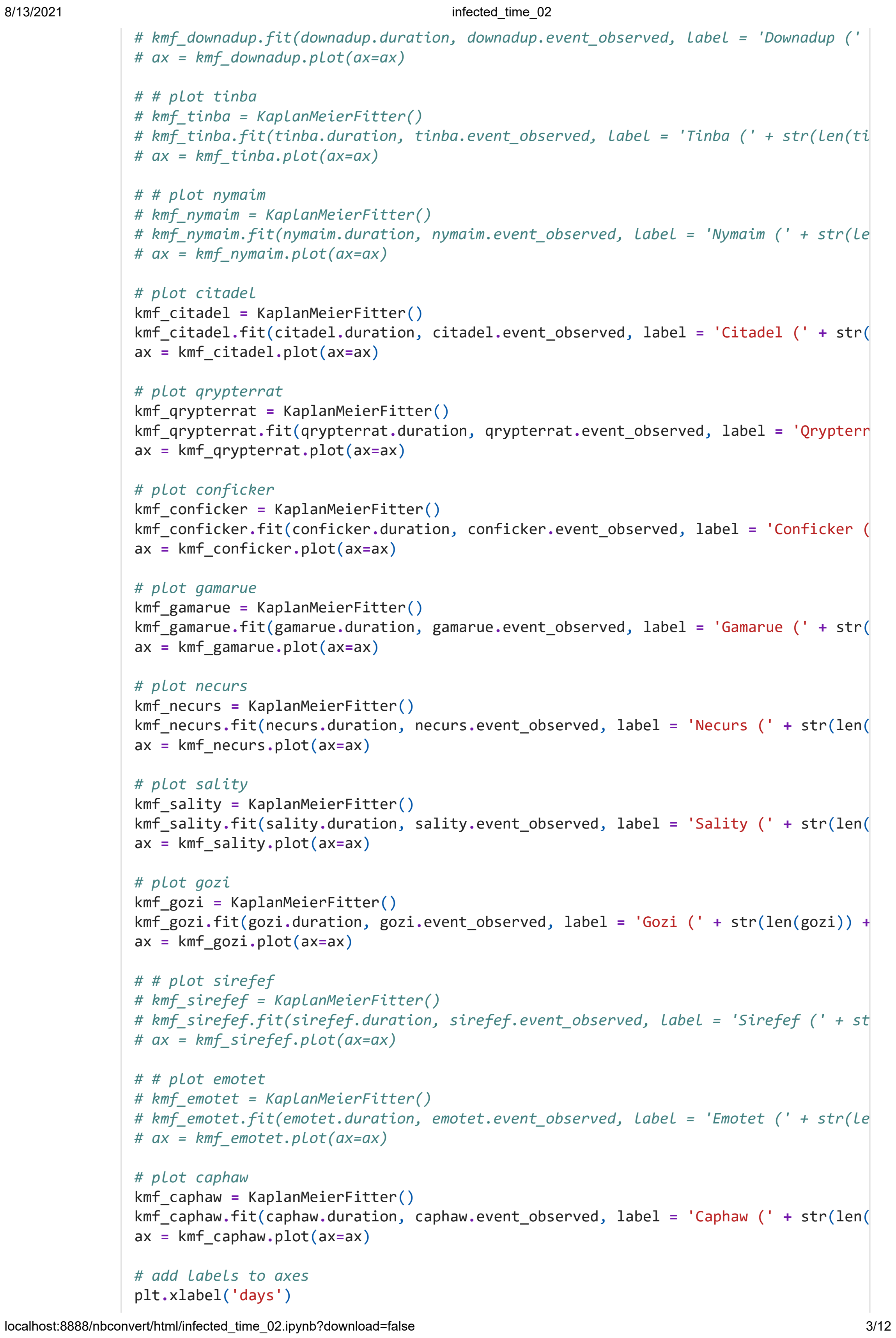
localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 1/12



8/13/2021 infected_time_02

```
# there are malware combinations in the dataset, in total 11 cases
malware_combinations = ['mirai|avalanche-generic|loader|zeus|foxbantrix-unknown|gozi
                        'mirai|sality|kronos|zeus|virut|plasma-tomas', 'mirai|sality
                        'mirai|tinba', 'mirai|tinba|http-scan', 'qsnatch|avalanche-g
                        'ramnit|android.hummer|avalanche-andromeda|stantinko|android
                        'vawtrak|luder|jadtre|minr|vpnfilter']
# drop malware combinations
df = df[~df.malware_types.isin(malware_combinations)]

# build different dataframes for each malware
qsnatch_df = df[df['malware_types'] == 'qsnatch']
vpnfilter_df = df[df['malware_types'] == 'vpnfilter']
mirai_df = df[df['malware_types'] == 'mirai']
avalanche = df[df['malware_types'] == 'avalanche']
andromeda = df[df['malware_types'] == 'avalanche-andromeda']
bladabindi = df[df['malware_types'] == 'bladabindi']
conficker = df[df['malware_types'] == 'conficker']
gamarue = df[df['malware_types'] == 'gamarue']
necurs = df[df['malware_types'] == 'necurs']
sality = df[df['malware_types'] == 'sality']
gozi = df[df['malware_types'] == 'gozi']
sirefef = df[df['malware_types'] == 'sirefef']
emotet = df[df['malware_types'] == 'emotet']
caphaw = df[df['malware_types'] == 'caphaw']
zeroaccess = df[df['malware_types'] == 'zeroaccess']
ramnit = df[df['malware_types'] == 'ramnit']
win = df[df['malware_types'] == 'win32/']
downadup = df[df['malware_types'] == 'downadup']
tinba = df[df['malware_types'] == 'tinba']
cutwail = df[df['malware_types'] == 'cutwail']
nymaim = df[df['malware_types'] == 'nymaim']
nivdort = df[df['malware_types'] == 'nivdort']
gamarue = df[df['malware_types'] == 'gamarue']
citadel = df[df['malware_types'] == 'citadel']
caphaw = df[df['malware_types'] == 'caphaw']
qrypterrat = df[df['malware_types'] == 'qrypter.rat']
pushdo = df[df['malware_types'] == 'pushdo']
emotet = df[df['malware_types'] == 'emotet']
zeus = df[df['malware_types'] == 'zeus']
kovter = df[df['malware_types'] == 'kovter']

plt.figure(figsize=(10,8))

# plot qsnatch
kmf_qsnatch = KaplanMeierFitter()
kmf_qsnatch.fit(qsnatch_df.duration, qsnatch_df.event_observed, label = 'QSnatch ('
ax = kmf_qsnatch.plot()

# plot mirai
kmf_mirai = KaplanMeierFitter()
kmf_mirai.fit(mirai_df.duration, mirai_df.event_observed, label = 'Mirai (' + str(le
ax = kmf_mirai.plot(ax=ax)

# plot ramnit
kmf_ramnit = KaplanMeierFitter()
kmf_ramnit.fit(ramnit.duration, ramnit.event_observed, label = 'Ramnit (' + str(len(
ax = kmf_ramnit.plot(ax=ax)

# plot kovter
kmf_kovter = KaplanMeierFitter()
kmf_kovter.fit(kovter.duration, kovter.event_observed, label = 'Kovter (' + str(len(
ax = kmf_kovter.plot(ax=ax)

# # plot downadup
# kmf_downadup = KaplanMeierFitter()
```

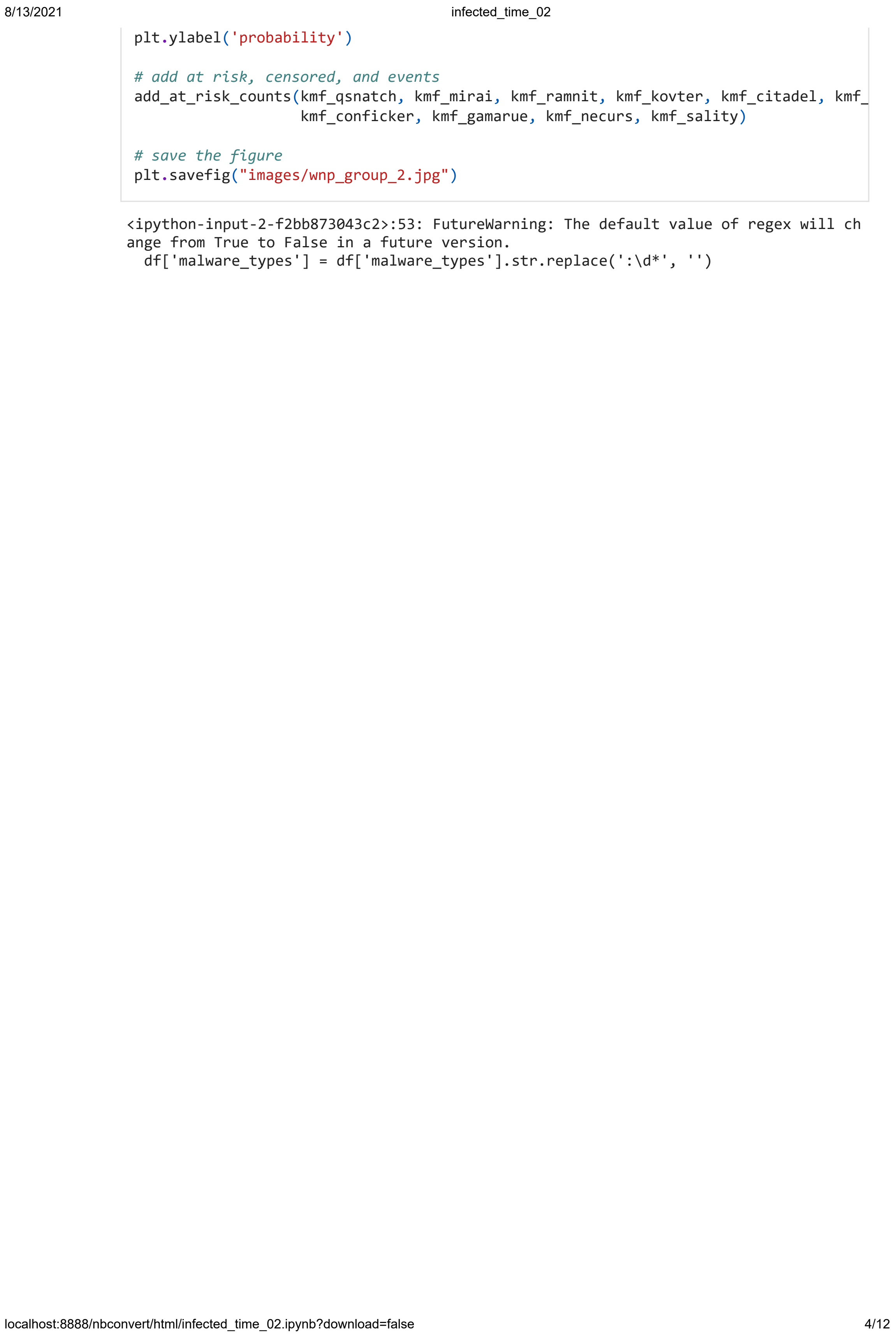
localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 2/12



8/13/2021 infected_time_02

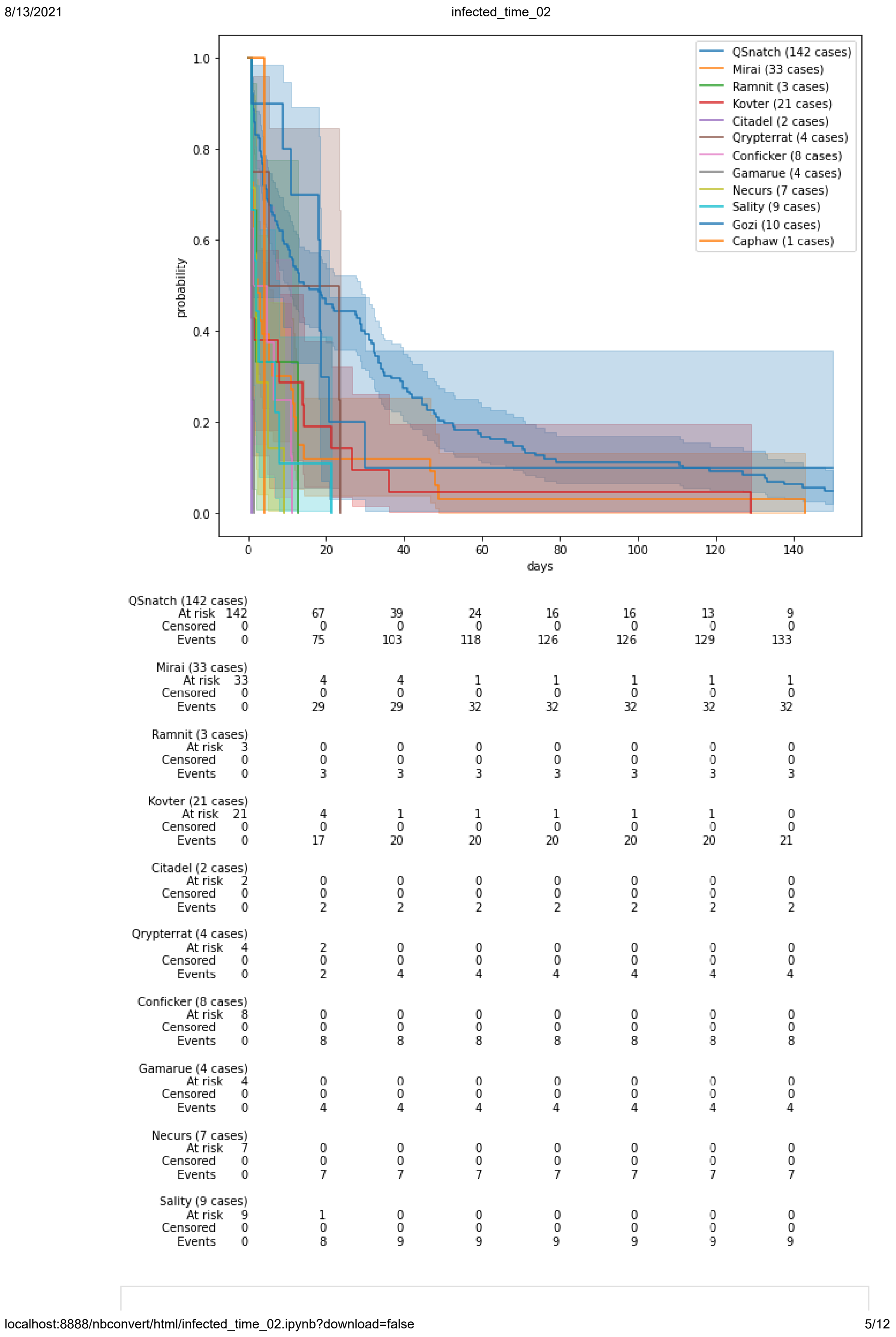
```
# kmf_downadup.fit(downadup.duration, downadup.event_observed, label = 'Downadup ('
# ax = kmf_downadup.plot(ax=ax)

# # plot tinba
# kmf_tinba = KaplanMeierFitter()
# kmf_tinba.fit(tinba.duration, tinba.event_observed, label = 'Tinba (' + str(len(ti
# ax = kmf_tinba.plot(ax=ax)

# # plot nymaim
# kmf_nymaim = KaplanMeierFitter()
# kmf_nymaim.fit(nymaim.duration, nymaim.event_observed, label = 'Nymaim (' + str(le
# ax = kmf_nymaim.plot(ax=ax)

# plot citadel
kmf_citadel = KaplanMeierFitter()
kmf_citadel.fit(citadel.duration, citadel.event_observed, label = 'Citadel (' + str(
ax = kmf_citadel.plot(ax=ax)

# plot qrypterrat
kmf_qrypterrat = KaplanMeierFitter()
kmf_qrypterrat.fit(qrypterrat.duration, qrypterrat.event_observed, label = 'Qrypterr
ax = kmf_qrypterrat.plot(ax=ax)

# plot conficker
kmf_conficker = KaplanMeierFitter()
kmf_conficker.fit(conficker.duration, conficker.event_observed, label = 'Conficker (
ax = kmf_conficker.plot(ax=ax)

# plot gamarue
kmf_gamarue = KaplanMeierFitter()
kmf_gamarue.fit(gamarue.duration, gamarue.event_observed, label = 'Gamarue (' + str(
ax = kmf_gamarue.plot(ax=ax)

# plot necurs
kmf_necurs = KaplanMeierFitter()
kmf_necurs.fit(necurs.duration, necurs.event_observed, label = 'Necurs (' + str(len(
ax = kmf_necurs.plot(ax=ax)

# plot sality
kmf_sality = KaplanMeierFitter()
kmf_sality.fit(sality.duration, sality.event_observed, label = 'Sality (' + str(len(
ax = kmf_sality.plot(ax=ax)

# plot gozi
kmf_gozi = KaplanMeierFitter()
kmf_gozi.fit(gozi.duration, gozi.event_observed, label = 'Gozi (' + str(len(gozi)) +
ax = kmf_gozi.plot(ax=ax)

# # plot sirefef
# kmf_sirefef = KaplanMeierFitter()
# kmf_sirefef.fit(sirefef.duration, sirefef.event_observed, label = 'Sirefef (' + st
# ax = kmf_sirefef.plot(ax=ax)

# # plot emotet
# kmf_emotet = KaplanMeierFitter()
# kmf_emotet.fit(emotet.duration, emotet.event_observed, label = 'Emotet (' + str(le
# ax = kmf_emotet.plot(ax=ax)

# plot caphaw
kmf_caphaw = KaplanMeierFitter()
kmf_caphaw.fit(caphaw.duration, caphaw.event_observed, label = 'Caphaw (' + str(len(
ax = kmf_caphaw.plot(ax=ax)

# add labels to axes
plt.xlabel('days')
```

localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 3/12



```
plt.ylabel('probability')

# add at risk, censored, and events
add_at_risk_counts(kmf_qsnatch, kmf_mirai, kmf_ramnit, kmf_kovter, kmf_citadel, kmf_
                   kmf_conficker, kmf_gamarue, kmf_necurs, kmf_sality)

# save the figure
plt.savefig("images/wnp_group_2.jpg")
```

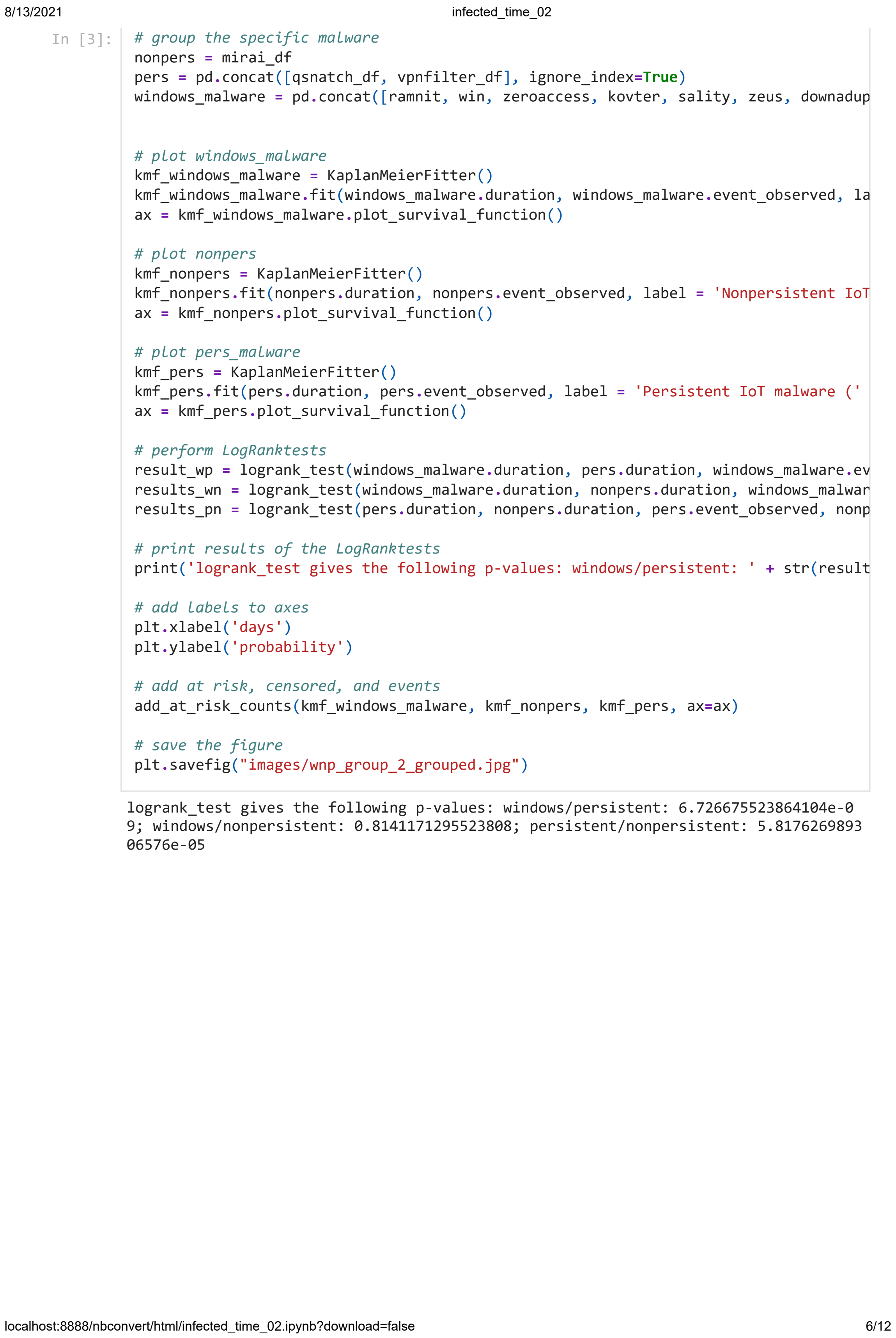
```
<ipython-input-2-f2bb873043c2>:53: FutureWarning: The default value of regex will ch
ange from True to False in a future version.
  df['malware_types'] = df['malware_types'].str.replace(':\d*', '')
```



8/13/2021 infected_time_02

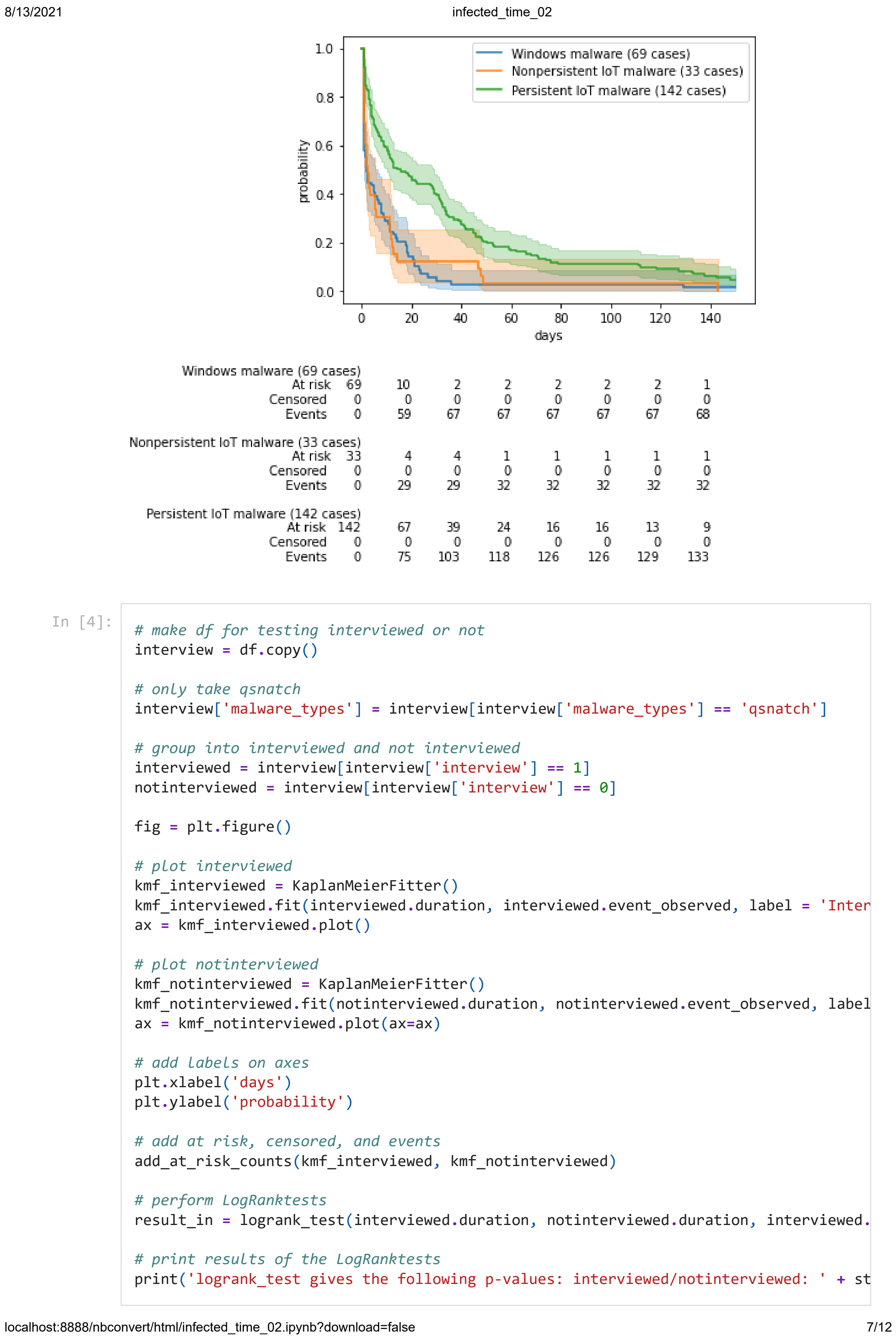
```
QSnatch (142 cases)
      At risk  142    67    39    24    16    16    13     9
     Censored    0     0     0     0     0     0     0     0
       Events    0    75   103   118   126   126   129   133

  Mirai (33 cases)
      At risk   33     4     4     1     1     1     1     1
     Censored    0     0     0     0     0     0     0     0
       Events    0    29    29    32    32    32    32    32

 Ramnit (3 cases)
      At risk    3     0     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     3     3     3     3     3     3     3

 Kovter (21 cases)
      At risk   21     4     1     1     1     1     1     0
     Censored    0     0     0     0     0     0     0     0
       Events    0    17    20    20    20    20    20    21

 Citadel (2 cases)
      At risk    2     0     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     2     2     2     2     2     2     2

Qrypterrat (4 cases)
      At risk    4     2     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     2     4     4     4     4     4     4

Conficker (8 cases)
      At risk    8     0     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     8     8     8     8     8     8     8

Gamarue (4 cases)
      At risk    4     0     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     4     4     4     4     4     4     4

 Necurs (7 cases)
      At risk    7     0     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     7     7     7     7     7     7     7

  Sality (9 cases)
      At risk    9     1     0     0     0     0     0     0
     Censored    0     0     0     0     0     0     0     0
       Events    0     8     9     9     9     9     9     9
```

localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 5/12



```
In [3]: # group the specific malware
        nonpers = mirai_df
        pers = pd.concat([qsnatch_df, vpnfilter_df], ignore_index=True)
        windows_malware = pd.concat([ramnit, win, zeroaccess, kovter, sality, zeus, downadup


        # plot windows_malware
        kmf_windows_malware = KaplanMeierFitter()
        kmf_windows_malware.fit(windows_malware.duration, windows_malware.event_observed, la
        ax = kmf_windows_malware.plot_survival_function()

        # plot nonpers
        kmf_nonpers = KaplanMeierFitter()
        kmf_nonpers.fit(nonpers.duration, nonpers.event_observed, label = 'Nonpersistent IoT
        ax = kmf_nonpers.plot_survival_function()

        # plot pers_malware
        kmf_pers = KaplanMeierFitter()
        kmf_pers.fit(pers.duration, pers.event_observed, label = 'Persistent IoT malware ('
        ax = kmf_pers.plot_survival_function()

        # perform LogRanktests
        result_wp = logrank_test(windows_malware.duration, pers.duration, windows_malware.ev
        results_wn = logrank_test(windows_malware.duration, nonpers.duration, windows_malwar
        results_pn = logrank_test(pers.duration, nonpers.duration, pers.event_observed, nonp

        # print results of the LogRanktests
        print('logrank_test gives the following p-values: windows/persistent: ' + str(result

        # add labels to axes
        plt.xlabel('days')
        plt.ylabel('probability')

        # add at risk, censored, and events
        add_at_risk_counts(kmf_windows_malware, kmf_nonpers, kmf_pers, ax=ax)

        # save the figure
        plt.savefig("images/wnp_group_2_grouped.jpg")
```

```
logrank_test gives the following p-values: windows/persistent: 6.726675523864104e-0
9; windows/nonpersistent: 0.8141171295523808; persistent/nonpersistent: 5.8176269893
06576e-05
```


probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
days
Windows malware (69 cases)
Nonpersistent IoT malware (33 cases)
Persistent IoT malware (142 cases)

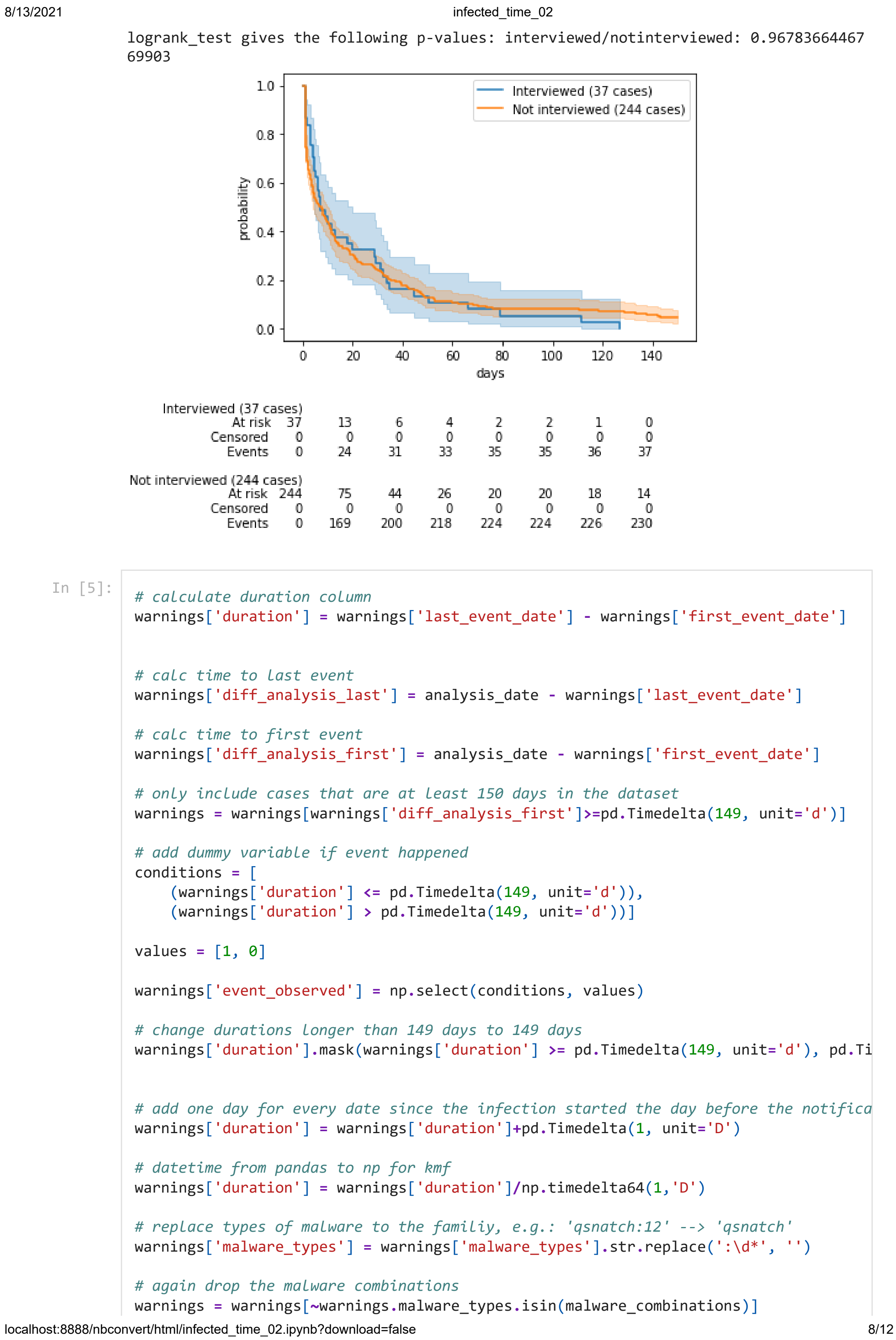
8/13/2021 infected_time_02

```
             Windows malware (69 cases)
                        At risk   69    10     2     2     2     2     2     1
                       Censored    0     0     0     0     0     0     0     0
                         Events    0    59    67    67    67    67    67    68

   Nonpersistent IoT malware (33 cases)
                        At risk   33     4     4     1     1     1     1     1
                       Censored    0     0     0     0     0     0     0     0
                         Events    0    29    29    32    32    32    32    32

      Persistent IoT malware (142 cases)
                        At risk  142    67    39    24    16    16    13     9
                       Censored    0     0     0     0     0     0     0     0
                         Events    0    75   103   118   126   126   129   133
```

In [4]:

```
# make df for testing interviewed or not
interview = df.copy()

# only take qsnatch
interview['malware_types'] = interview[interview['malware_types'] == 'qsnatch']

# group into interviewed and not interviewed
interviewed = interview[interview['interview'] == 1]
notinterviewed = interview[interview['interview'] == 0]

fig = plt.figure()

# plot interviewed
kmf_interviewed = KaplanMeierFitter()
kmf_interviewed.fit(interviewed.duration, interviewed.event_observed, label = 'Inter
ax = kmf_interviewed.plot()

# plot notinterviewed
kmf_notinterviewed = KaplanMeierFitter()
kmf_notinterviewed.fit(notinterviewed.duration, notinterviewed.event_observed, label
ax = kmf_notinterviewed.plot(ax=ax)

# add labels on axes
plt.xlabel('days')
plt.ylabel('probability')

# add at risk, censored, and events
add_at_risk_counts(kmf_interviewed, kmf_notinterviewed)

# perform LogRanktests
result_in = logrank_test(interviewed.duration, notinterviewed.duration, interviewed.

# print results of the LogRanktests
print('logrank_test gives the following p-values: interviewed/notinterviewed: ' + st
```

localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 7/12

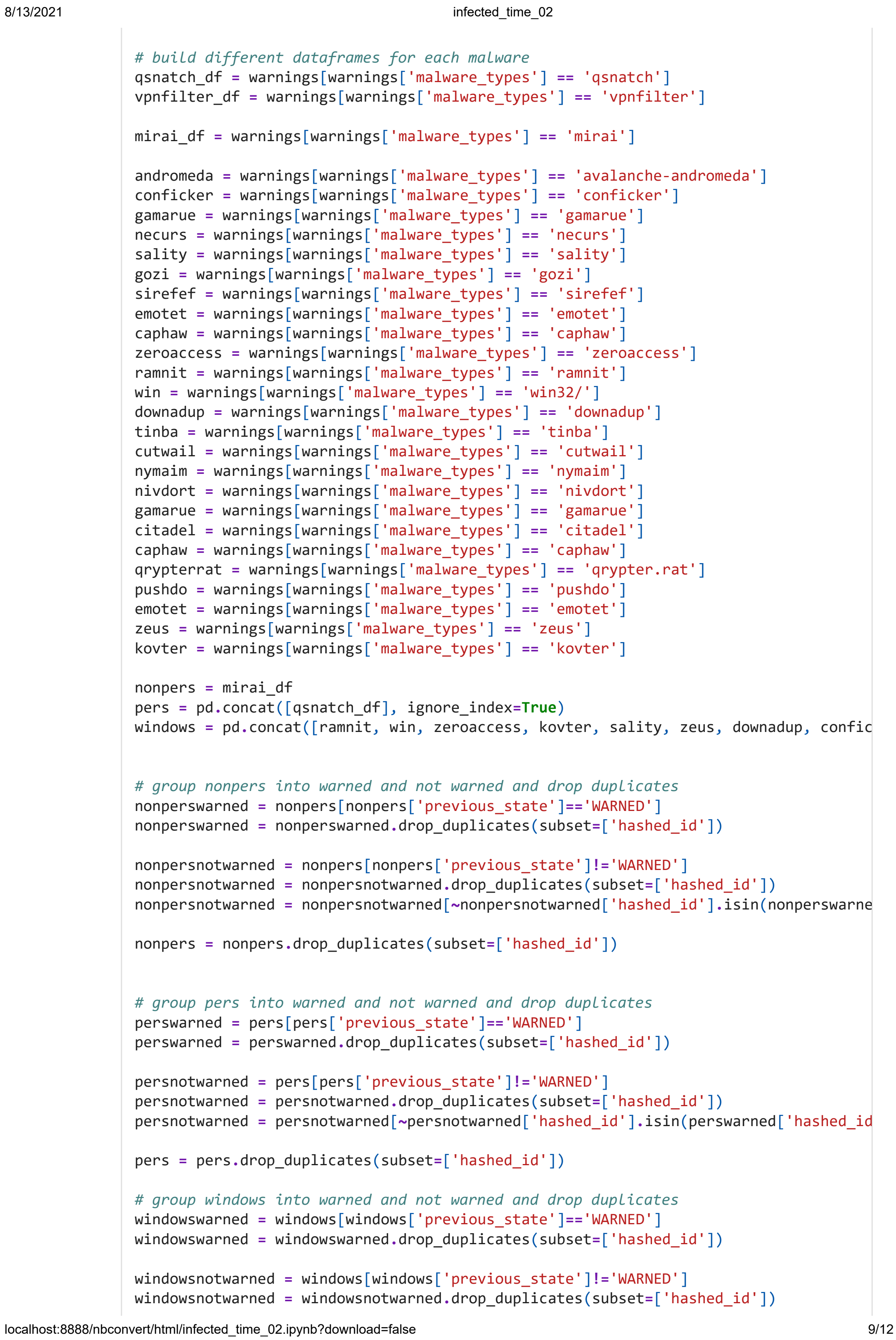


```
logrank_test gives the following p-values: interviewed/notinterviewed: 0.96783664467
69903
```

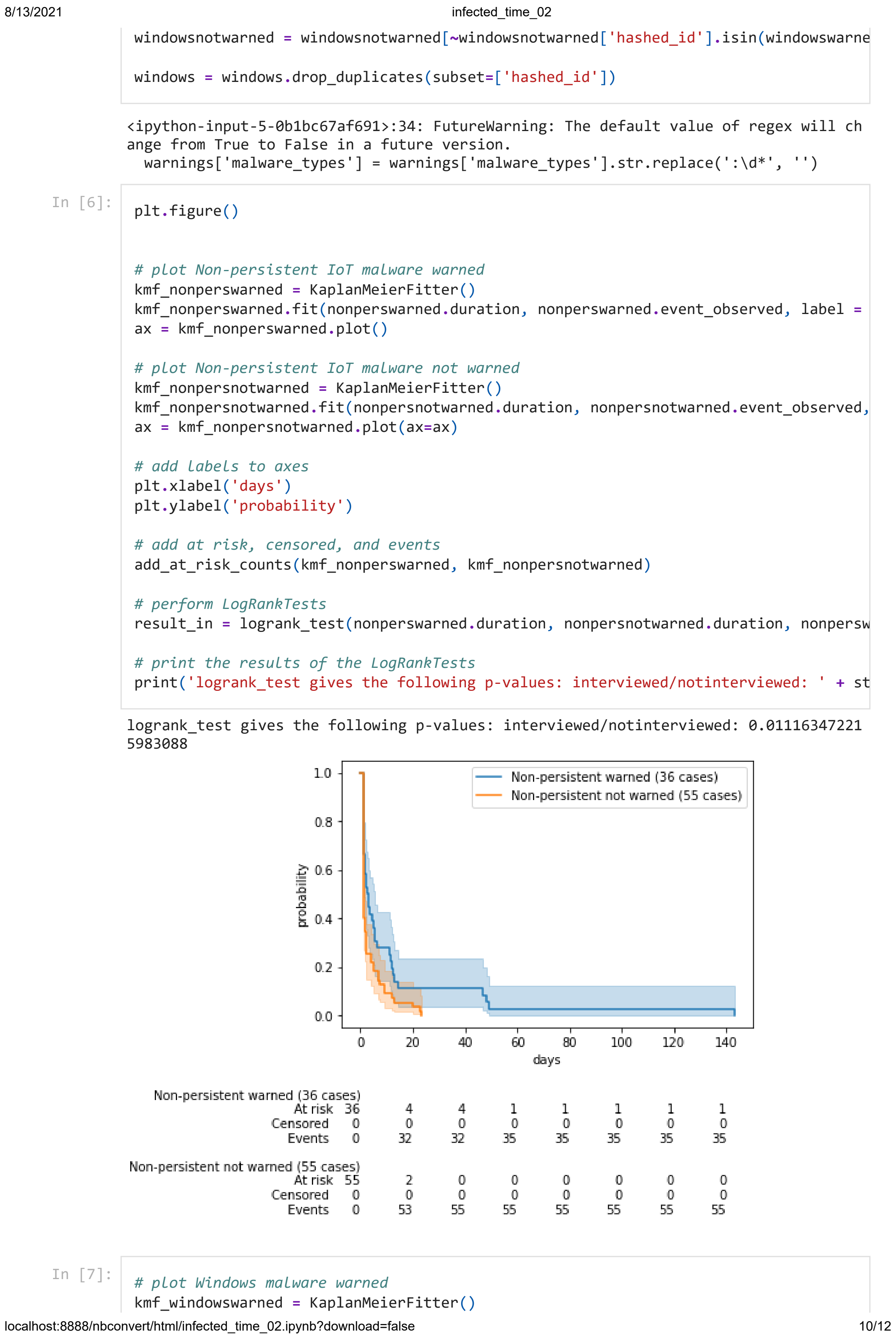
```
Interviewed (37 cases)
           At risk   37       13      6      4      2      2      1      0
          Censored    0        0      0      0      0      0      0      0
            Events    0       24     31     33     35     35     36     37

Not interviewed (244 cases)
           At risk  244       75     44     26     20     20     18     14
          Censored    0        0      0      0      0      0      0      0
            Events    0      169    200    218    224    224    226    230
```

In [5]:

```python
# calculate duration column
warnings['duration'] = warnings['last_event_date'] - warnings['first_event_date']


# calc time to last event
warnings['diff_analysis_last'] = analysis_date - warnings['last_event_date']

# calc time to first event
warnings['diff_analysis_first'] = analysis_date - warnings['first_event_date']

# only include cases that are at least 150 days in the dataset
warnings = warnings[warnings['diff_analysis_first']>=pd.Timedelta(149, unit='d')]

# add dummy variable if event happened
conditions = [
    (warnings['duration'] <= pd.Timedelta(149, unit='d')),
    (warnings['duration'] > pd.Timedelta(149, unit='d'))]

values = [1, 0]

warnings['event_observed'] = np.select(conditions, values)

# change durations longer than 149 days to 149 days
warnings['duration'].mask(warnings['duration'] >= pd.Timedelta(149, unit='d'), pd.Ti


# add one day for every date since the infection started the day before the notifica
warnings['duration'] = warnings['duration']+pd.Timedelta(1, unit='D')

# datetime from pandas to np for kmf
warnings['duration'] = warnings['duration']/np.timedelta64(1,'D')

# replace types of malware to the familiy, e.g.: 'qsnatch:12' --> 'qsnatch'
warnings['malware_types'] = warnings['malware_types'].str.replace(':\d*', '')

# again drop the malware combinations
warnings = warnings[~warnings.malware_types.isin(malware_combinations)]
```



```
# build different dataframes for each malware
qsnatch_df = warnings[warnings['malware_types'] == 'qsnatch']
vpnfilter_df = warnings[warnings['malware_types'] == 'vpnfilter']

mirai_df = warnings[warnings['malware_types'] == 'mirai']

andromeda = warnings[warnings['malware_types'] == 'avalanche-andromeda']
conficker = warnings[warnings['malware_types'] == 'conficker']
gamarue = warnings[warnings['malware_types'] == 'gamarue']
necurs = warnings[warnings['malware_types'] == 'necurs']
sality = warnings[warnings['malware_types'] == 'sality']
gozi = warnings[warnings['malware_types'] == 'gozi']
sirefef = warnings[warnings['malware_types'] == 'sirefef']
emotet = warnings[warnings['malware_types'] == 'emotet']
caphaw = warnings[warnings['malware_types'] == 'caphaw']
zeroaccess = warnings[warnings['malware_types'] == 'zeroaccess']
ramnit = warnings[warnings['malware_types'] == 'ramnit']
win = warnings[warnings['malware_types'] == 'win32/']
downadup = warnings[warnings['malware_types'] == 'downadup']
tinba = warnings[warnings['malware_types'] == 'tinba']
cutwail = warnings[warnings['malware_types'] == 'cutwail']
nymaim = warnings[warnings['malware_types'] == 'nymaim']
nivdort = warnings[warnings['malware_types'] == 'nivdort']
gamarue = warnings[warnings['malware_types'] == 'gamarue']
citadel = warnings[warnings['malware_types'] == 'citadel']
caphaw = warnings[warnings['malware_types'] == 'caphaw']
qrypterrat = warnings[warnings['malware_types'] == 'qrypter.rat']
pushdo = warnings[warnings['malware_types'] == 'pushdo']
emotet = warnings[warnings['malware_types'] == 'emotet']
zeus = warnings[warnings['malware_types'] == 'zeus']
kovter = warnings[warnings['malware_types'] == 'kovter']

nonpers = mirai_df
pers = pd.concat([qsnatch_df], ignore_index=True)
windows = pd.concat([ramnit, win, zeroaccess, kovter, sality, zeus, downadup, confic


# group nonpers into warned and not warned and drop duplicates
nonperswarned = nonpers[nonpers['previous_state']=='WARNED']
nonperswarned = nonperswarned.drop_duplicates(subset=['hashed_id'])

nonpersnotwarned = nonpers[nonpers['previous_state']!='WARNED']
nonpersnotwarned = nonpersnotwarned.drop_duplicates(subset=['hashed_id'])
nonpersnotwarned = nonpersnotwarned[~nonpersnotwarned['hashed_id'].isin(nonperswarne

nonpers = nonpers.drop_duplicates(subset=['hashed_id'])


# group pers into warned and not warned and drop duplicates
perswarned = pers[pers['previous_state']=='WARNED']
perswarned = perswarned.drop_duplicates(subset=['hashed_id'])

persnotwarned = pers[pers['previous_state']!='WARNED']
persnotwarned = persnotwarned.drop_duplicates(subset=['hashed_id'])
persnotwarned = persnotwarned[~persnotwarned['hashed_id'].isin(perswarned['hashed_id

pers = pers.drop_duplicates(subset=['hashed_id'])

# group windows into warned and not warned and drop duplicates
windowswarned = windows[windows['previous_state']=='WARNED']
windowswarned = windowswarned.drop_duplicates(subset=['hashed_id'])

windowsnotwarned = windows[windows['previous_state']!='WARNED']
windowsnotwarned = windowsnotwarned.drop_duplicates(subset=['hashed_id'])
```



8/13/2021 infected_time_02

```
windowsnotwarned = windowsnotwarned[~windowsnotwarned['hashed_id'].isin(windowswarne

windows = windows.drop_duplicates(subset=['hashed_id'])
```

```
<ipython-input-5-0b1bc67af691>:34: FutureWarning: The default value of regex will ch
ange from True to False in a future version.
  warnings['malware_types'] = warnings['malware_types'].str.replace(':\d*', '')
```

In [6]:

```
plt.figure()


# plot Non-persistent IoT malware warned
kmf_nonperswarned = KaplanMeierFitter()
kmf_nonperswarned.fit(nonperswarned.duration, nonperswarned.event_observed, label =
ax = kmf_nonperswarned.plot()

# plot Non-persistent IoT malware not warned
kmf_nonpersnotwarned = KaplanMeierFitter()
kmf_nonpersnotwarned.fit(nonpersnotwarned.duration, nonpersnotwarned.event_observed,
ax = kmf_nonpersnotwarned.plot(ax=ax)

# add labels to axes
plt.xlabel('days')
plt.ylabel('probability')

# add at risk, censored, and events
add_at_risk_counts(kmf_nonperswarned, kmf_nonpersnotwarned)

# perform LogRankTests
result_in = logrank_test(nonperswarned.duration, nonpersnotwarned.duration, nonpersw

# print the results of the LogRankTests
print('logrank_test gives the following p-values: interviewed/notinterviewed: ' + st
```

```
logrank_test gives the following p-values: interviewed/notinterviewed: 0.01116347221
5983088
```

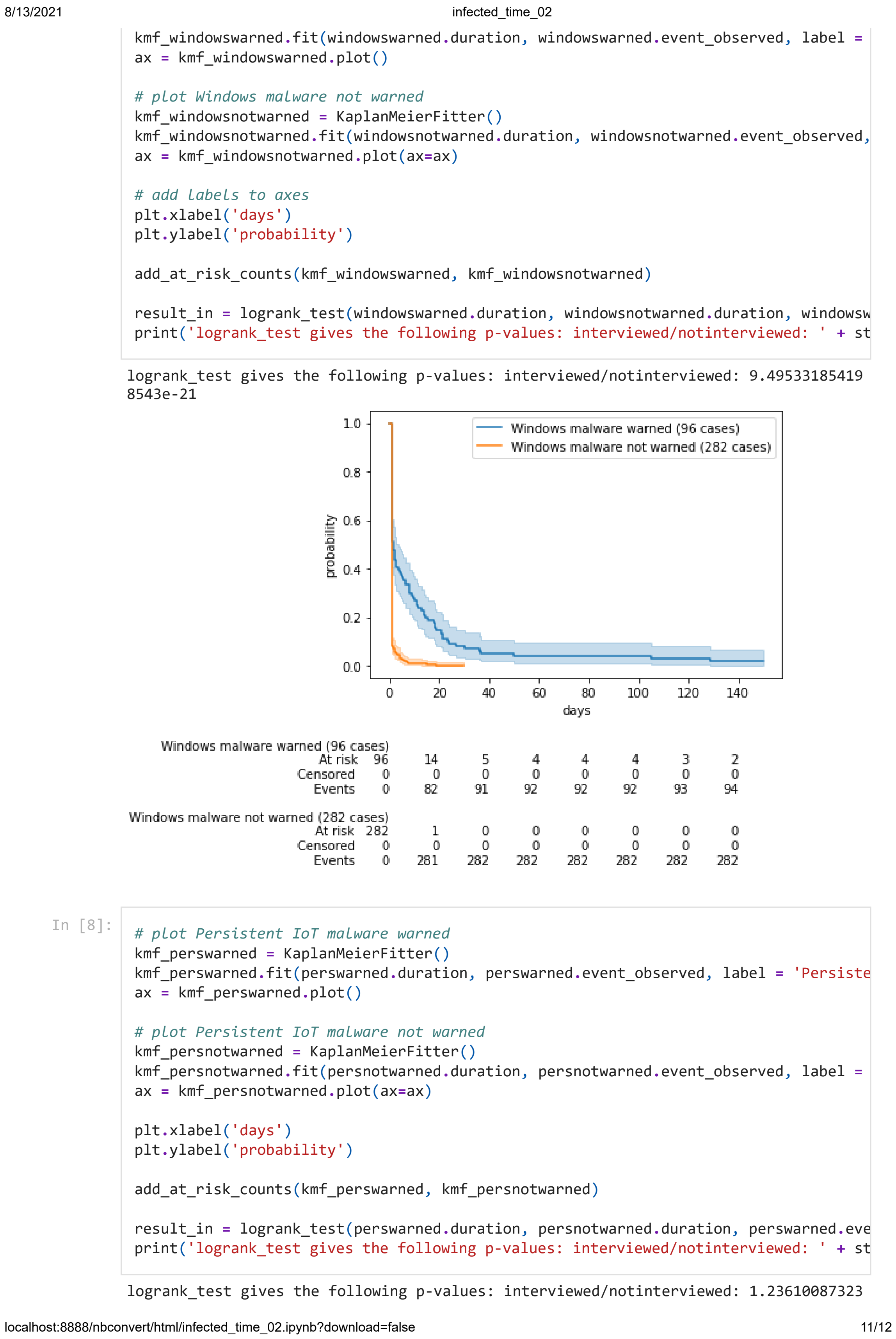
Non-persistent warned (36 cases)
Non-persistent not warned (55 cases)
probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
days

```
Non-persistent warned (36 cases)
                         At risk  36     4     4     1     1     1     1     1
                        Censored   0     0     0     0     0     0     0     0
                          Events   0    32    32    35    35    35    35    35

Non-persistent not warned (55 cases)
                         At risk  55     2     0     0     0     0     0     0
                        Censored   0     0     0     0     0     0     0     0
                          Events   0    53    55    55    55    55    55    55
```

In [7]:

```
# plot Windows malware warned
kmf_windowswarned = KaplanMeierFitter()
```

localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 10/12



8/13/2021 infected_time_02

```
kmf_windowswarned.fit(windowswarned.duration, windowswarned.event_observed, label =
ax = kmf_windowswarned.plot()

# plot Windows malware not warned
kmf_windowsnotwarned = KaplanMeierFitter()
kmf_windowsnotwarned.fit(windowsnotwarned.duration, windowsnotwarned.event_observed,
ax = kmf_windowsnotwarned.plot(ax=ax)

# add labels to axes
plt.xlabel('days')
plt.ylabel('probability')

add_at_risk_counts(kmf_windowswarned, kmf_windowsnotwarned)

result_in = logrank_test(windowswarned.duration, windowsnotwarned.duration, windowsw
print('logrank_test gives the following p-values: interviewed/notinterviewed: ' + st
```

```
logrank_test gives the following p-values: interviewed/notinterviewed: 9.49533185419
8543e-21
```

Windows malware warned (96 cases)
Windows malware not warned (282 cases)
probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
days

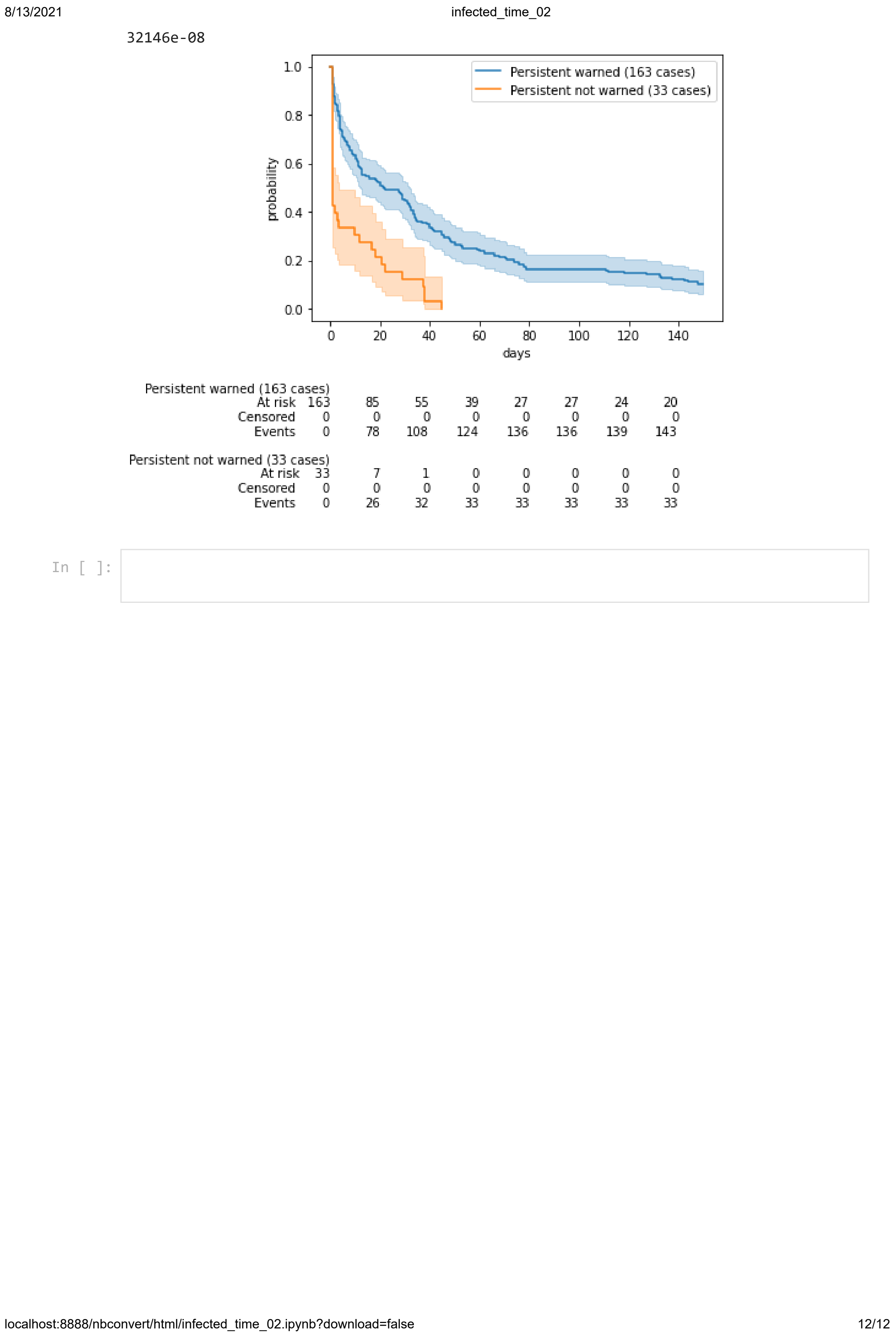
```
         Windows malware warned (96 cases)
                          At risk   96    14     5     4     4     4     3     2
                        Censored     0     0     0     0     0     0     0     0
                          Events     0    82    91    92    92    92    93    94

     Windows malware not warned (282 cases)
                          At risk  282     1     0     0     0     0     0     0
                        Censored     0     0     0     0     0     0     0     0
                          Events     0   281   282   282   282   282   282   282
```

In [8]:

```
# plot Persistent IoT malware warned
kmf_perswarned = KaplanMeierFitter()
kmf_perswarned.fit(perswarned.duration, perswarned.event_observed, label = 'Persiste
ax = kmf_perswarned.plot()

# plot Persistent IoT malware not warned
kmf_persnotwarned = KaplanMeierFitter()
kmf_persnotwarned.fit(persnotwarned.duration, persnotwarned.event_observed, label =
ax = kmf_persnotwarned.plot(ax=ax)

plt.xlabel('days')
plt.ylabel('probability')

add_at_risk_counts(kmf_perswarned, kmf_persnotwarned)

result_in = logrank_test(perswarned.duration, persnotwarned.duration, perswarned.eve
print('logrank_test gives the following p-values: interviewed/notinterviewed: ' + st
```

```
logrank_test gives the following p-values: interviewed/notinterviewed: 1.23610087323
```

localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 11/12



8/13/2021 infected_time_02

```
32146e-08
```

Persistent warned (163 cases)
Persistent not warned (33 cases)
probability
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
days

```
Persistent warned (163 cases)
             At risk  163    85    55    39    27    27    24    20
            Censored    0     0     0     0     0     0     0     0
              Events    0    78   108   124   136   136   139   143

Persistent not warned (33 cases)
             At risk   33     7     1     0     0     0     0     0
            Censored    0     0     0     0     0     0     0     0
              Events    0    26    32    33    33    33    33    33
```

```
In [ ]:
```

localhost:8888/nbconvert/html/infected_time_02.ipynb?download=false 12/12

Appendix I Interview Questions

Components	Subcomponents	Definition	Questions
Capability	Physical capability	The capability to engage in essential physical processes	Did you have any physical or bodily limitations that made the steps challenging?
	Psychological capability	The capacity to engage in the necessary thought processes - comprehension and reasoning	Did you understand the steps?
			Did you find the steps challenging?
			Did you know what malware is?
			Did you know the difference between persistent and non-persistent malware?
			Did you think you could perform the steps?
			Did you have previous experience and skills with IT?
Motivation	Reflective processes	The evaluations and plans	Do you find the steps useful?
			What do you think would happen if someone does not follow the steps?
			Did you think you are responsible for performing the steps?
	Automatic processes	The emotions and impulses that arise from associative learning and/or innate dispositions	What did you feel during the steps?
			Did an impulse help you perform the steps?
Opportunity	Physical opportunity	The opportunity afforded by the environment	Did you receive the notification email?
			Did you use any tools to perform the steps?
			Did you have enough time to perform the steps?
			Was the location of the device or any of the tools you used an issue to access it?
	Social opportunity	The cultural environment that directs the way an individual thinks	Have any people helped you perform the steps?
			Do some people you know have a strong opinion on performing the steps?

Table 49: Components, Subcomponents, and Interview questions for customers that followed the steps

	Did perform the steps	**Did not/partially perform the steps**
Physical capability	Did you have any physical or bodily limitations that made the steps challenging?	Did you have any physical or bodily limitations that prevented you from finishing the steps?
Psychological capability	Did you understand the steps?	Did you understand the steps?
	Did you find the steps challenging?	Did you find the steps challenging?
	Did you know what malware is?	Did you know what malware is?
	Did you know the difference between persistent and non-persistent malware?	Did you know the difference between persistent and non-persistent malware?
	Did you think you could perform the steps?	Did you think you could perform the steps?
	Did you have previous experience with IT systems?	Did you have previous experience with IT systems?
Reflective processes	Do you find the steps useful?	Do you find the steps useful?
	What do you think would happen if someone does not follow the steps?	What do you think would happen if someone does not follow the steps?
	Did you think you are responsible for performing the steps?	Did you think you are responsible for performing the steps?
Automatic processes	What did you feel while you performed the steps?	What did you feel when you received the notification email?
	Did an impulse helped you perform the steps?	Did an impulse prevent you from performing the steps?
Physical opportunity	Did you receive the notification email?	Did you receive the notification email?
	Did you use any tools to perform the steps?	Did you lacked any tools to perform all of the steps?
	Did you have enough time to perform the steps?	Did you not have enough time to perform the steps?
	Was the location of the device or any of the tools you used an issue to access it?	Was the location of the device or any of the tools you used an issue to access it?
Social opportunity	Have any people helped you perform the steps?	Did any people tried to help you perform the steps?
	Do some people you know have a strong opinion on performing the steps?	Do some people you know have a strong opinion on performing the steps?

Table 50: Questions for the different performances on the steps

Appendix J Flow charts of the interviews

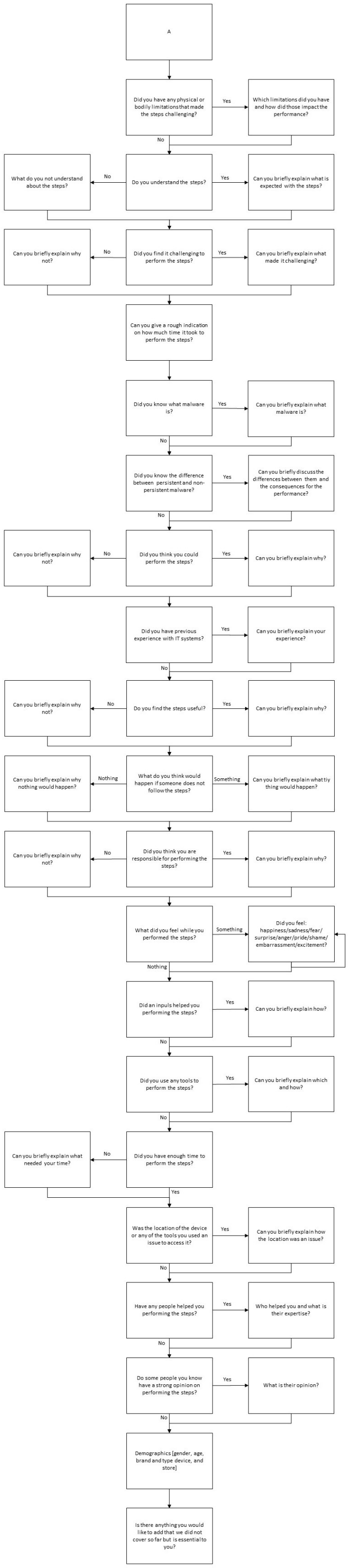

A
Did you have any physical or bodily limitations that made the steps challenging?
Yes
Which limitations did you have and how did those impact the performance?
No
What do you not understand about the steps?
No
Do you understand the steps?
Yes
Can you briefly explain what is expected with the steps?
Can you briefly explain why not?
No
Did you find it challenging to perform the steps?
Yes
Can you briefly explain what made it challenging?
Can you give a rough indication on how much time it took to perform the steps?
Did you know what malware is?
Yes
Can you briefly explain what malware is?
No
Did you know the difference between persistent and non-persistent malware?
Yes
Can you briefly discuss the differences between them and the consequences for the performance?
No
Can you briefly explain why not?
No
Did you think you could perform the steps?
Yes
Can you briefly explain why?
Did you have previous experience with IT systems?
Yes
Can you briefly explain your experience?
No
Can you briefly explain why not?
No
Do you find the steps useful?
Yes
Can you briefly explain why?
Can you briefly explain why nothing would happen?
Nothing
What do you think would happen if someone does not follow the steps?
Something
Can you briefly explain what tiy thing would happen?
Can you briefly explain why not?
No
Did you think you are responsible for performing the steps?
Yes
Can you briefly explain why?
What did you feel while you performed the steps?
Something
Did you feel: happiness/sadness/fear/ surprise/anger/pride/shame/ embarrassment/excitement?
Nothing
Did an inpuls helped you performing the steps?
Yes
Can you briefly explain how?
No
Did you use any tools to perform the steps?
Yes
Can you briefly explain which and how?
No
Can you briefly explain what needed your time?
No
Did you have enough time to perform the steps?
Yes
Was the location of the device or any of the tools you used an issue to access it?
Yes
Can you briefly explain how the location was an issue?
No
Have any people helped you performing the steps?
Yes
Who helped you and what is their expertise?
No
Do some people you know have a strong opinion on performing the steps?
Yes
What is their opinion?
No
Demographics [gender, age, brand and type device, and store]
Is there anything you would like to add that we did not cover so far but is essential to you?

Figure 20: The interview questions and flow for a customer that followed all the steps

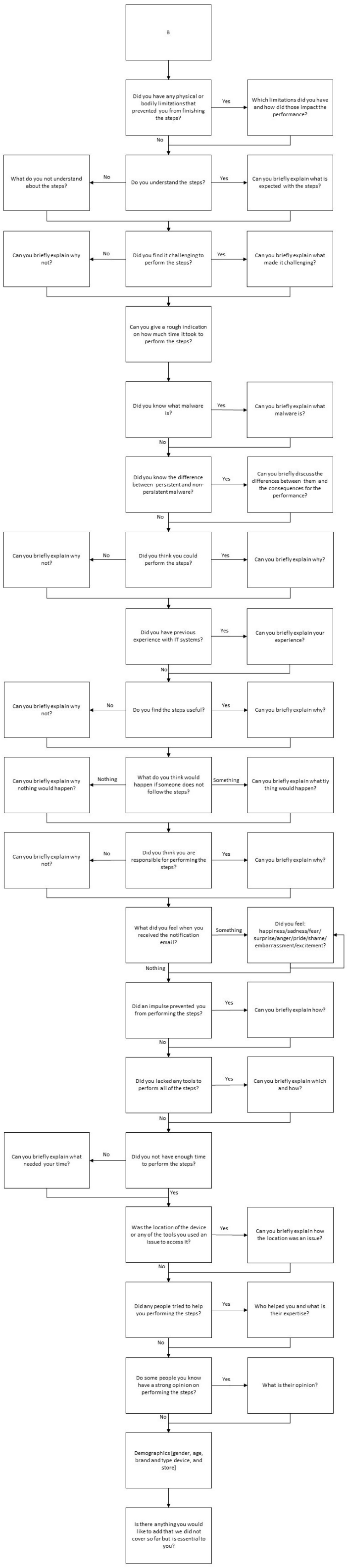
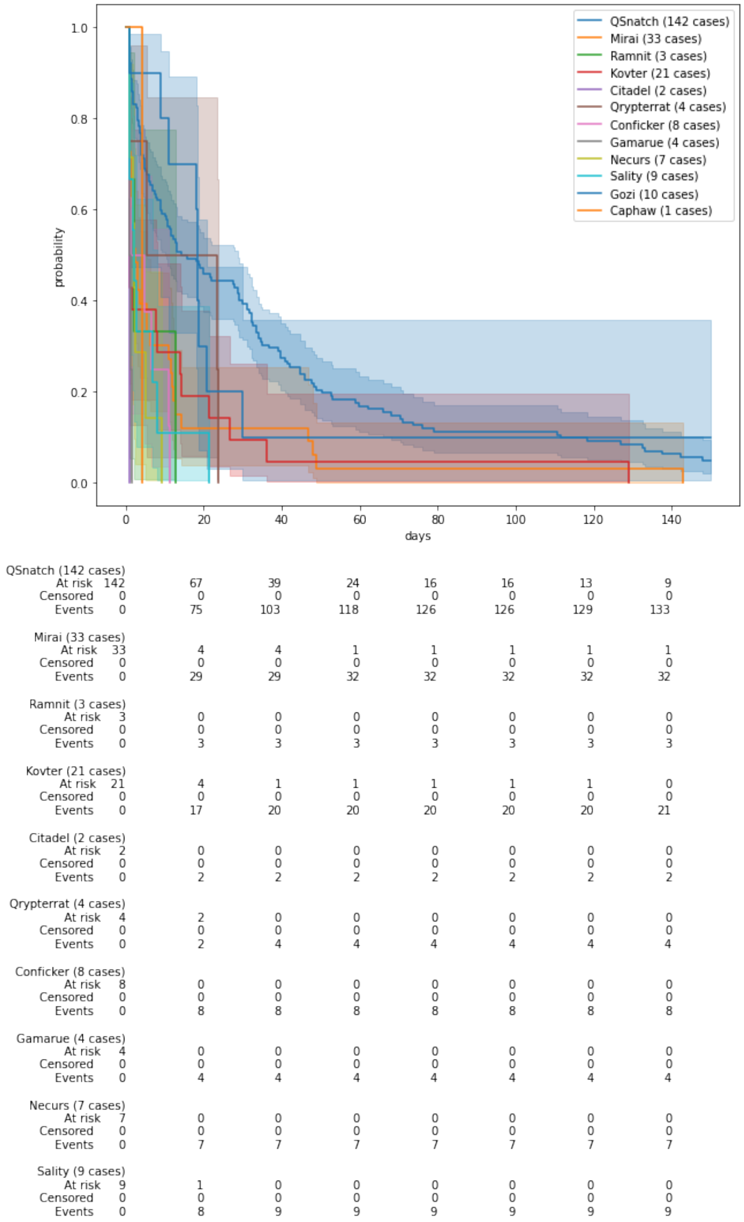

B
Did you have any physical or bodily limitations that prevented you from finishing the steps?
Yes
Which limitations did you have and how did those impact the performance?
No
What do you not understand about the steps?
No
Do you understand the steps?
Yes
Can you briefly explain what is expected with the steps?
Can you briefly explain why not?
No
Did you find it challenging to perform the steps?
Yes
Can you briefly explain what made it challenging?
Can you give a rough indication on how much time it took to perform the steps?
Did you know what malware is?
Yes
Can you briefly explain what malware is?
No
Did you know the difference between persistent and non-persistent malware?
Yes
Can you briefly discuss the differences between them and the consequences for the performance?
No
Can you briefly explain why not?
No
Did you think you could perform the steps?
Yes
Can you briefly explain why?
Did you have previous experience with IT systems?
Yes
Can you briefly explain your experience?
No
Can you briefly explain why not?
No
Do you find the steps useful?
Yes
Can you briefly explain why?
Can you briefly explain why nothing would happen?
Nothing
What do you think would happen if someone does not follow the steps?
Something
Can you briefly explain what tiy thing would happen?
Can you briefly explain why not?
No
Did you think you are responsible for performing the steps?
Yes
Can you briefly explain why?
What did you feel when you received the notification email?
Something
Did you feel: happiness/sadness/fear/ surprise/anger/pride/shame/ embarrassment/excitement?
Nothing
Did an impulse prevented you from performing the steps?
Yes
Can you briefly explain how?
No
Did you lacked any tools to perform all of the steps?
Yes
Can you briefly explain which and how?
No
Can you briefly explain what needed your time?
No
Did you not have enough time to perform the steps?
Yes
Was the location of the device or any of the tools you used an issue to access it?
Yes
Can you briefly explain how the location was an issue?
No
Did any people tried to help you performing the steps?
Yes
Who helped you and what is their expertise?
No
Do some people you know have a strong opinion on performing the steps?
Yes
What is their opinion?
No
Demographics [gender, age, brand and type device, and store]
Is there anything you would like to add that we did not cover so far but is essential to you?

Figure 21: The interview questions and flow for a customer that followed none or some of the steps

Appendix K Kaplan-Meier curves and risk, censored, and event counts



```
QSnatch (142 cases)
         At risk  142     67     39     24     16     16     13      9
        Censored    0      0      0      0      0      0      0      0
          Events    0     75    103    118    126    126    129    133

    Mirai (33 cases)
         At risk   33      4      4      1      1      1      1      1
        Censored    0      0      0      0      0      0      0      0
          Events    0     29     29     32     32     32     32     32

   Ramnit (3 cases)
         At risk    3      0      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      3      3      3      3      3      3      3

  Kovter (21 cases)
         At risk   21      4      1      1      1      1      1      0
        Censored    0      0      0      0      0      0      0      0
          Events    0     17     20     20     20     20     20     21

   Citadel (2 cases)
         At risk    2      0      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      2      2      2      2      2      2      2

Qrypterrat (4 cases)
         At risk    4      2      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      2      4      4      4      4      4      4

 Conficker (8 cases)
         At risk    8      0      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      8      8      8      8      8      8      8

  Gamarue (4 cases)
         At risk    4      0      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      4      4      4      4      4      4      4

   Necurs (7 cases)
         At risk    7      0      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      7      7      7      7      7      7      7

    Sality (9 cases)
         At risk    9      1      0      0      0      0      0      0
        Censored    0      0      0      0      0      0      0      0
          Events    0      8      9      9      9      9      9      9
```

Figure 22: K-M curves: cases at risk, censored, and events counts

Appendix L Results of the respondents

[h]

Table 51: Dates of invitation, interviewing, notifying, and detection traffic of the respondents

r.	*invited*	*interviewed*	*warned*	*detected*

1	22-04-2021	23-04-2021	23-10-20, 05-11-20	"20 October 2020 21 October 2020 22 October 2020 23 October 2020 24 October 2020 01 November 2020 02 November 2020 04 November 2020 05 November 2020 06 November 2020 07 November 2020 09 November 2020 10 November 2020 13 November 2020 15 November 2020 17 November 2020 18 November 2020 23 November 2020 27 November 2020 28 November 2020 30 November 2020 01 December 2020 02 December 2020 03 December 2020 04 December 2020 05 December 2020 06 December 2020 07 December 2020 09 December 2020 12 December 2020 18 December 2020 19 December 2020 22 December 2020 23 December 2020 31 December 2020 06 January 2021 08 January 2021 09 January 2021 10 January 2021 11 January 2021 12 January 2021 14 January 2021 15 January 2021 17 January 2021 20 January 2021 21 January 2021 23 January 2021 24 January 2021 25 January 2021 28 January 2021 29 January 2021 31 January 2021 01 February 2021 09 February 2021 10 February 2021 11 February 2021 15 February 2021 16 February 2021 17 February 2021 18 February 2021 23 February 2021 02 March 2021 06 March 2021"
2	22-04-2021	23-04-2021	16-12-20, 06-01-21, 15-01-21, 19-01-21, 22-01-21, 26-01-21, 05-02-21, 11-02-21, 15-02-21	"07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 24 August 2020 26 August 2020 27 August 2020 21 January 2021 22 January 2021 24 January 2021 25 January 2021 26 January 2021 28 January 2021 29 January 2021 30 January 2021 31 January 2021 02 February 2021 03 February 2021 04 February 2021 05 February 2021 06 February 2021 07 February 2021 08 February 2021 09 February 2021 11 February 2021 12 February 2021 13 February 2021 14 February 2021 15 February 2021 16 February 2021 17 February 2021"
3	22-04-2021	23-04-2021	15-02-21	"12 February 2021 13 February 2021 14 February 2021 15 February 2021"
4	22-04-2021	23-04-2021	15-02-21	14 February 2021
5	22-04-2021	23-04-2021	27-01-21	"26 December 2020 27 December 2020 31 December 2020 02 January 2021 03 January 2021 06 January 2021 10 January 2021 12 January 2021 17 January 2021 18 January 2021 22 January 2021 25 January 2021 26 January 2021"
6	22-04-2021	23-04-2021	14-01-21	13 January 2021
7	22-04-2021	23-04-2021	11-01-21	"08 January 2021 09 January 2021 10 January 2021"

8	22-04-2021	23-04-2021	11-01-21, 14-01-21, 18-01-21, 25-01-21, 29-01-21, 01-02-21, 04-02-21, 08-02-21, 11-02-21, 10-05-21, 18-05-21, 02-06-21, 07-06-21, 14-06-21, 17-06-21, 28-06-21, 02-07-21,, 12-07-21, 15-07-21	"09/01/2021 10/01/2021 11/01/2021 13/01/2021 15/01/2021 17-Jan 18-Jan 24/01/2021 25/01/2021 26/01/2021 28/01/2021 31/01/2021 02/02/2021 03/02/2021 04/02/2021 05/02/2021 06/02/2021 07/02/2021 10/02/2021 11/02/2021 13/02/2021 15/02/2021 16/02/2021 17/02/2021 18/02/2021 22/02/2021 24/02/2021 25/02/2021 26/02/2021 27/02/2021 28/02/2021 01/03/2021 02/03/2021 03/03/2021 04/03/2021 05/03/2021 06/03/2021 07/03/2021 08/03/2021 09/03/2021 10/03/2021 11/03/2021 12/03/2021 14/03/2021 16/03/2021 17/03/2021 18/03/2021 20/03/2021 21/03/2021 22/03/2021 24/03/2021 25/03/2021 27/03/2021 28/03/2021 29/03/2021 30/03/2021 31/03/2021 01/04/2021 02/04/2021 03/04/2021 05/04/2021 07/04/2021 08/04/2021 11/04/2021 13 April 2021 14 April 2021 15 April 2021 17 April 2021 18 April 2021 19 April 2021 20 April 2021 21 April 2021 22 April 2021 24 April 2021 25 April 2021 26 April 2021 28 April 2021 10 June 2021 11 June 2021 12 June 2021 13 June 2021 14 June 2021 15 June 2021 16 June 2021 17 June 2021 18 June 2021 19 June 2021 21 June 2021 23 June 2021 25 June 2021 27 June 2021 29 June 2021 01 July 2021 02 July 2021 03 July 2021 05 July 2021 09 July 2021 11 July 2021 12 July 2021 13 July 2021 14 July 2021"
9	22-04-2021	23-04-2021	06-01-21	"03 January 2021 04 January 2021 05 January 2021"
10	22-04-2021	23-04-2021	no warning	21 December 2020
11	22-04-2021	23-04-2021	11-02-21	"06 February 2021 08 February 2021 10 February 2021"
12	22-04-2021	23-04-2021	data not available	
13	22-04-2021	23-04-2021	data not available	
14	22-04-2021	23-04-2021	01-12-20	"30 November 2020"
15	22-04-2021	23-04-2021	30-11-20	"27 November 2020 28 November 2020 29 November 2020"

16	22-04-2021	23-04-2021	16-11-20, 21-10-20, 29-10-20, 12-11-20, 18-11-20, 18-11-20, 27-11-20	"18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 04 August 2020 05 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 20 August 2020 21 August 2020 24 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 17 September 2020 20 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 03 November 2020 04 November 2020 05 November 2020 06 November 2020 07 November 2020 10 November 2020 11 November 2020 12 November 2020 13 November 2020 14 November 2020 15 November 2020 16 November 2020 17 November 2020 18 November 2020 19 November 2020 20 November 2020 21 November 2020 22 November 2020 23 November 2020 24 November 2020 25 November 2020 26 November 2020 27 November 2020 28 November 2020 29 November 2020"
17	22-04-2021	23-04-2021	26-11-20	"24 November 2020 25 November 2020 26 November 2020"
18	23-04-2021	26-04-2021	26-11-20, 02-12-20	"14 September 2020 20 September 2020 26 November 2020 02 December 2020"
19	23-04-2021	26-04-2021	31-07-20, 04-08-20, 10-08-20, 13-08-20, 21-08-20, 25-08-20 , 31-08-20, 06-08-20, 12-10-20, 21-10-20, 27-10-20, 05-11-20, 10-11-20, 17-11-20, 20-11-20, 27-11-20	"12 July 2020 31 July 2020 04 August 2020 10 August 2020 13 August 2020 20 August 2020 24 August 2020 21 August 2020 25 August 2020 31 August 2020 06 September 2020 22 September 2020 09 October 2020 10 October 2020 12 October 2020 19 October 2020 21 October 2020 27 October 2020 05 November 2020 10 November 2020 17 November 2020 20 November 2020 27 November 2020"

20	23-04-2021	26-04-2021	06-08-20, 12-08-20, 13-10-20, 20-10-20, 30-11-20, 30-11-20, 11-12-20, 15-12-20, 05-01-21, 08-01-21, 13-01-21, 19-01-21, 22-01-21, 25-01-21, 29-01-21, 02-02-21, 08-02-21, 11-02-21	"17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 11 August 2020 12 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 24 August 2020 25 August 2020 27 August 2020 28 August 2020 29 August 2020 02 September 2020 03 September 2020 07 September 2020 12 September 2020 17 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 17 January 2021 18 January 2021 19 January 2021 20 January 2021 21 January 2021 22 January 2021 23 January 2021 24 January 2021 25 January 2021 26 January 2021 27 January 2021 28 January 2021 29 January 2021 30 January 2021 31 January 2021 01 February 2021 02 February 2021 03 February 2021 04 February 2021 05 February 2021 06 February 2021 07 February 2021 08 February 2021 09 February 2021 10 February 2021 12/02/2021"
21	23-04-2021	26-04-2021	12-10-20, 25-11-20, 30-11-20	"22 July 2020 23 July 2020 24 July 2020 25 July 2020 25 August 2020 26 August 2020 27 August 2020 28 February 2020 29 August 2020 02 September 2020 03 September 2020 08 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020 13 October 2020 24 November 2020 25 November 2020 29 November 2020"
22	23-04-2021	26-04-2021	21-10-20, 28-10-20, 13-11-20	"16 October 2020 17 October 2020 18 October 2020 20 October 2020 21 October 2020 22 October 2020 23 October 2020 26 October 2020 27 October 2020 28 October 2020 29 October 2020 30 October 2020 31 October 2020 02 November 2020 03 November 2020 04 November 2020 05 November 2020 06 November 2020 07 November 2020 08 November 2020 09 November 2020 10 November 2020 13 November 2020"

23	23-04-2021	26-04-2021	12-11-20	"20 July 2020 21 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 24 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 31 August 2020 02 September 2020 05 September 2020 09 September 2020 10 September 2020 13 September 2020 15 September 2020 16 September 2020 17 September 2020 19 September 2020 20 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 5-20-2020 06 October 2020 07 October 2020 08 October 2020 18 October 2020 19 October 2020 20 October 2020 21 October 2020 22 October 2020 23 October 2020 24 October 2020 25 October 2020 27 October 2020 28 October 2020 29 October 2020 30 October 2020 31 October 2020 01 November 2020 02 November 2020 03 November 2020 04 November 2020 05 November 2020 06 November 2020 07 November 2020 08/11/2020 09/11/2020 10/11/2020 11/11/2020 12/11/2020"
24	23-04-2021	26-04-2021	05-08-20, 21-10-20, 28-10-20, 11-11-20	"13 July 2020 14 July 2020 15 July 2020 16 July 2020 17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 04 September 2020 06 September 2020 11 September 2020 16 September 2020 18 September 2020 19 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 12 October 2020 13 October 2020 14 October 2020 15 October 2020 16 October 2020 17 October 2020 18 October 2020 19 October 2020 20 October 2020 21 October 2020 22 October 2020 23 October 2020 24 October 2020 27 October 2020 28 October 2020 30 October 2020 01 November 2020 02 November 2020 03 November 2020 06 November 2020 10 November 2020"
25	22-04-2021	26-04-2021	same as 4	same as 4

26	23-04-2021	26-04-2021	31-07-20, 04-08-20, 10-08-20, 10-08-20, 14-08-20, 24-08-20, 28-08-20, 03-09-20, 28-10-20, 06-11-20, 1-11-20	"18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 22 August 2020 23 August 2020 24 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 02 September 2020 05 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020 13 October 2020 14 October 2020 15 October 2020 16 October 2020 17 October 2020 18 October 2020 21 October 2020 22 October 2020 24 October 2020 25 October 2020 26 October 2020 27 October 2020 28 October 2020 29 October 2020 30 October 2020 31 October 2020 01 November 2020 02 November 2020 04 November 2020 05 November 2020 06 November 2020"
27	23-04-2021	26-04-2021	30-07-20, 06-11-20	"13 July 2020 14 July 2020 15 July 2020 16 July 2020 17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 October 2020 01 November 2020 02 November 2020 03 November 2020 04 November 2020 05 November 2020 06 November 2020"
28	23-04-2021	26-04-2021	no warning	"03 November 2020"
29	23-04-2021	26-04-2021	13-10-20, 19-10-20	"20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 31 August 2020 02 September 2020 03 September 2020 05 September 2020 06 September 2020 07 September 2020 12 September 2020 14 September 2020 15 September 2020 17 September 2020 19 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020 13 October 2020 14 October 2020 15 October 2020 16 October 2020 17 October 2020 18 October 2020 19 October 2020 20 October 2020 21 October 2020"
30	23-04-2021	26-04-2021	21-10-20, 27-10-20	"17 July 2020 27 August 2020 20 October 2020"

31	23-04-2021	26-04-2021	20-10-20	"19 October 2020 20 October 2020"
32	23-04-2021	26-04-2021	12-10-20, 16-10-20	"12 July 2020 16 July 2020 17 July 2020 19 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 30 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 22 August 2020 23 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 31 August 2020 01 September 2020 21 September 2020 23 September 2020 25 September 2020 26 September 2020 27 September 2020 29 September 2020 30 September 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 11 October 2020 13 October 2020 14 October 2020 15 October 2020 16 October 2020 17 October 2020 18 October 2020"
33	23-04-2021	26-04-2021	19-10-20	"12 October 2020 13 October 2020 14 October 2020 15 October 2020 16 October 2020 18 October 2020 19 October 2020 21 October 2020"
34	23-04-2021	26-04-2021	31-07-20, 07-08-20, 11-08-20, 14-10-20	"19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 24 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 31 August 2020 02 September 2020 03 September 2020 08 September 2020 11 September 2020 22 September 2020 24 September 2020 25 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020 13 October 2020 14 October 2020 15 October 2020"

35	23-04-2021	26-04-2021	14-10-20	"17 July 2020 18 July 2020 19 July 2020 21 July 2020 22 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 19 August 2020 20 August 2020 22 August 2020 23 August 2020 24 August 2020 25 August 2020 26 August 2020 28 August 2020 30 August 2020 31 August 2020 04 September 2020 07 September 2020 09 September 2020 10 September 2020 16 September 2020 19 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020 13-20-2020 14 October 2020"
36	23-04-2021	26-04-2021	13-10-20	"17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 07 August 2020 08 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 23 August 2020 24 August 2020 29 August 2020 02 September 2020 03 September 2020 05 September 2020 08 September 2020 11 September 2020 17 September 2020 22 September 2020 25 September 2020 28 September 2020 29 September 2020 01 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 09 October 2020 10 October 2020 12 October 2020 13 October 2020"
37	23-04-2021	26-04-2021	12-10-20	"22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020"
38	26-04-2021	28-04-2021	12-10-20	"21 September 2020 22 September 2020 23 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020"
39	26-04-2021	28-04-2021	31-07-20, 05-08-20, 11-08-20, 14-08-20, 20-08-20, 21-08-20	"16 July 2020 17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 02 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 10 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020"

40	26-04-2021	28-04-2021	11-08-20, 26-08-20, 21-08-20	"12 July 2020 13 July 2020 14 July 2020 15 July 2020 16 July 2020 18 July 2020 20 July 2020 23 July 2020 24 July 2020 25 July 2020 28 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 07 August 2020 11 August 2020 13 August 2020 20 August 2020 21 August 2020 23 August 2020 24 August 2020 25 August 2020 26 August 2020 30 August 2020"
41	26-04-2021	28-04-2021	21-08-20	"20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020"
42	26-04-2021	29-04-2021	30-07-20, 03-08-20	"13 July 2020 14 July 2020 15 July 2020 16 July 2020 17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 01 August 2020 02 August 2020 03 August 2020"
43	26-04-2021	29-04-2021	data not available	
44	26-04-2021	29-04-2021	data not available	
45	23-04-2021	29-04-2021	15-10-20	"17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 09 August 2020 10 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 31 August 2020 01 September 2020 02 September 2020 05 September 2020 07 September 2020 08 September 2020 09 September 2020 14 September 2020 15 September 2020 18 September 2020 19 September 2020 20 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 09 October 2020 10 October 2020 11 October 2020 12 October 2020 13 October 2020 14 October 2020 15 October 2020"
46	22-04-2021	29-04-2021	11-12-20	"09 December 2020 10 December 2020 11 December 2020"

47	23-04-2021	29-04-2021	12-10-20, 19 Oct, 20, 23-10-20, 27-10-20	"18 July 2020 19 July 2020 21 July 2020 22 July 2020 23 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 01 August 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 24 August 2020 25 August 2020 26 August 2020 27 August 2020 29 August 2020 30 August 2020 02 September 2020 05 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 05 October 2020 06 October 2020 07 October 2020 08 October 2020 10 October 2020 11 October 2020 12 October 2020 13 October 2020 14 October 2020 16 October 2020 17 October 2020 18 October 2020 19 October 2020 21 October 2020 22 October 2020 23 October 2020 24 October 2020 26 October 2020 27 October 2020 28 October 2020 29 October 2020"
48	23-04-2021	29-04-2021	22-10-20, 18-11-20	"13 July 2020 16 July 2020 17 July 2020 19 November 2020 30 July 2020 31 July 2020 04 August 2020 05 August 2020 09 August 2020 11 August 2020 23 August 2020 25 August 2020 26 August 2020 29 August 2020 02 September 2020 03 September 2020 05 September 2020 07 September 2020 10 September 2020 13 September 2020 15 September 2020 18 September 2020 21 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 28 September 2020 02 October 2020 03 October 2020 04 October 2020 06 October 2020 10 October 2020 22 October 2020 31 October 2020 01 November 2020 06 November 2020 09 November 2020 10 November 2020 14 November 2020 17 November 2020 18 November 2020"
49	22-04-2021	04-05-2021	01-02-21	"31 January 2021"
50	22-04-2021	04-05-2021	10-10-20, 12-10-20, 20-10-20, 26-10-20, 29-10-20, 06-11-20, 12-11-20, 17-11-20, 20-11-20, 25-11-20, 30-11-20, 03-12-20, 07-12-20, 10-12-20, 14-12-20, 17-12-20, 06-01-21, 11-01-21, 15-01-21, 18-01-21, 21-01-21, 25-01-21, 28-01-21, 04-02-21	"19 July 2020 20 July 2020 21 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 February 2020 30 July 2020 31 July 2020 02 August 2021 03 August 2020 04 August 2020 05 August 2020 06 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 21 August 2020 22 August 2020 24 August 2020 25 August 2020 26 August 2020 27 August 2020 28 August 2020 29 August 2020 30 August 2020 17 September 2020 18 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 30 September 2020 01 October 2020 02 October 2020 03 October 2020 04 October 2020 05 October 2020 06 October 2020 08 October 2020 03 January 2021 05 January 2021 06 January 2021 07 January 2021 08 January 2021 09 January 2021 10 January 2021 11 January 2021 12 January 2021 13 January 2021 14 January 2021 15 January 2021 16 January 2021 17 January 2021 18 January 2021 19 January 2021 20 January 2021 21 January 2021 22 January 2021 23 January 2021 24 January 2021 25 January 2021 26 January 2021 27 January 2021 28 January 2021"

51	22-04-2021	04-05-2021	16-12-20, 12-01-21, 15-01-21, 19-01-21, 22-01-21, 26-01-21, 05-02-21, 08-02-21, 11-02-21, 15-02-21	"07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 13 August 2020 19 August 2020 20 August 2020 21 August 2020 22 August 2020 23 August 2020 24 August 2020 26 August 2020 27 August 2020 21 January 2021 22 January 2021 24 January 2021 25 January 2021 26 January 2021 28 February 2021 29 January 2021 30 January 2021 31 January 2021 02 February 2021 03 February 2021 04 February 2021 05 February 2021 06 February 2021 07 February 2021 08 February 2021 09 February 2021 10 February 2021 11 February 2021 12 February 2021 13 February 2021 14 February 2021 15 February 2021 16 February 2021 17 February 2021"
52	22-04-2021	04-05-2021	16-10-20, 22-10-20, 27-10-20, 06-11-20, 09-11-20, 13-11-20, 17-11-20, 25-11-20, 01-12-20, 04-12-20, 07-12-20, 10-12-20, 15-12-20, 18-12-20, 01-02-21	"17 July 2020 18 July 2020 19 July 2020 20 July 2020 21 July 2020 22 July 2020 23 July 2020 24 July 2020 25 July 2020 26 July 2020 27 July 2020 28 July 2020 29 July 2020 30 July 2020 31 July 2020 02 August 2020 03 August 2020 04 August 2020 05 August 2020 07 August 2020 08 August 2020 09 August 2020 10 August 2020 11 August 2020 12 August 2020 19 August 2020 20 August 2020 22 August 2020 23 August 2020 24 August 2020 25 August 2020 27 August 2020 28 August 2020 02 September 2020 03 September 2020 06 September 2020 11 September 2020 13 September 2020 19 September 2020 20 September 2020 22 September 2020 23 September 2020 24 September 2020 25 September 2020 26 September 2020 27 September 2020 28 September 2020 29 September 2020 30 September 2020 02 October 2020 03 October 2020 04 October 2020 06 October 2020 07 October 2020 08 October 2020 29 November 2020 30 November 2020 01 December 2020 02 December 2020 03 December 2020 04 December 2020 05 December 2020 06 December 2020 08 December 2020 09 December 2020 10 December 2020 11 December 2020 14 December 2020 15 December 2020 16 December 2020 17 December 2020 18 December 2020 19 December 2020 20 December 2020 21 December 2020 22 December 2020 23 December 2020 24 December 2020 25 December 2020 26 December 2020"
53	23-04-2021	04-05-2021	11-11-20, 18-11-20, 25-11-20	"10 November 2020 12 November 2020 13 November 2020 14 November 2020 15 November 2020 16 November 2020 17 November 2020 18 November 2020"
54	22-04-2021	04-05-2021	24-12-20, 05-01-21	"23 December 2020 24 December 2020 25 December 2020 26 December 2020 27 December 2020 28 December 2020 29 December 2020 31 December 2020 02 January 2021 03 January 2021 04 January 2021 05 January 2021"
55	30-04-2021	10-05-2021	16-11-20	"15 November 2020"
56	30-04-2021	10-05-2021	29-04-21	"17 April 2021 18 April 2021 19 April 2021 20 April 2021 21 April 2021 22 April 2021 23 April 2021 24 April 2021 25 April 2021 26 April 2021 27 April 2021 28 April 2021 29 April 2021"
57	30-04-2021	10-05-2021	10-05-21, 19-05-21	"10 April 2021 11 April 2021 13 April 2021 14 April 2021 15 April 2021 17 April 2021 18 April 2021 20 April 2021 21 April 2021 22 April 2021 25 April 2021 27 April 2021 28 April 2021 07 May 2021 08 May 2021 09 May 2021 10 May 2021 11 May 2021 12 May 2021 13 May 2021 14 May 2021 15 May 2021 16 May 2021 17 May 2021 18 May 2021"
58	30-04-2021	10-05-2021	29-04-21	"07 April 2021 13 April 2021 16 April 2021 17 April 2021"

r.	*interview*	*1st warning*	*1st infection*			*2nd infection*		
1	23/04/2021	23/10/2020	20/10/2020	until	06/03/2021			
2	23/04/2021	16/12/2020	07/08/2020	until	27/08/2020	21/01/2021	until	17/02/2021
3	23/04/2021	15/02/2021	12/02/2021	until	15/02/2021			
4	23/04/2021	15/02/2021	14/02/2021					
5	23/04/2021	27/01/2021	26/12/2020	until	26/01/2021			
6	23/04/2021	14/01/2021	13/01/2021					
7	23/04/2021	11/01/2021	08/01/2021	until	10/01/2021			
8	23/04/2021	11/01/2021	09/01/2021	until	28/04/2021	10/06/2021	until	14/07/2021
9	23/04/2021	06/01/2021	03/01/2021	until	05/01/2021			
10	23/04/2021	no warning	21/12/2020					
11	23/04/2021	11/02/2021	06/02/2021	until	10/02/2021			
12	23/04/2021	data not available						
13	23/04/2021	data not available						
14	23/04/2021	01/12/2020	30/11/2020					
15	23/04/2021	30/11/2020	27/11/2020	until	29/11/2020			
16	23/04/2021	16/11/2020	18/07/2020	until	29/11/2020			
17	23/04/2021	26/11/2020	24/11/2020	until	26/11/2020			
18	26/04/2021	26/11/2020	14/09/2020	until	20/09/2020	26/11/2020	until	02/12/2020
19	26/04/2021	31/07/2020	12/07/2020	until	27/11/2020			
20	26/04/2021	06/08/2020	17/07/2020	until	08/10/2020	17/01/2021	until	12/02/2021
21	26/04/2021	12/10/2020	22/07/2020	until	29/11/2020			
22	26/04/2021	21/10/2020	16/10/2020	until	13/11/2020			
23	26/04/2021	12/11/2020	20/07/2020	until	12/11/2020			
24	26/04/2021	05/08/2020	13/07/2020	until	10/11/2020			
25	26/04/2021	same as 4						
26	26/04/2021	31/07/2020	18/07/2020	until	06/11/2020			
27	26/04/2021	30/07/2020	13/07/2020	until	10/11/2020			
28	26/04/2021	no warning	03/11/2020					
29	26/04/2021	13/10/2020	20/07/2020	until	21/10/2020			
30	26/04/2021	21/10/2020	27/08/2020			20/10/2020		
31	26/04/2021	20/10/2020	19/10/2020	until	20/10/2020			
32	26/04/2021	12/10/2020	12/07/2020	until	18/10/2020			
33	26/04/2021	19/10/2020	12/10/2020	until	21/10/2020			
34	26/04/2021	31/07/2020	19/07/2020	until	15/10/2020			
35	26/04/2021	14/10/2020	17/07/2020	until	14/10/2020			
36	26/04/2021	13/10/2020	17/07/2020	until	13/10/2020			
37	26/04/2021	12/10/2020	22/09/2020	until	12/10/2020			
38	28/04/2021	12/10/2020	21/09/2020	until	12/10/2020			
39	28/04/2021	31/07/2020	16/07/2020	until	23/08/2020			
40	28/04/2021	11/08/2020	12/07/2020	until	30/08/2020			
41	28/04/2021	21/08/2020	20/07/2020	until	21/08/2020			
42	29/04/2021	30/07/2020	13/07/2020	until	03/08/2020			
43	29/04/2021	data not available						
44	29/04/2021	data not available						
45	29/04/2021	15/10/2020	17/07/2020	until	15/10/2020			
46	29/04/2021	11/12/2020	09/12/2020	until	11/12/2020			
47	29/04/2021	12/10/2020	18/07/2020	until	29/10/2020			
48	04/05/2021	22/10/2020	13/07/2020	until	18/11/2020			
49	04/05/2021	01/02/2021	31/01/2021					
50	04/05/2021	10/10/2020	19/07/2020	until	28/01/2021			
51	04/05/2021	16/12/2020	07/08/2020	until	27/08/2020	21/01/2021	until	17/02/2021
52	04/05/2021	16/10/2020	17/07/2020	until	26/12/2020			
53	04/05/2021	11/11/2020	10/11/2020	until	18/11/2020			
54	04/05/2021	24/12/2020	23/11/2020	until	05/01/2021			
55	10/05/2021	16/11/2020	15/11/2020					
56	10/05/2021	29/04/2021	17/04/2021	until	29/04/2021			
57	10/05/2021	10/05/2021	10/04/2021	until	18/05/2021			
58	10/05/2021	29/04/2021	07/04/2021	until	17/04/2021			

Table 52: Results of the respondents concerning the interview, first warning, first infection, and second infection periods and dates

Appendix M SSH script to remove QSnatch provided by a respondent

```
#!/bin/sh

######################
# Derek Be Gone v1.4 #
# Author: qnapd      #
######################

set_mutable() {
    if [ ! -e "$1" ]; then
        return 0
    fi
    if [ -e /etc/IS_64BITS ]; then
        # 64bit set mutable
        SET_M_64="$1"
        #echo " [*] Setting mutable 64bit on $SET_M_64"
        python -c "import os,fcntl,sys,struct;fd = os.open('${SET_M_64}', os.O_RDONLY); rec = struct.pack('L', 0); x = fcntl.ioctl(fd, 0x80086601, rec); flags = struct.unpack('L',x)[0]; was_immutable = flags & 0x00000010; flags = flags & ~0x00000010; f = struct.pack('i', flags); fcntl.ioctl(fd, 0x40086602, f); os.close(fd)"
    else
        # 32bit set mutable
        SET_M_32="$1"
        #echo " [*] Setting mutable 32bit on $SET_M_32"
        python -c "import os,fcntl,sys,struct;fd = os.open('${SET_M_32}', os.O_RDONLY); rec = struct.pack('L', 0); x = fcntl.ioctl(fd, 0x80046601, rec); flags = struct.unpack('L',x)[0]; was_immutable = flags & 0x00000010; flags = flags & ~0x00000010; f = struct.pack('i', flags); fcntl.ioctl(fd, 0x40046602, f); os.close(fd)"
    fi
}

remove_bad_thing() {
    badpath="$1"
    if [ -e "$badpath" ]; then
        echo " [*] Removing $badpath"
        rm -rf "$badpath"
        if [ $? -eq 0 ]; then
            echo "  [+] Success!"
        else
            echo "  [-] Failed"
        fi
    fi
}

sterilise() {
 FILE="$1"
 KEY='7C0vK4SzMO15zBxLD7XCi5hbjgP1ZjkJ'
 if grep -q $KEY $FILE; then
  set_mutable "$FILE"
  echo " [*] Sterilise $FILE"
  sed -i 's/'"$KEY"'/NOPE/g' "$FILE"
  if grep $KEY $FILE; then
   echo "  [-] Failed"
  else
   echo "  [+] Success!"
  fi
 fi
}

echo ">>> derek-be-gone v1.4"

# clear fake qpkg
bdir=
test -f "${confdir}/smb.conf" && for i in homes Public Download Multimedia Web Recordings; do bdir=`getcfg "$i
" path -f "${confdir}/smb.conf"` && test ! -z "$bdir" && bdir=`dirname "$bdir"` && test -d "$bdir" && testwriteab
le=$(mktemp "${bdir}/.tmp.XXXXXX") && rm "${testwriteable}" && break; bdir="; done
test -z "${bdir}" || test ! -d "${bdir}" && { command -v readlink >/dev/null 2>&1 || ln -sf /bin/busybox /usr/bin/read
link; for i in Public Download Multimedia Web Recordings homes; do bdir=`readlink "/share/${i}" 2>/dev/null` &&
 test ! -z "$bdir" && bdir=`dirname "$bdir"` && bdir=/share/${bdir##*/} && test -d "$bdir" && break; done;
test -z "${bdir}" || test ! -d "${bdir}"; } && { bdir=`getcfg SHARE_DEF defVolMP -f "${confdir}/def_share.info"`
test -z "${bdir}" || test ! -d "${bdir}"; } && { bdir=`mount | sed -n "s/.*\(\/share\/[^ /]\+\) .*/\1/gp" | head -n 1`
test -z "${bdir}" || test ! -d "${bdir}"; } && { for i in CACHEDEV3_DATA CACHEDEV2_DATA CACHEDEV1
_DATA MD0_DATA; do test -d "/share/${i}" && bdir="/share/${i}" && break; done;
test -z "${bdir}" || test ! -d "${bdir}" && bdir=/mnt/HDA_ROOT; }

echo "[o] System path: ${bdir}"
echo "[o] Removing fake qpkg"

set_mutable "${bdir}/.qpkg/.liveupdate/liveupdate.sh"
set_mutable "${bdir}/.qpkg/.liveupdate/"
remove_bad_thing "${bdir}/.qpkg/.liveupdate/liveupdate.sh"
remove_bad_thing "${bdir}/.qpkg/.liveupdate/"
ln -sf /dev/null "${bdir}/.qpkg/.liveupdate"

set_mutable "${bdir}/.qpkg/.config/backup_conf.sh"
set_mutable "${bdir}/.qpkg/.config/"
remove_bad_thing "${bdir}/.qpkg/.config/backup_conf.sh"
remove_bad_thing "${bdir}/.qpkg/.config/"
ln -sf /dev/null "${bdir}/.qpkg/.config"

for i in /etc/rcK_init.d/K0*.sh; do
 remove_bad_thing "$i"
done
echo "---"

# infected qpkg
echo "[o] Sterilising infected QPKG"
for i in $(grep -i shell /etc/config/qpkg.conf | cut -d'=' -f2 | grep -v null); do
 sterilise "$i"
done

# clear dom
mdir=/tmp/config
__BOOT_DEV=
__model=`getcfg System "Internal Model"`
CONFIG_DEV_NODE=`getcfg "CONFIG STORAGE" DEVICE_NODE -f /etc/platform.conf`
CONFIG_DEV_PART=`getcfg "CONFIG STORAGE" FS_ACTIVE_PARTITION -f /etc/platform.conf`
CONFIG_DEV_FS=`getcfg "CONFIG STORAGE" FS_TYPE -f /etc/platform.conf`
__BOOT_CONF=`test -f /etc/default_config/BOOT.conf && cat /etc/default_config/BOOT.conf 2>/dev/null || cat "
${confdir}/BOOT.conf"` || { test "$arch_o" = arm && __BOOT_CONF=TS-NASARM; }
command -v hal_app > /dev/null 2>&1 && { __BOOT_DEV=$(hal_app --get_boot_pd port_id=0); }
```

```
test "${__BOOT_CONF}" = TS-NASARM || test "$arch_o" = arm && { test -f /etc/IS_TAS && __BOOT_DEV="
${__BOOT_DEV:-/dev/mtdblock}7" || __BOOT_DEV="${__BOOT_DEV:-/dev/mtdblock}5"; } || __BOOT_DEV
="${__BOOT_DEV:-/dev/sdx}6"
test "x${CONFIG_DEV_NODE}" != "x" && { ubiattach -m "${CONFIG_DEV_PART}" -d 2; mount -t ubifs ubi2:
config "${mdir}" > /dev/null 2>&1 || { test -f /etc/IS_TAS && mount -t ext4 /dev/mmcblk0p7 "${mdir}"; } } || mou
nt ${__BOOT_DEV} -t ext2 ${mdir} || { test "${__model}" = "TS-201" && mount -t ext2 /dev/mtdblock4 ${mdir}
; } || { ubiattach -m "${CONFIG_DEV_PART}" -d 2; mount -t ubifs ubi2:config "${mdir}"; mount -t ext4 /dev/mm
cblk0p7 "${mdir}"; } || { test "${__model}" = "TS-269L" && mount -t ext2 /dev/sdc6 ${mdir}; } || { test "${__mod
el}" = "TS-869" && mount -t ext2 /dev/sdi6 ${mdir}; } || { test "$arch_o" = arm || ${__BOOT_CONF} = "TS-NAS
ARM" && { for i in 5 7 4 6 3 8; do mount -t ext2 "/dev/mtdblock${i}" ${mdir} && break; done; }; } || { test "$arch
_o" = x86 && for n in /dev/sdc /dev/sdx /dev/sdi $__BOOT_DEV; do for i in 6 $CONFIG_DEV_PART; do mount
-t ext2 ${n}${i} ${mdir} && break 2; done; done; } || { mount -t ext2 $(/sbin/hal_app --get_boot_pd port_id=0)6 ${
mdir}; }

echo "[o] Cleaning DOM"
set_mutable "$mdir/autorun.sh"
remove_bad_thing "$mdir/autorun.sh"
for i in $mdir/K0*.sh; do
 set_mutable "$i"
   remove_bad_thing "$i"
done
umount "$mdir"
echo "---"

# naughty xml
echo "[o] Remove bad XML"
for i in /etc/config/rssdoc/qpkgcenter_*.xml; do
 set_mutable "$i"
   remove_bad_thing "$i"
done
echo "---"

# reinstall MR
echo "[o] Remove old MR"
mrpath="${bdir}/.qpkg/MalwareRemover/"
set_mutable "$mrpath"
set_mutable "$mrpath/modules/10_derek_3.pyc"
set_mutable "$mrpath/modules/12_derek_3.pyc"
remove_bad_thing "$mrpath"
set_mutable /etc/config/qpkg.conf
rmcfg MalwareRemover -f /etc/config/qpkg.conf
echo "---"

echo "[o] Install new MR"
mrpkg=MalwareRemover_3.4.1_20190125_182348
echo "" > /etc/hosts
wget -nv "https://download.qnap.com/QPKG/${mrpkg}.zip"
unzip "${mrpkg}.zip"
sh "${mrpkg}.qpkg" > /dev/null 2>&1
getcfg MalwareRemover Enable -f /etc/config/qpkg.conf > /dev/null
if [ $? -eq 0 ]; then
   echo " [+] Success!"
else
   echo " [-] Failed"
```

```
fi
rm -f "${mrpkg}.zip" "${mrpkg}.qpkg"
echo "---"
echo "Finished!"
rm -f "$0"
```