EXPLORING EFFECTIVE NOTIFICATION MECHANISMS FOR INFECTED IOT DEVICES

MASTER THESIS

E.M. ALTENA

EXPLORING EFFECTIVE NOTIFICATION MECHANISMS FOR INFECTED IOT DEVICES

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

MASTER OF SCIENCE

in Management of Technology
Faculty of Technology, Policy and Management

by

Elisabeth Maria Altena
Student number: 1526413

To be defended in public on April 26th, 2018

PUBLIC VERSION

Graduation committee

Chairperson	: Prof. M.J.G. van Eeten, Section Organisation & Governance
First Supervisor	: Dr. C. H. Gañán, Section Organisation & Governance
Second Supervisor	: Dr. G.A. de Reuver, Section Information and Communication Technology
Advisor	: O. Çetin, Section Organisation & Governance
External Supervisor	: D. van Beusekom, KPN

PREFACE

This master thesis presents my final project with which I conclude the master Management of Technology at the faculty of Technology, Policy and Management of the Delft University of Technology. The research project was carried out in cooperation with KPN. At KPN's Abuse team, I investigated the most effective notification mechanism for customers with infected IoT devices.

Working on this project has been a great experience. Having the time and opportunity to dive into such a much-discussed topic has been very interesting. Moreover, being given the opportunity to collaborate on three scientific publications during this research project is something I could have not wished for in advance. This has served as a great motivation to work a little harder.

It would have been impossible to complete this thesis without the help and support I received from some people. I would like to thank them for their guidance during the project. First of all I want to thank Folkert Visser and Dennis van Beusekom for offering me the opportunity to perform this project at KPN's Abuse team. I also want to thank the Abuse team for their kind and indispensable contribution to the experiment and willingness to answer all my questions.

From the TU Delft, I want to thank Orçun Çetin, my daily supervisor, for his excellent guidance and support during my thesis. You have taught me a lot. Furthermore, I would like to thank my supervisors Michel van Eeten, Carlos Gañán and Mark de Reuver. Their comments and feedback have been very valuable.

Lastly, I would like to express my thanks to my family and friends for the much needed distraction, the pleasant coffee breaks in the library and for encouraging me throughout the process. Lieuwe Thys, thank you for all your love, patience, support and ability to put things in perspective.

Lisette Altena
Delft, April 2018

SUMMARY

Many Internet of Things (IoT) devices that are currently on the market lack security and therefore many of them got infected with malware to launch powerful distributed denial of service (DDoS) attacks. Notifications from Internet Service Providers (ISPs) to their customers play a crucial role in the fight to clean up the malware infected IoT devices. It is, however, difficult for the employees of abuse departments to explain how to cleanup an IoT malware infection to a non-technical customer and provide usable action steps to clean up the infection. In particular, because there is no "one size fits all" cleanup solution, due to the heterogeneous nature of the IoT devices. The abuse department of the Dutch ISP KPN would like to know how to notify customers with IoT malware infections and how to explain the cleanup of infected IoT devices to the customers. Therefore, the objective of this research is to make a recommendation to KPN on what notification mechanism to adopt by providing insight into: (1) how to increase the effectiveness of IoT malware notifications from an ISP to its customers in terms of IoT malware cleanup; and (2) how users perceive an IoT malware notification from their ISP. The main research question that is answered in this research is: *What notification mechanism is the most effective in terms of both IoT malware cleanup and improving the reactions of customers?* To this end, an experiment has been conducted with 190 retail customers with infected IoT devices to measure the difference in cleanup among IoT malware notifications sent via different channels and with different messages. To explore the reactions of the customers to the different notification mechanisms, telephone interviews have been conducted and the communication logs between KPN and the customers in the experiment have been analysed. We have compared the influence of the notification channel on cleanup and the reactions of customers by comparing customers that received: (1) email notifications; and (2) a combination of walled garden and email notifications. The different notification messages that have been compared in this study include: (1) the walled garden notification content that KPN's abuse department uses to notify its customers with an IoT malware infection; and (2) a newly composed more actionable walled garden notification message which clearly defines the steps that need to be taken and avoids technical terms.

Firstly, in the experiment we found that a walled garden notification with a more actionable content significantly reduces the infection time of the IoT malware infection compared to the infection of quarantined customers that received the old content and customers in the email-only treatment group. Surprisingly, we found no measurable differences in terms of infection time and cleanup ware when comparing email notifications and walled garden notifications with the old content to the control group. Secondly, the analysis of the customer reactions showed that quarantining improves the customer reaction time and reaction rate after an IoT malware notification significantly compared to email notifications. In addition, walled garden notifications have a higher probability of being read and more often encourage people to disconnect their device from the Internet. However, in some cases the quarantine event leads to complaints over the disruption. Regarding the notification content, it is found that the more actionable content of a walled garden notification does not make a difference in the reaction time and reaction rate compared to a less actionable content of the notification. Though, the newly composed notification content improves the understanding and trust compared to the old notification content. Lastly, an analysis of the correlation between variables related to the customer's understanding of the notification and the cleanup showed that customers' apparent misunderstanding of the IoT malware notification does not always correlate with a longer infection period. Only the customers that requested additional help to clean up the IoT malware infection have a significantly longer infection period than the customers who did not request additional help.

From these results, we can conclude that a combination of a walled garden and email notification with an actionable content is the most effective in terms of IoT malware cleanup. Furthermore, the walled garden notification is most effective in getting customers to read and react to the IoT malware notification, yet it sometimes results in customers having a low satisfaction with the service they receive. The more actionable notification content results in better understanding and trust from the customer compared to a less actionable content of the notification. However, customers' understanding of the notification content and the satisfaction with the quarantine event remain a challenge.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

C-HIP	Communication-Human Information Processing
CPU	Central processing unit
C&C	Command & Control
DDoS	Distributed Denial of Service
DNS	Domain Name System
DVR	Digital Video Recorder
HTTP	Hypertext Transfer Protocol
IM	Instant message
IP	Internet Protocol
IoT	Internet of Things
IoTPot	Internet of Things Honeypot
ISP	Internet Service Provider
PMT	Protection Motivation Theory
SSH	Secure Shell
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

1 INTRODUCTION

1.1 RESEARCH CONTEXT

The Internet of Things (IoT) is an abstract term that encloses different definitions. The IoT refers to physical objects which are connected to the Internet, such as a PlayStation. Others restrict the IoT to sensor networks. Gartner (n.d.) defines the IoT as "the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment." An often used example is a smart fridge which could calculate the freshness of food or automatically order essential items when you are running low. On a larger scale, the IoT can be implemented to improve production processes. Also, it might provide solutions for problems concerning energy, the environment, crime, healthcare and education. In smart cities where multiple devices and systems are connected, information sharing leads to improved processes. For example energy savings, improved traffic flow and the detection of broken city lights. Cisco expects that by 2020, 50 billion networked devices will be in use. Assuming a world population of 7,6 billion, this corresponds with 6,58 connected devices per person (Evans, 2011).

The IoT is not just a blessing, IoT devices present a variety of security risks that could be exploited to cause harm. These risks are mainly caused by the fact that manufacturers of the IoT devices are rushing to make their devices Internet connected. However, in many cases, it is done with little thought and knowledge about security. Many IoT devices, such as digital video recorders (DVRs), IP cameras and routers, are equipped with default passwords. Since many people do not change the passwords, criminals can access the devices by using a list of default passwords. This lack of security could enable intruders to access and misuse personal information, which is collected and transmitted to or from the device. Secondly, the exploitation of vulnerabilities might create safety risks. For example, when an insulin pump is hacked, the intruder is able to change settings so it no longer delivers medicine (Federal Trade Commission, 2015). In addition, the hijacked devices can be used in botnets. These networks of hijacked devices have been used in distributed denial of service (DDoS) attacks, to shut down websites and extort companies. In October 2016, a large DDoS attack on Dyn, a company that manages domain name system (DNS) servers, took place. In turn, many websites were unavailable. The attack appeared to have relied on hundreds of thousands of infected IoT devices. "Though it was not the first time hackers used the Internet of Things to power an attack, the scale of the effort against Dyn was a revelation to people who didn't realize that having Internet-connected things knitted into daily life would come with new risks" (Markoff, 2016).

In order to mitigate the spread of IoT malware, interventions are required. An emerging consensus is that Internet service providers (ISPs) play an important role in controlling the spread of malware infecting conventional computers, as they are best positioned to intervene (Van Eeten & Bauer, 2008). They receive information about malicious traffic in their network which they can act upon. Once the ISP has identified or been informed about infections in its network, it can link the IP address to customer details and is therefore able to notify and assist customers to perform cleanup (Hofmeyr, Moore, Forrest, Edwards, & Stelle, 2013). ISPs can take a number of different actions to resolve the problem. ISPs can notify the customer through a variety of channels, such as email or phone, and inform the customer about where to find anti-virus software or directly provide customer with the tools they need to disinfect their computers. Moreover, ISPs can use what is arguably the most costly notification channel: placing an infection customer into a quarantine environment, also known as a 'walled garden', which only gives access to a small set of white-listed sites. The different notification mechanisms embody varying degrees of intervention and differ in terms of costs and effectiveness. ISPs need to balance these different considerations in order to decide which notification mechanism to adopt.

1.2 PROBLEM STATEMENT

This study is executed at KPN, a Dutch ISP. Since 2003, KPN has an abuse department that acts upon the information about malicious traffic in their network. They notify customers with malware infected machines by quarantining them and sending an email simultaneously. In the case of malware infections on desktop computers and laptops, they provide the customers with anti-virus software which is known to be effective in terms of removing the infection. However, the abuse department encounters issues with notifying customers infected with IoT malware, as such cleanup tools do not exist for IoT malware infections. In addition, employees of the abuse department notice that customers often do not realise which of their devices, other than computers, laptops, tablets and smartphones, are connected to the Internet. Let alone being aware of the possibility of malware infections on these IoT devices. It is difficult for the employees of the abuse department to explain how to clean up an IoT malware infection to a non-technical customer and provide usable action steps to clean up the infection. In particular, because there is no "one size fits all" cleanup solution, due to the heterogeneous nature of the IoT devices. Therefore, KPN's abuse department would like to know how to notify customers with IoT malware infections and how to explain the cleanup of infected IoT devices to the customers.

Previous research has shown that notifications about vulnerabilities and compromises, improve remediation and cleanup. These studies analysed multiple variables that could impact the notification effectiveness, including the the verbosity and level of detail of the notification message on cleanup (Çetin, Gañán, Korczyski, & Van Eeten, 2017; Çetin, Hanif Jhaveri, Gañán, van Eeten, & Moore, 2016; Li, Durumeric, et al., 2016; Vasek & Moore, 2012). Moreover, the impact of the recipient (Çetin et al., 2017; Li, Durumeric, et al., 2016) and the sender reputation (Çetin et al., 2016) on vulnerability remediation and malware cleanup have been analysed. Recently, Stock, Pellegrino, Li, Backes, & Rossow (2018) also investigated how users perceive vulnerability notifications. The end users in these studies are typically webmasters, server admins and network operators. A study by Zhang, Duan, Liu, and Yao (2017) was the first that analysed factors which could affect the effectiveness of vulnerability notifications from an ISP to its customers.

As far as we are aware, there is no prior work on the effectiveness of notifying retail customers and asking them to clean up malware infections on their machines. This makes the effectiveness of mechanisms analysed in previous studies difficult to compare to malware notifications and cleanup by consumers. Moreover, previous work has never focussed on IoT malware infections. These studies addressed malware infections for which appropriate cleanup solutions exist. However, for IoT malware such "one size fits all" cleanup solutions do not exist due to the heterogeneous nature of the IoT devices. In addition, it has never been investigated how users experience a malware notification, or more specifically an IoT malware notification, from their ISP. These knowledge gaps will be addressed in this research project.

1.3 RESEARCH OBJECTIVE AND QUESTIONS

Following from the practical and scientific problem statement, the objective of this research is to make a recommendation to a Dutch ISP on what notification mechanism to adopt by providing insight into: (1) how to increase the effectiveness of IoT malware notifications from an ISP to its customers in terms of IoT malware cleanup; and (2) how users perceive an IoT malware notification from their ISP. In this recommendation, it is described which notification channel and content an ISP should use to notify its IoT malware infected customers in order to promote end user cleanup efforts and to improve the reactions of customers. In this study, the desired customer reaction is that customers: (1) quickly react to the notification; (2) trust the notification; (3) understand the notification content; and (4) are satisfied with the notification mechanism.

To achieve the defined objective, the main research question that needs to be answered is:

What notification mechanism is the most effective in terms of both IoT malware cleanup and improving the reactions of customers?

To answer this question, the following sub-questions have been formulated. Each sub-question will be briefly discussed.

SQ 1. What are possible IoT malware notification channels for an ISP?

To be able to analyse which notification mechanism is the most effective in terms of cleanup and improving the reactions of customers, we need to map which communication channels can be used by ISPs to notify infected customers.

SQ. 2 Which factors influence the actionability of an IoT malware notification content?

This question concerns the other aspect of the notification mechanism: the notification content. To be able to compose a more actionable notification content, information about how to write an actionable notification content is needed.

SQ 3. What is KPN's current notification process for customers with infected IoT devices?

a. What IoT abuse data does KPN have?
b. What is the priority of IoT abuse?
c. What notification channel does KPN use to notify customers with infected IoT devices?
d. What notification message does KPN use to notify customers with infected IoT devices?
e. What are the strengths and weaknesses of KPN's current notification process?

A clear overview and evaluation of KPN's current notification process for customers with infected IoT devices is needed to understand which aspects could be improved. Moreover, we need this information to be able to create a feasible empirical research design to determine the influence of different notification mechanisms on IoT malware cleanup and the reactions of customers.

SQ 4. How can the effectiveness of notification mechanisms on IoT malware cleanup be measured quantifiably?

The primary interest and therefore the dependent variable in this research is IoT malware cleanup. Metrics have to be created in order to compare the effect of different notification mechanisms on cleanup in terms of concrete numbers.

SQ 5. What is the added value of quarantining in terms of IoT malware cleanup?

This sub-question investigates the relation between one of the independent variables (notification channel) and the dependent variable (IoT malware cleanup).

SQ 6. What is the influence of a more actionable walled garden notification content on IoT malware cleanup?

This sub-question investigates the relation between the other independent variable (notification content) and the dependent variable (IoT malware cleanup).

SQ 7. What are the reactions of the customers to different notification mechanisms for IoT malware infections?

a. What is the influence of the notification channel on the reactions of the customers?

b. What is the influence of a more actionable content on the reactions of the customers?

The mediating variable in this study, which explains the relation between the independent and dependent variable, is the reaction of the customers. This sub-question examines how the different notification mechanisms influence the reaction of the customer.

SQ 8. What is the impact of the customer's understanding of the notification on IoT malware cleanup?

Lastly, this sub-question addresses the relation between a specific aspect of the mediation variable (reactions of the customers) and the dependent variable (IoT malware cleanup). We investigate whether a better understanding of the notification content promotes the cleanup.

1.4 RESEARCH APPROACH

In order to answer the previously discussed sub-questions, the study is divided into 3 phases (see Figure 1): (I) preparation; (II) empirical analysis; and (III) evaluation of the results. The results of each phase are the input for the next phase.

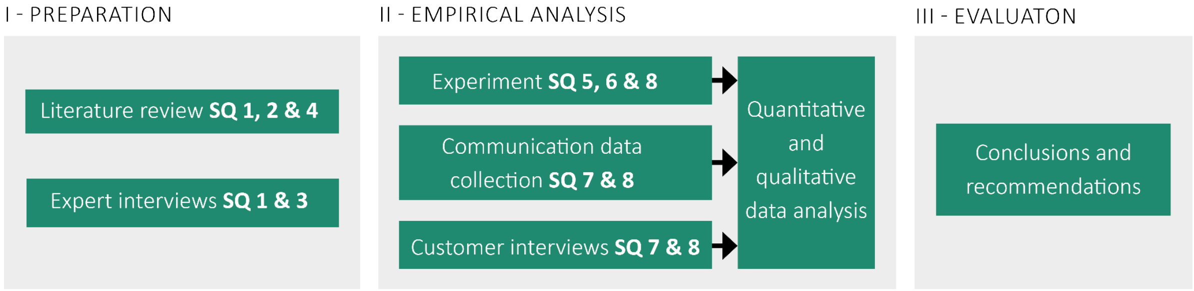

I - PREPARATION
II - EMPIRICAL ANALYSIS
III - EVALUATON
Literature review **SQ 1, 2 & 4**
Expert interviews **SQ 1 & 3**
Experiment **SQ 5, 6 & 8**
Communication data collection **SQ 7 & 8**
Customer interviews **SQ 7 & 8**
Quantitative and qualitative data analysis
Conclusions and recommendations

Figure 1 Research framework

1.4.1 PHASE I: PREPARATION

During the preparatory phase we conduct a literature review and interviews with abuse experts from KPN. By means of a literature review the possible notification channels for an ISP to notify its customers are identified (SQ 1). Furthermore, we obtain insights about composing a more actionable notification content (SQ 2) from studies about persuasive communication in the field of security warnings and malware notifications. Lastly, it is analysed which metrics previous studies have used to measure the effectiveness of notifications on cleanup quantifiably (SQ 4). Afterwards, we carry out the expert interviews. During these interviews with abuse experts from KPN's abuse department, the different notification channels found in the literature are evaluated (SQ 1). The notification channels are assessed based on the general strengths and weaknesses and the feasibility of implementation for KPN. In addition, KPN's current notification process for customers with infected IoT devices is mapped and evaluated (SQ 3). During these interviews it is also discussed how the abuse department prioritizes IoT abuse and which difficulties IoT abuse mitigation entails in general and compared to conventional abuse mitigation.

In short, the results of this phase are: (1) an overview of possible IoT malware notification channels for an ISP;

(2) an evaluation of KPN's current notification process for customers with infected IoT devices; and (3) a set of metrics to measure cleanup quantifiably.

1.4.2 PHASE II: EMPIRICAL ANALYSIS

The information obtained during the preparatory phase, is used as an input for the design of the second phase of the study. In the empirical phase, we examine the impact of the different notification mechanisms on cleanup and the reactions of customers. Moreover, the relation between the reactions of customers and cleanup is investigated. To be able to assess the influence of the notification channel and actionability of the walled garden notification content on IoT malware cleanup (SQ 5 and 6), an experiment is conducted at KPN. To this end, infected customers are randomly assigned to an experimental group, including a control group. By means of a statistical analysis we compare the cleanup of the different treatment groups. Additionally, to assess the reactions of the customers to the different notification mechanisms (SQ 7), interviews with the customers in the experiment are carried out. Next to that, the communication data between these customers and KPN is collected and analysed for this purpose. Lastly, we investigate the relation between the reactions of customers and cleanup (SQ 8). To this end the results of the experiment, the interviews and communication data analysis are combined in a statistical analysis.

In short, the results of this phase are insights into: (1) the influence of the notification channel on IoT malware cleanup, (2) the influence of a more actionable walled garden notification content on IoT malware cleanup, (3) the influence of the notification channel on the reactions of customers, (4) the influence of a more actionable walled garden notification content on the reactions of customers and (5) the relation between the reactions of customers and cleanup.

1.4.3 PHASE III: EVALUATION

In the last phase, we draw conclusions based on the results of this research project and reflect on the process and outcomes of this study. Additionally, based on the results of this study, recommendations for KPN's abuse department are formulated. Lastly, we provide our recommendations for future work.

1.5 SCIENTIFIC AND PRACTICAL RELEVANCE

From a scientific perspective, this research contributes to existing research in the field of abuse notifications. In previous studies that investigated the factors influencing the effectiveness of abuse notifications, the end users were typically webmasters, server admins and network operators. In contrast, this is the first study that provides insight into the context of an ISP's abuse department that sends abuse notifications to home users. Moreover, this is the first empirical study analysing which factors influence the effectiveness of IoT malware notifications in terms of end user cleanup. Additionally, the collaboration with KPN creates other opportunities. Firstly, the difficulties regarding IoT malware notifications to resource owners are mapped by expert interviews. Secondly, the customer reactions to IoT malware notifications are investigated and therefore we can also examine the relation between the reactions of customers and the IoT malware cleanup. Something that, to our knowledge, has not been researched before.

In practical terms, this research could contribute to mitigating cybercrime. A notification mechanism which is successful in getting customers to act against IoT abuse could decrease the number of infected machines. Secondly, when customers take action faster and better understand how to clean up the infection, the abuse department will spend less time on communicating with the customer and sending subsequent notifications. Thirdly, the notification mechanism that is the most effective in improving the level of customer satisfaction, could promote customer retention.

1.6 THESIS OUTLINE

The structure of this report is as follows (see Figure 2). First, in Chapter 2 the literature review is described. Background information on IoT security is provided and prior work on the effectiveness of abuse and vulnerability notifications is discussed. Furthermore we elaborate on studies into browser security warnings, as these studies focus on the end user behaviour as to security. In Chapter 3 the procedure of KPN's abuse team is described. Moreover, in this chapter the various data sources available for this study are described. Chapter 4 specifies the research design for this study. The empirical hypotheses are formulated and an explanation of the research methods and data collection strategy is given. Chapter 5 provides the results of the experiment. The result of the analysis of the customer reactions to the different notification mechanisms, based on the communication data and customer interviews, are described in Chapter 6. Then, in Chapter 7, we describe the results of the analysis to investigate the relation between the understanding of the customer and cleanup. Lastly, in Chapter 8 the conclusions of this research project are described. Here, we will reflect on the outcomes of this study. Additionally, this chapter includes recommendations for KPN about what notification mechanism to adopt and recommendations for future research.

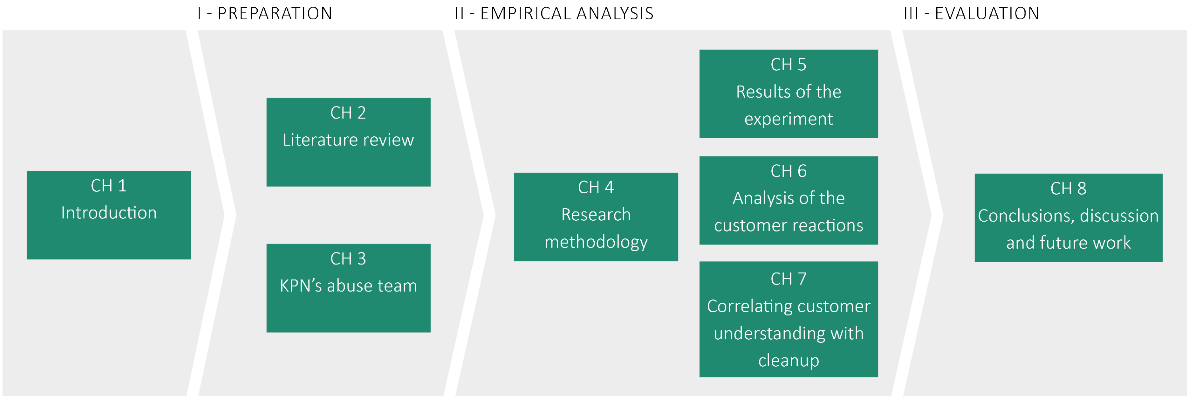


Figure 2 Thesis outline

2 LITERATURE REVIEW

In the previous chapter the research objective and research questions of this study were formulated. The goal of this chapter is to develop a conceptual model to analyse the problem, as presented in the previous chapter, systematically. Furthermore, this chapter provides an answer to the first sub-question: What are possible IoT malware notification channels for an ISP? Furthermore, it is examined which factors influence the actionability of an IoT malware notification content (SQ 2) and how previous studies have measured notification effectiveness quantifiably (SQ 4).

First, in Section 2.1, a research context is provided. The security challenges for the IoT are described and it is explained why IoT systems are at a higher security risk compared to conventional computing systems. Second, we identify the stakeholders involved in IoT abuse and elaborate on the role of ISPs in the control of malware. In Section 2.4, possible notification channels for ISPs are discussed. Next, we elaborate on previous studies into the effectiveness of notifications to affected parties. These studies investigated the influence of various factors on cleanup and remediation. Additionally, we briefly discuss studies into browser security warnings, as these studies focus on the end user behaviour as to security. In Section 2.5, we identify factors that influence the actionability of an IoT walled garden notification content by combining the communication-human information processing (C-HIP) model and the protection motivation theory (PMT). Based on the insights obtained from the literature, a conceptual model is developed and presented in Section 2. Lastly, an answer to the sub-questions is provided in the conclusions in Section 2.7.

2.1 SECURITY CHALLENGES FOR THE IOT

IoT devices present a variety of security risks that could be exploited to cause harm. Firstly, a lack of security could enable intruders to access and misuse personal information, which is collected and transmitted to or from the device. Secondly, vulnerabilities in a device might facilitate attacks on the network to which the device is connected or on other systems, like DDoS attacks. Thirdly, the exploitation of vulnerabilities might create safety risks. For example, when an insulin pump is hacked, the intruder is able to change settings so it no longer delivers medicine (Federal Trade Commission, 2015).

Even though these security risks in the context of information systems are not new, the IoT creates new and unique challenges. Compared to conventional computing systems, such as laptops and smartphones, IoT systems are at a higher security risk for several reasons. There is not a 'one-size-fits-all' solution for securing IoT. At one end of the spectrum, there are small appliances like IP cameras and smart lights, while on the other end there are larger devices. Each type needs its own approach. Moreover, many different communication media and protocols and platforms are used. Besides, many devices are not designed to be connected to the Internet, because in their original design they were intended to be stand-alone. In addition, IoT systems do not have well-defined perimeters and continuously change due to the mobility of the device and user (Bertino, 2016).

Besides the above-mentioned reasons, there is a market failure at work. Many of the devices are inexpensive and essentially disposable (Federal Trade Commission, 2015). Often, they are designed and built offshore, and then rebranded and resold. Low costs have priority. Manufacturers choose a chip based on price and features. The teams building the devices do not have the security expertise one has come to expect from the major computer and smartphone manufacturers, because the market does not stand for the additional costs that would require.

Even though the devices are often used for years and decades, when a vulnerability is found after manufacturing, it may be difficult or impossible to update the software or apply a patch (Schneier, 2014, 2016). An additional market failure is that neither the seller nor the buyer has an incentive to invest in security. The owners of the devices want a functioning device for a good price. When a device gets infected and becomes part of a botnet, the owner might not notice. The machine keeps working fine and there is no integrated security software that could detect it (Barcena & Wueest, 2015). The seller does not care either. After selling a product, the motivation to maintain the software is limited, because the manufacturer already received its money. This holds especially for products with a low price. After all, they want to make a profit. Since the insecurity primarily affects other people, there seems to be no market solution (Jacobs, 2016; Schneier, 2016).

Consequently, many IoT systems lack even basic security. A study by Hewlett-Packard Development Company (2014) showed that six out of ten of the most popular connected devices had common vulnerabilities, which are listed in Table 1, and 70% did not encrypt communications over the Internet.

Table 1 Common Internet of Things vulnerabilities. From Bertino and Islam (2017, p. 78)

Vulnerability	**Examples**
Insecure web/mobile/cloud interface	Inability to change default usernames and passwords; weak passwords; lack of robust password recovery mechanisms; exposed credentials; lack of account lockout; susceptibility to cross-site scripting, cross-site request forgery, and/or SQL injection
Insufficient authentication/ authorization	Privilege escalation; lack of granular access control
Insecure network services	Vulnerability to denial-of-service, buffer overflow, and fuzzing attacks; network ports or services unnecessarily exposed to the Internet
Lack of transport encryption/ integrity verification	Transmission of unencrypted data and credentials
Privacy concerns	Collection of unnecessary user data; exposed personal data; insufficient controls on who has access to user data; sensitive data not de-identified or anonymized; lack of data retention limits
Insufficient security configurability	Lack of granular permissions model; inability to separate administrators from users; weak password policies; no security logging; lack of data encryption options; no user notification of security events
Insecure software/firmware	Lack of secure update mechanism; update files not encrypted; update files not verified before upload; insecure update server; hardcoded credentials
Poor physical security	Device easy to disassemble; access to software via USB ports; removable storage media

2.1.1 IOT MALWARE

With the growth of the IoT market, "malware targeting the Internet of Things (IoT) has come of age and the number of attack groups focusing on IoT has multiplied over the past year" (Symantec, 2016). Attackers tend to have low interest in the owner of the compromised device, the majority wants to hijack the device to add it to a botnet. "IoT botnets are cheap, easy to construct, and lack significant functionality aside from DDoS attacks" (Scott & Spaniel, 2016, p. 10). The malware distribution consists in most cases of a scan for random IP addresses with open Telnet or secure shell (SSH) ports, followed by a brute-force attempt to login with frequently used credentials. Telnet and SSH are both network protocols which enable remotely logging in and controlling a system. The main difference between the two is that SSH has encryption, while Telnet has not. In many cases, using the default username and password to login suffices. This might be as simple as ‘root’ and ‘admin’.

The IoT devices run on a variety of Central Processing Unit (CPU) architectures. Therefore, the IoT malware either can try to randomly download bot executables for multiple architectures and run them one after another, until one is successful or a module of the malware first checks for the existing devices’ platform and only downloads the correct bot binary. When the bot binary is executed, a connection to a C&C server will be established (Symantec, 2016). This enables the control server to perform different actions remotely. For example, at a certain moment, the bots can be used to launch a DDoS attack by simultaneously directing traffic from parallel bots against a single victim (Bertino & Islam, 2017).

Malware families

On October 21, 2016, thousands infected IoT devices were responsible for a DDoS attack against DNS provider Dyn. Most of these IoT devices were enslaved by Mirai, a self-spreading malware for IoT devices. Mirai is however not the first IoT botnet to make headlines. BASHLITE botnets, a predecessor to Mirai, was also responsible for enslaving over 1 million devices (Gallagher, 2016; Scott & Spaniel, 2016; Symantec, 2016). In this subsection the characteristics of these two most prevalent malware families are identified.

BASHLITE, which is also known as Lizkebab, Torlus and gafgyt, infects Linux systems to perform DDoS attacks. The original version of BASHLITE originates from 2014. After the leakage of the source code in 2015, it has been adapted into different variants. Level 3 Threat Research Labs estimates that almost 96 percent of the identifiable devices participating in the botnets were IoT devices, of which 95 percent were IP cameras and DVR units, 4 percent were home routers and less than 1 percent were Linux servers (Level 3 communications, n.d.). The majority of the infected devices were located in Taiwan, Brazil and Columbia. DVRs are valuable bots, because these devices “are configured with open Telnet and other web interfaces, often rely on default credentials, and are able to process high bandwidth, as is required to stream a video” (Scott & Spaniel, 2016, p. 27). To find vulnerable devices to infect, the malware conducts two scans. Firstly, the bots are used to port scan IP ranges for Telnet servers. Subsequently the bot is instructed to perform a brute-force attack, using a build-in dictionary of common usernames and passwords, in order to access and infect the device. The second attack vector employs external scanners to detect and infect vulnerable devices. The exact capabilities of the BASHLITE DDoS attacks differ between variants. Most are UDP and TCP floods, though it does support a less used feature to spoof source addresses and some variants support HTTP attacks (Scott & Spaniel, 2016).

Mirai was first found by a white hat malware research group in August 2016. Shortly after, on September 30, 2016, a script kiddie using the name Anna-senpai posted the botnet and C&C server code on hacker forums, which gave researchers insight into how Mirai operated. On the down side, it enabled new threat actors to adopt the malware and adapt its functionality. Basically, everyone who has access to Internet connected servers and can compile the code is able to build a botnet. Mirai, is the successor of BASHLITE and it works, to a great extent, the same way. One of the enhancements that Mirai implemented, is that the communication with its C&C servers is now encrypted, so the traffic is less visible to firewalls and other security systems. Furthermore, the malware contains scripts to kill any other processes that run SSH, Telnet or HTTP ports and remove competing infections or malware. When a device is infected with Mirai, port 48101 is used as a sign to prevent wasting scanning activity and to prevent multiple Mirai infections (Gallagher, 2016; Scott & Spaniel, 2016). The Mirai threat seems to stabilize. However, Mirai is expected to be the first of a category of botnets that exploit IoT devices and systems, as history shows that the deployment of defences against a security threat is soon followed by new attack vectors (Bertino & Islam, 2017).

2.2 STAKEHOLDERS

IoT abuse involves many different stakeholders. In Table 2, the different stakeholders are classified into categories, as proposed by Sheng, Kumaraguru, Acquisti, Cranor, & Hong (2009):

Primary victims: they suffer direct losses from IoT infections. The end user's device works slower or stops working. Companies can be victim of a DDoS attack. Furthermore, intruders could access and misuse companies' and device owners' sensitive data, which is collected and transmitted to or from the device.

Infrastructure providers: they have technical capabilities to mitigate the problem. ISPs and hosting providers are able to detect malicious traffic in their network and to act upon subsequently. IoT vendors could tackle the vulnerability issues by, for example, bundling security software on their machines.

Defenders: the goal of these public protectors is to protect the society at large and mitigate illegal activities. AbuseHub and Shadowserver aim to bring different stakeholders together to fight more effectively against abuse.

Table 2 IoT abuse stakeholders

Categories	**Roles**
Infected IoT device owners	Primary victims
Companies	
ISPs	Infrastructure providers
Hosting providers	
IoT vendors	
CERTs	Defenders
Academia	
Law enforcement	
AbuseHub	
Shadowserver	

2.2.1 THE ROLE OF ISPS

As this research is conducted in collaboration with KPN, the research focusses on the role of ISPs as defenders. An emerging consensus is that ISPs play an important role in controlling the spread of malware infecting conventional computers, as they are best positioned to intervene. "The term ISP is used to cover a variety of businesses, typically ISPs are defined as providers that offer individuals and organisations access to the Internet" (M. J. G. van Eeten & Bauer, 2008, p. 26). In the role of Internet access providers, ISPs are able to detect malicious traffic in their network and to subsequently act upon. ISPs are able to detect infected devices by means of scanning for outgoing connections to known C&C servers used by botnet operators. A crucial advantage of ISPs is that they can link the IP address to customer details and are therefore able to notify and assist customers (Hofmeyr et al., 2013).

Pijpker and Vranken (2016) studied how ISPs are involved in botnet mitigation in the Netherlands. Therefore, they created a reference model, which summarises measures for botnet mitigation from scientific literature that ISPs can take. This reference model is structured according to the anti-botnet lifecycle. The Online Trust Alliance (OTA) identified the five elements of the anti-botnet lifecycle, as shown in Figure 3:

- Prevention: proactive activities initiated by an ISP that can reduce the vulnerability of a user's device.
- Detection: actions/activities aimed at identifying threats on a device or network.
- Notification: action/activities conducted by an ISP to inform a customer.
- Remediation: actions/activities initiated by an ISP to remove malicious software from a compromised device.
- Recovery: actions/activities supported by an ISP to resolve the impact of an attack.

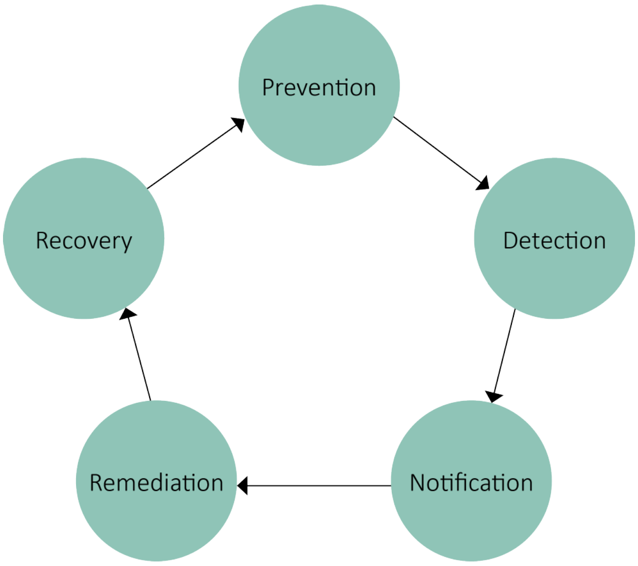


Figure 3 Anti-botnet life cycle. From Online Trust Alliance (2012)

The study showed that Dutch ISPs spend most effort on prevention. Firstly, because this is the most efficient and effective approach and secondly, because Dutch ISPs are obligated to do so according to the Telecommunication act. According to this act, ISPs are required to take technical and organisational measures to protect their customers against cybercrime. ISPs are also required to inform their customers about the risks associated with the use of the offered Internet services and what the customer themselves can do to reduce these risks ("Telecommunicatiewet," 1998). This legal framework mainly addresses preventive measures. The act does not prescribe what ISPs should do in case of infections in their network. ISPs are not obliged to take actions against compromised machines, such actions are voluntary. Moreover, it is concluded that ISPs are currently well informed about botnet threats, as large-scale detection becomes more feasible due to systems such as AbuseHUB. Nevertheless, Pijpker and Vranken (2016) state that the information sharing with customers could be improved.

ISP's Incentives

As described before, no entity is responsible for acting on abuse data. Incentives determine why ISPs receive abuse reports and then take action on them. The incentives for ISPs to implement security countermeasures are weak, as much of the harm caused by the infected devices affects other people, while the cost of notification and clean-up would fall largely on the ISP. On the other hand, there are costs associated with inaction. Infected customers might contact their ISP for help, which raises the ISP's cost of customer support. Furthermore, public reputation could be an incentive for the ISP. An ISP that behaves ‘responsibly’ by remediating compromises might be able to increase their share of the overall market. However, there is mixed evidence whether this holds in practice (Hofmeyr et al., 2013; Jhaveri, Cetin, Gañán, Moore, & Eeten, 2017).

2.3 ISP’S NOTIFICATION CHANNELS

In this section we describe the possible malware notification channels for an ISP discussed in the literature (SQ 1). Livingood, Mody, & O’Reirdan (2012) published a document with recommendations on how ISPs can use various notification channels. Here, the process of notification to internet users that may have a bot-related problem will be discussed. Moreover, the complications of certain methods will be described.

After detection of a bot, or strong likelihood of a bot, the internet user should be informed that they may have a bot-related problem. This message could also include information on remediation tools that can be applied to solve the problem of the infection. The ISP has to decide on the most appropriate method or methods to notify their customers taking into account a range of factors including “the technical capabilities of the ISP, the technical attributes of its network, financial considerations, available server resources, available organizational resources, the number of likely infected hosts detected at any given time, and the severity of any possible threats” (Livingood et al., 2012, p. 11). Such notification methods include one or more of the following methods:

Email notification: This method is commonly used by ISPs. However, a major drawback is that it is not assured that the email is read in a timely manner, if read at all. Firstly, a user might use a different primary email address than the one provided to the ISP. Secondly, the user’s email server could classify the email as spam, causing it to be deleted or filed into a folder which is read irregularly. Furthermore, bot masters could impersonate the ISP or a trusted sender and send fraudulent emails to the users. Lastly, when the user’s email credentials are compromised, the hacker or a bot could access the email account and delete the email before it is read.

Telephone call notification: This is an effective means of communication in high risk situations. Nevertheless, due to the high cost of making a large number of calls, it may be unfeasible. Moreover, clients may not answer the call and if they do, interpret it as a telemarketing call or lack the technical expertise to understand or be able to deal with the threat.

Postal mail notification: This is indicated to be the least effective and popular means of communication, due to the preparation and delivery time, and the cost of printing, paper and postage.

Instant message (IM) notification: This provides the ISP a simple means to communicate with the customer. The cost-effectiveness is a major advantage. When a user subscribes to the ISP’s IM service, the user can be notified automatically or by a manual process involving the ISP’s support staff. There are, however, several drawbacks. Firstly, not every user uses IM or the user might not want to share its IM identity with the ISP. In that case, an alternative means has to be used. Secondly, a message might be interpreted as spam and therefore ignored. Furthermore, the client may not be signed onto the IM system when notification is attempted. Lastly, there could be a privacy concern on the part of the users, when a message has to be transmitted over a third-party network and/or IM service. Therefore, the notification should be discrete and not include any personal identifiable information.

Short message service (SMS) notification: This method allows the ISP to send a brief description of the problem to the user’s mobile phone. Therefore, the client has to register his mobile number and grant permission to be contacted via this means. The major advantage is that users are likely to read the message. However, if they are not near the device, they may not act on the notification immediately. Moreover, an additional means of notification is needed, as not all the necessary information can be conveyed in one message. Another drawback is the cost associated with it and the fact that the client might change its number without notifying the ISP. Furthermore, the user might ignore the message, because it is interpreted as spam. Also, not every user uses SMS and some might not be willing to share its mobile number. Even if the user provides his number, the mobile phone may not be powered on when the notification is attempted. Lastly, there could be a privacy concern on the part of the users, when a message has to be transmitted over a third-party network. Therefore, the notification should be discrete and not include any personal identifiable information in the notification itself.

Walled garden notification: An ISP could place the customer in a so-called 'walled garden'. "A 'walled garden' refers to an environment that controls the information and services that a subscriber is allowed to utilize and what network access permissions are granted" (Livingood et al., 2012, p. 13). This method is effective, as the user is notified and simultaneously the communication between the bot and C&C server is blocked. The Messaging, Malware and Mobile Anti-Abuse Working Group published best practices related to the implementation of a walled garden (Messaging Anti-Abuse Working Group, 2007). According to them, ISPs are required to use a walled garden as a more proactive measure in an effort to protect their network, since subscriber-originating network abuse increases. In this document a distinction is made between which characteristics the walled garden must, should and may have. According to the authors, the walled garden system should manage all outbound SMTP to a quarantine area, to a honeypot MTA (Message Transfer Agent) or should block altogether during this process. In addition, the walled garden system should allow instant escape based on trust. Lastly, the walled garden system may redirect HTTP to a quarantine website, may redirect botnet C&C traffic to a honey network for analysis and may provide exit if certain security software is downloaded and installed.

Since each method has its own limitations, Livingood et al. (2012) recommend the use of multiple notification methods. Moreover, it is emphasised that a notification is time-sensitive. If the user does not receive or view the notification in a timely manner, a bot could have already caused harm. Therefore, Livingood et al. (2012) recommend ISPs to establish a preferred means of notification when the subscriber first signs up for the service. It is recommended that the client chooses the method on an opt-in basis and the client should not be allowed to opt-out of notification entirely. ISPs can also decide to notify other stakeholders, such as peer ISPs or governmental agencies that aggregate threat data, about infections. Livingood et al. (2012) point out that an ISP needs approval from a client when sharing personal identifiable information with third parties. This should be done on an opt-in basis.

Lastly, it should be noted that the effectiveness of the notification heavily depends on the expertise of the end user and the wording of the notification. The latter issue is addressed in next section.

2.4 EFFECTIVENESS OF NOTIFICATIONS

As far as we are aware, there is no prior work on the effectiveness of notifying customers in an ISP's network and asking them to clean up malware infections on their devices. In this section, we describe three related areas of work. Previous studies into abuse and vulnerability notifications have studied similar mechanisms. In Section 2.4.1.1, we explain which metrics were used in previous studies to measure the effectiveness of abuse notification quantifiably (SQ 3). Prior work had typically a different type of end user, namely webmasters, server admins and network operators, not home users. This makes the effectiveness of those mechanisms difficult to compare to malware notifications and cleanup by consumers. Another area of related work concerns security warnings (Section 2.4.3). These notifications are aimed at preventing compromise, trying to steer the user back to safety. In contrast, we study a notification mechanism where the action is not avoiding a danger, but dealing with the damage that already has occurred. Also, the action required of the user in case of compromise is not a single decision for or against a potentially dangerous action, but the execution of a rather complicated set of steps to resolve the incident that has already manifested itself.

2.4.1 ABUSE NOTIFICATIONS

Because little is known about the factors that drive higher response rates to abuse reports, a few researchers have recently investigated how abuse notifications can promote cleanup.

Jhaveri et al. (2017) constructed a model of the abuse reporting infrastructure. The research, described in this report, concerns intermediary remediation. The ISP, referred to as intermediary (INT), receives an abuse report from the abuse notifier (AN) and decides whether or not to send a notification to the infected resource owner (RO). Building on this model of the abuse reporting infrastructure, the authors created a list of factors that might affect the success or failure of cleanup efforts. Table 3 shows a selection of the attributes relevant for this study. It should be noted that in this research project both the recipient (RO), sender (INT) and abuse type are constant. Seeing that the channel and content are variable, the influence of these factors can be investigated in this research project.

Table 3 Attributes of the abuse reporting infrastructure that may influence the effectiveness of cleanup efforts

Attribute	Description	Possible values
Channel	How is information shared?	Unsolicited Email, Phone Call, SMS, Walled garden, Public Post
Recipient	Who receives the notification?	Abuse notifier (AN), Intermediary (INT), Resource owner (RO)
Content	How is the abuse report transformed?	Legalese, None, Education, Simplification, Explanation, Threats
Type	What is the type of abuse?	Malware, Spam, Phishing, etc.
Sender reputation	How well known is the organization and what's the credibility?	High, Medium, Low, Anonymous

A range of studies have investigated if and how abuse notifications impact the cleanup of compromised websites. Notifications can be send to the affected owners of the site or to their hosting provider. In an observational study, Li et al. (2016) used data of over 700,000 infected websites that were detected by Google Safe Browsing and Search Quality. The researchers found that direct notifications to webmasters via Google Webmaster Console increased the likelihood of cleanup by over 50% and, furthermore, that the infection lifetime decreased by at least 62%. Vasek and Moore (2012) conducted an experimental study on malicious URLs submitted to the StopBadware community feeds to investigate the impact of abuse reports and how the level of detail in the reports influenced the cleanup rate. They found that only abuse reports with detailed information result in higher cleanup rates of compromised websites compared to those not receiving a notice, 62% compared to 45% after 16 days. Notably, they found that sending a minimal report is roughly as effective as not sending a notification at all. Çetin, Jhaveri, Gañán, van Eeten, and Moore (2016) reaffirmed the finding that detailed notices work. They found that around half of all compromised sites got cleaned up after a notification to the hosting provider. The authors did not find a statistically significant difference between the abuse notifications of senders with varying levels of reputation. Canali, Balzarotti, and Francillon (2013) studied how hosting provider handle abuse notifications. They created vulnerable webservers on 22 hosting services and ran five different attacks on them that simulated infections and then notified the providers about these attacks. They observed that only 36% reacted to the abuse notifications.

Similarly, Nappa, Zubair Rafique, and Caballero (2013) issued abuse reports for 19 long-lived exploit servers. However, only 7 providers took action towards cleaning up the malicious servers.

Measuring effectiveness of abuse notifications

In previous experimental studies (Çetin et al., 2016; Vasek & Moore, 2012), quantitative metrics were used to measure the impact of different notification mechanisms on the cleanup of compromised websites: (1) cleanup rate; and (2) median time to cleanup across the various treatment groups relative to the control group. The cleanup rate is defined as the percentage of notified parties that is clean at the end of the investigation period. The second metric is the median number of days required to clean up those sites that were successfully remediated.

2.4.2 VULNERABILITY NOTIFICATIONS

How security notifications can expedite vulnerability remediation has recently been a subject of several studies. For example, Durumeric et al. (2014) discovered and notified system owners vulnerable to the Heartbleed vulnerability. The study revealed that the rate of patching for the notified groups was 47% higher than the control group, 39.5% versus 26.8%. Similarly, Kührer, Hupperich, Rossow, and Holz (2014) issued notifications to administrators of vulnerable Network Time Protocol (NTP) servers, in collaboration with CERTs, clearinghouses and afflicted vendors. Though their study lacks a control group to assess the impact of the campaign itself, they found that 92% of NTP server were remediated in 13 weeks. Li et al. (2016) investigated which factors have the greatest impact on vulnerability remediation rates. They found notifications addressed directly to the vulnerable resource owners to be more effective than those sent to national CERTs and US-CERT. In addition, the study showed that notifications with detailed information increased the remediation rate compared to terse notifications. However, the majority of contacts did not take action and when they did, remediation was often only partial. A study by Stock, Pellegrino, Rossow, Johns, and Backes (2016) into the effectiveness of large-scale vulnerability notification campaigns for vulnerable Web servers also observed the challenge of reaching an appropriate point of contact. They found that only around 6% of the affected parties could be reached. Of that small fraction, around 40% were remediated upon notification. Due to the poor deliverability of email-based notifications, Çetin et al. (2017) also proposed to move away from email as the main notification medium and search for other notification channels to drive remediation rates. Stock, Pellegrino, Li, Backes, & Rossow (2018) later tested the effectiveness of other channels such as postal mail, social media, and phone and concluded that the slightly higher remediation rates of these channels do not justify the additional work and costs. Recently, Zhang et al. (2017) focussed on the effectiveness of telephone, email and instant message (IM) notifications in the scope of an ISP, whose main customers are educational institutions instead of home users like in our study. They conclude that IM is the most appropriate notification mode for such an ISP.

2.4.3 SECURITY WARNINGS

More research into user behaviour regarding security issues has been conducted in the field of security warnings. Neupane, Saxena, Kuruvilla, Georgescu, and Kana (2014) used neuropsychological measures to investigate security behaviour. The user's neural activity in phishing detection and malware warnings showed that users are actively engaged in these security tasks. Nevertheless, users often ignore warning. A large body of literature focused on why users ignore warnings and how this could be avoided. Almuhimedi, Felt, Reeder, and Consolvo (2014) studied user reactions to Google Chrome malware warnings. Up to half of the warnings were ignored, under certain circumstances. Some users confused the malware warnings with SSL warnings. Sunshine, Egelman, Almuhimedi, Atri, and Cranor (2009) examined users' reactions to existing and newly designed SSL warnings. The authors suggested that, although existing SSL warnings can be improved, minimizing the use of SSL warnings by blocking users from making insecure connections proves to be more effective. Finally, Mathur, Engel, Sobti, Chang, and Chetty (2016) concluded that one of the reasons why users ignore software updates is that updates regularly interrupt users who often lack sufficient basic information to decide whether or not to update. A closely related topic is the problem of habituation of users to ignore warnings after they have learned that this does not seem to cause any harm (Egelman, Cranor, & Hong, 2008; Kim & Wogalter, 2009). Bravo-Lillo et al. (2013) and Bravo-Lillo, Cranor, Komanduri, Schechter, and Sleeper (2014) tested the effectiveness of user-interface modifications to draw users' attention to the most important information required for decisions.

Multiple studies have demonstrated that end users have difficulty securing their computers, either because of lack of knowledge or ignoring security advice that is hard to understand. In a study conducted by Wash, Rader, Vaniea, and Rizor (2014) on how users perceive automated software updates, the authors observed that the majority of users do not correctly understand the automatic update settings on their computer and cannot manage software updates the way they intend to. A study by Krol et al. (2012) also showed that misunderstanding is a reason for ignoring security warnings. This mismatch between intention and behaviour frequently led to computers being more or less secure than intended. Fagan, Maifi, & Khan (2016) studied user motivations regarding their decisions on following common security advice (i.e., update software, use password manager, change passwords) and concluded that the majority of users follow the usability/security trade-off. Finally, Forget et al. (2016) developed a Security Behaviour Observatory to collect data on users' behaviour and their machine configurations. Their findings highlighted the importance of content, presentation, and functionality of security notifications provided to users who have different expertise, expectations, and computer security engagement.

2.5 MALWARE NOTIFICATION CONTENT

The aim of this section is to find an answer to sub-question 2: Which factors influence the actionability of an IoT malware walled garden notification content? The content of the notification is an important aspect to get infected customers to take the desired cleanup actions. As pointed out in the previous section, even if people actually read a warning message, they may reject it based on its content. Therefore, it is critical to compose a message that increases attention and understanding regarding the security issue and presents ways to cope with the issue.

In the past years, a range of best practices and guidelines for ISPs around the content of malware notifications, have been published by leading industry associations (Livingood et al., 2012; Messaging Anti-Abuse Working Group, 2007; Online Trust Alliance, 2012). In these articles it is described how to communicate a technical message to a wide variety of users with the objective to encourage the reader to take action. However, to the best of our knowledge, there is no prior research into the content of malware notifications to end users based on communication and persuasion theory. Therefore, we discuss a closely related area of work: the design of security warnings. We briefly deliberate on persuasion and communication theory in the field of security warnings. Based on the theory and findings from the literature, the we formulate guidelines for writing an actionable malware notification message.

A difference between security warnings and malware notifications is that a warning is mostly meant to prevent compromise, whereas a malware notification deals with the damage that has already occurred. In addition, the action required in case of an IoT malware infection is not a single decision for or against a potentially dangerous action, but the execution of a rather complicated set of steps to solve the issue. Nevertheless, the aim of both a malware notification and security warning is to steer the reader towards complying with the requested procedure in order to prevent harm.

2.5.1 C-HIP MODEL

Because of the growth of research in the field of warnings, in 1999, the communication-human information processing (C-HIP) model was developed. This framework is useful for organizing and structuring the findings in warning research and it can serve as a tool to help determine why a warning fails to be effective (Wogalter, 2006). The model, which is represented in Figure 4, shows that to communicate a message, you have to consider the entity delivering the message (source), the channel and multiple aspects of the receiver. When the message is presented to the receiver, there are successive stages that could affect the ultimate behaviour of the receiver: (1) gaining and retaining attention; (2) comprehension; (3) attitudes and beliefs; and (4) motivation. These aspects can be influenced by the content of the notification. In the following, we will briefly discuss these factors. It should be noted that, in addition to the notification itself, all factors can always be affected by the receiver's personal characteristics and environmental factors (Wogalter, 2006). These variables are, however, not taken into account in this section.

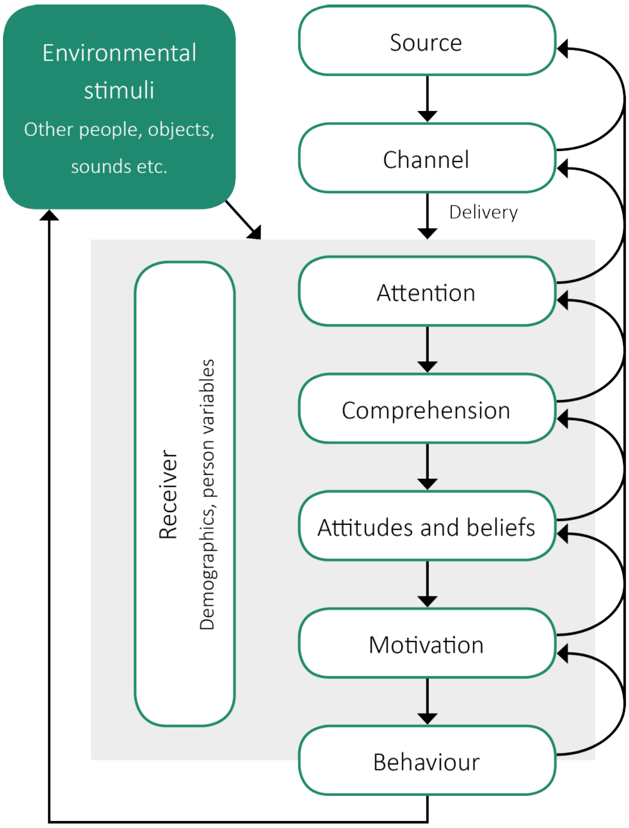


Figure 4 C-HIP model. Adapted from Wogalter (2006)

Attention

Following the C-HIP model, an effective warning must attract attention. Even though, paying attention does not have a direct effect on compliance. Attention is necessary to affect attitudes and beliefs of the receiver. Without this attention, the message will have no effect. Wogalter and Laughery (1996) described design factors that influence how well warnings attract attention. First of all, a colour which is distinctive in its environment is an important attribute that can facilitate attention attraction. A study on webbrowser warnings by Egelman & Schechter (2013) showed that altering text and colour significantly increased user's attention. Moreover, symbols such as an exclamation mark and signal words like "Danger" can be useful to draw attention. Lastly, formatting can improve the attention maintenance. People prefer to read a message in list format as opposed to continuous text (Desaulniers, 1987).

In short, guidelines to attract and maintain the user's attention include, the use of: (1) colour such as red, which signals danger; (2) symbols such as an exclamation point; (3) signal words such as "Danger"; and (4) a list format to present the information.

Comprehension

After capturing the user's attention, the next step is comprehension. Given that users have varying levels of technical expertise, they behave differently based on their level of comprehension (Sunshine et al., 2009). Therefore, the messages should be crafted towards the least skilled reader, to ensure that all readers understand the message (Wogalter & Laughery, 1996). To this end technical jargon should be avoided where possible and consistent terminology should be used. Technical terms must be replaced by phrases or expressions that might be better understood by the user (Bauer, Bravo-Lillo, Cranor, & Fragkaki, 2013; Livingood et al., 2012; Messaging Anti-Abuse Working Group, 2007; Modic & Anderson, 2014; Online Trust Alliance, 2012). Moreover, the message should be presented in the user's primary language (Online Trust Alliance, 2012).

In short, guidelines to improve the reader's comprehension include: (1) avoiding technical terms; (2) using consistent terminology; and (3) presenting the message in the reader's primary language.

Attitudes and beliefs, motivation and behaviour

Once a user pays attention and comprehends the message content, their attitudes and beliefs can be changed. "Beliefs and attitudes refer to an individual's knowledge that is accepted as true, although some of it may actually be untrue" (Wogalter, 2006, p. 57). According to the C-HIP model, the user must be motivated by attitudes and beliefs to change their behaviour. Different theories exist regarding the effect of attitudes and beliefs on ultimate behaviour. Here we describe the Protection Motivation Theory (PMT) (Rogers, 1983).

Protection Motivation Theory

The PMT was developed to explain the effects of fear appeals on attitude change. The theory has conventionally been applied in personal health contexts. More recently, the application of PMT has extended to research focussing on information security (Silic, Barlow, & Ormond, 2015). According to PMT, to change user's behaviour, four conditions must be met: (1) perceived vulnerability, (2) perceived severity, (3) perceived response efficacy and (4) perceived self-efficacy. These conditions can be sub-divided into threat appraisal and coping appraisal. The threat appraisal consists of both the vulnerability and the severity of the situation. The perceived vulnerability is the user's perceived probability that one will experience harm. Severity refers to the degree of harm from not complying with the recommended action. The coping appraisal consists of both efficacy and self-efficacy. The perceived response efficacy is the user's belief that the recommended action is effective in removing or preventing possible harm. The self-efficacy is the belief that one can successfully perform the recommended action.

These conditions for protection motivation can be translated to guidelines for the notification content. Regarding the threat appraisal, clear and non-technical communication is required regarding potential negative outcomes if not complying with the intended course of action (Modic & Anderson, 2014; Seiders, Flynn, Berry, & Haws, 2015). The message should clearly describe why the customer was notified. It may include an explanation of what bots are and the threats that they pose (Livingood et al., 2012; Messaging Anti-Abuse Working Group, 2007; Online Trust Alliance, 2012). Secondly, as to the coping appraisal, the message should include easy to understand steps that the customers must take in order to clean up the infected device and prevent future infections, including links to online tools and security updates. In addition, the message should provide support or abuse contact information in order to increase the user's belief in one's ability to execute the recommended courses of action successfully (Livingood et al., 2012; Messaging Anti-Abuse Working Group, 2007; Online Trust Alliance, 2012).

In short, guidelines to motivate the receiver of the notification to take the requested actions include: (1) clearly specifying the underlying risk; (2) clearly describing the potential negative outcomes of not complying with the intended course of action; (3) including easy to understand steps; and (4) providing support or abuse contact information.

Additional guidelines

Besides the guidelines following from the C-HIP, additional guidelines were presented in the articles describing best practices for remediating bots through end-user notification. In the following, we briefly discuss these guidelines. Firstly, the notification should be distinguishable from fraudulent notifications, as some customers tend to ignore warnings because criminals often use pop ups and redirects. To this end, the ability to verify the authenticity of the notification should be provided (Messaging Anti-Abuse Working Group, 2007; Online Trust Alliance, 2012). For example, by including the customer number or amount of the last invoice. Secondly, the ISP should attempt to identify the specific device that is infected. If the ISP is unable to identify the infected device, this should be clearly communicated to the customer if. Since, in that case, the cleanup advice is generic (Livingood et al., 2012).

2.5.2 SUMMARY GUIDELINES

The guidelines presented in the previous sections, applicable to abuse notifications from an ISP to resource owners, are summarized in Table 4. This table is supplemented with general guidelines for creating easy-to-understand messages to a wide variety of audiences (CDC, 2010).

Table 4 Guidelines and principles applicable to writing abuse notification messages for resource owners

Category	Guideline	Source
Attract attention	a) use colour such as red	[8]
	b) use symbols such as an exclamation point	[8]
	c) use signal words such as “Danger”	[8]
Describe the risk comprehensively	a) Clearly specify the underlying risk, why the notification is sent	[1,3,4,5,7]
	b) Clearly describe the consequences of not complying with the intended course of action	[1,3,7]
Be concise and accurate	a) Brief, remove redundant text	[1,2]
	b) Avoid technical jargon	[1,2,3,4,5,6]
	c) Use an active voice and short words and sentences	[2]
	d) Identify action steps or desired behaviours for the audience I. Steps to remove malware II. Steps to prevent future infections	[1,2,3,4,6]
	e) Avoid ambiguous terms	[1]
	f) Be polite, supportive, and encouraging	[1]
	g) Use the primary language of the user	[6]
	h) Communicate if type of device unknown	[3]
Follow a consistent layout	a) Present information in a logical order	[2]
	b) Put most important information at the beginning of the document	[2]
	c) Include headers and bullets for lists	[2]
	d) Use consistent words throughout the text	[6]
Improve trust	a) Include the ability to verify the authenticity of the notification	[4,6]

1. Bauer, L., Bravo-Lillo, C., Cranor, L., & Fragkaki, E. (2013). Warning Design Guidelines (CMU-CyLab-13-002). CyLab. Retrieved from http://repository.cmu.edu/cylab/113
2. CDC. (2010). Simply put; a guide for creating easy-to-understand materials. Retrieved from https://stacks.cdc.gov/view/cdc/11938
3. Livingood, J., Mody, N., & O’Reirdan, M. (2012). Recommendations for the Remediation of Bots in ISP Networks. RFC 6561 (Informational). Retrieved from http://www.ietf.org/rfc/rfc6561.txt
4. M3AAWG. (2007). Best Common Practices for the Use of a Walled Garden. Retrieved from https://www.m3aawg.org/sites/default/files/document/M3AAWG_Walled_Garden_BCP_Ver2_2015-03_0.pdf
5. Modic, D., & Anderson, R. J. (2014). Reading this May Harm Your Computer: The Psychology of Malware Warnings. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2374379
6. Online Trust Alliance. (2012). Combatting Botnets Through User Notification Across the Ecosystem: A View of Emerging Practices. Retrieved from https://otalliance.org/system/files/files/best-practices/documents/ota_botnet_notification_whitepaper12-7.pdf
7. Seiders, K., Flynn, A. G., Berry, L. L., & Haws, K. L. (2015). Motivating Customers to Adhere to Expert Advice in Professional Services. Journal of Service Research, 18(1), 39–58. https://doi.org/10.1177/1094670514539567
8. Wogalter, M. S., & Laughery, K. R. (1996). WARNING! Sign and Label Effectiveness. Current Directions in Psychological Science. Sage Publications, Inc.Association for Psychological Science. https://doi.org/10.2307/20182386

2.6 CONCEPTUAL MODEL

Based on the literature review, we developed the conceptual model to analyse the problem systematically (see Figure 5). The dependent variable in this model is IoT malware cleanup. The variables that are expected to influence the dependent variable are the independent variables: the notification channel and notification content. The mediating variable which explains the relation between the independent and dependent variables, is the reaction of the customers. The rationale behind the relations are briefly discussed below.

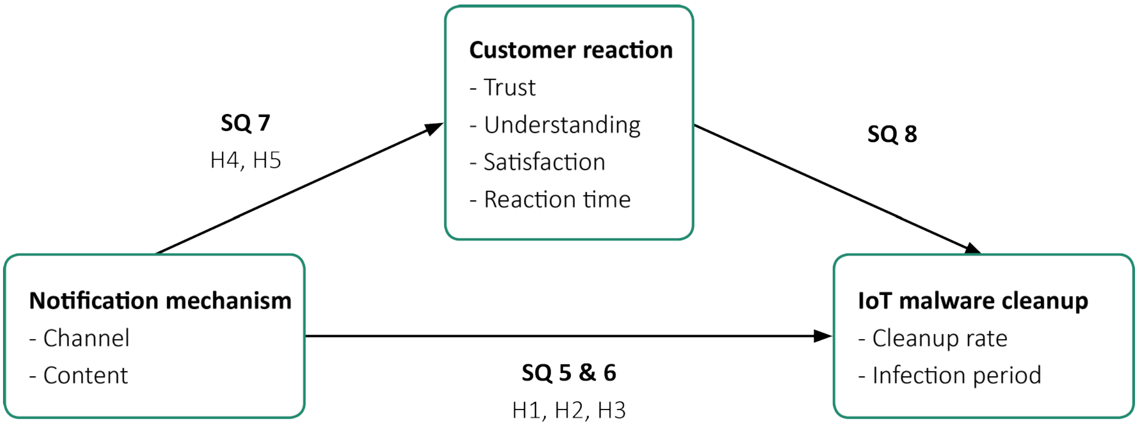


Figure 5 Conceptual model

2.6.1 RELATION BETWEEN THE INDEPENDENT AND DEPENDENT VARIABLES

As to the relation between the notification mechanism and the malware cleanup, Jhaveri et al. (2017) presented a list of factors that could influence the effectiveness of cleanup efforts (see Table 3). This research project investigates the effectiveness of IoT malware notifications from an ISP to its customers. Therefore the recipient, type of abuse and sender reputation are predetermined. This leaves the notification channel and notification content as a subject of analysis for this research project.

Influence of the notification channel on IoT malware cleanup

In Jhaveri et al. (2017), the notification channel is indicated as one of the factors that may influence the effectiveness of cleanup efforts. This statement is confirmed by recent studies that investigated the influence of the notification channel on remediation and cleanup (Gañán, Çetin, & van Eeten, 2015; Stock et al., 2018; Zhang et al., 2017). Moreover, Jhaveri et al. (2017) argue that connectivity restrictions, like in a walled garden, could influence the effectiveness of notifications in terms of cleanup as well. There are no empirical studies that measure the effectiveness of walled garden notifications. However, in the field of security warnings, a study by Egelman et al. (2008) indeed found that warnings that force users to notice by interrupting, are more effective. To analyse the relation between the notification channel and IoT malware cleanup in the scope of an ISP, we formulated sub-question 5: What is the added value of quarantining in terms of IoT malware cleanup? The empirical hypotheses that need to be tested are:

- H 1: Any kind of IoT malware notification reduces the duration of the infection.
- H 2: Walled garden notifications reduce the duration of IoT malware infections compared to email notifications.

Influence of the notification content on IoT malware cleanup

Jhaveri et al. (2017) suggested the notification content as another factor that may influence the effectiveness of cleanup efforts. However, studies in the field of notification effectiveness have contradictory results regarding the influence of a detailed notification content. Some studies found that a detailed message promotes cleanup and remediation (Çetin et al., 2016; Li, Durumeric, et al., 2016; Vasek & Moore, 2012). On the other hand, a study by Krol et al. (2012) into security warnings found no significant difference between a brief generic or longer specific warning. In order to investigate the influence of the notification content on IoT malware cleanup, we examine the impact of a more actionable walled garden notification content. To this end, we formulated sub-question 6: What is the influence of a more actionable walled garden notification content on IoT malware cleanup? The hypothesis that needs to be tested is:

- H 3: A more actionable walled garden notification message reduces the duration of the IoT malware infection.

2.6.2 RELATION BETWEEN THE INDEPENDENT AND MEDIATING VARIABLES

Both independent variables, notification channel and content, can influence the ultimate behaviour of the customer, in this case removing IoT malware. However, as illustrated by the C-HIP model (described in Section 2.5.1) a notification does not have a direct effect on the behaviour of the receiver. The reaction of the customer is a mediating variable. First, the notification should attract the attention, thereafter, the customer should trust and understand the notification in order to change the attitudes and beliefs of the customer. The receivers must be motivated by attitudes and beliefs in order to ultimately change their behaviour (Wogalter, 2006). The impact of the notification channel on the reaction of the customer is illustrated by different studies that pointed out the trust issues related to email notifications (Çetin et al., 2017; Livingood et al., 2012; Stock et al., 2018). After all, when a user does not trust a notification, the notification will be ignored. To examine the relation between the independent and mediating variables, we formulated sub-question 7: What are the reactions of the customers to different notification mechanisms for IoT malware infections? The hypotheses that need to be tested are:

- H 4: Customers respond faster to a walled garden notification compared to an email notification.
- H 5: Customers respond faster to a more actionable notification message.

2.6.3 RELATION BETWEEN THE MEDIATING AND DEPENDENT VARIABLES

Following the C-HIP model, it is expected that an IoT malware notification is more effective in terms of cleanup when a customer trusts and understands the notification. A number of studies in the field of security warnings indeed found that a lack of understanding negatively influences the effectiveness of security warnings (Krol et al., 2012; Mathur et al., 2016; Wash et al., 2014). However, as far as we know, the relation between the reaction of the customer to a malware notification and the cleanup has never been studied. In order to investigate this relation, we formulated sub-question 8: What is the impact of the customer’s understanding of the notification on IoT malware cleanup?

2.7 CONCLUSIONS

The defined goal of this chapter is to develop a conceptual model to analyse the problem systematically and to answer sub-question 1,2 and 4: (1). What are possible IoT malware notification channels for an ISP?; (2) Which factors influence the actionability of an IoT malware walled garden notification content?; and (4) How can the effectiveness of notification mechanisms on IoT malware cleanup be measured quantifiably? The conceptual model is presented in Section 2.6. In this section, an answer is provided to each of the sub-questions.

In the literature 6 possible notification channels for ISPs were described: (1) email; (2) telephone call; (3) postal mail; (4) instant message; (5) SMS; and (6) walled garden notification. Each method has its own limitations, either in terms of time to reach the customer, credibility, costs or reachability. To this end it is recommended to use multiple notification channels. In expert interviews with employees of KPN's abuse department, we evaluated the perceived effectiveness and feasibility of these notification channels for KPN. In the Chapter 3, we describe the results of these expert interviews.

In order to determine which factors influence the actionability of an IoT walled garden notification content, we combined the C-HIP model with PMT. Following the C-HIP model, we identified content characteristics that attract the attention of the receiver and improve the understanding. Moreover, factors were identified that drive the protection motivation. These factors were translated into guidelines, that can be used to compose an actionable walled garden notification message (see Table 4). In short, the guidelines include: (1) clearly specify the underlying risk; (2) write the message for the least technical user, therefore avoid technical terms; (3) provide clear and easily recognisable action steps; (4) write the message in primary language of the reader; and (5) include the ability to verify the authenticity of the notification.

In previous experimental studies, 2 quantitative metrics were used to measure the impact of different notification mechanisms on the cleanup of compromised websites: (1) cleanup rate; and (2) median time to cleanup. The cleanup rate is defined as the percentage of notified parties that is clean at the end of the investigation period. The second metric is the median number of days required to clean up those sites that were successfully remediated.

3 KPN'S ABUSE TEAM

In the previous chapter, we identified possible notification channels for an ISP by conducting a literature review. To validate the findings from the literature, we conducted expert interviews with abuse experts from KPN's abuse department. The rationale behind the expert interview design is discussed in Section 4.4 In these interviews we evaluated the perceived effectiveness and feasibility of the different notification channels found in the literature for KPN. Additionally, during the expert interviews, we map and evaluate KPN's current notification process for customers with infected IoT devices. This information enables us to understand which aspects of the notification process could be improved. Moreover, we need this information to be able to create a feasible empirical research design.

The result of this chapter is an answer to sub-question 1 and 3: (1) What are possible IoT malware notification channels for an ISP?; and (3) What is KPN's current notification process for customers with infected IoT devices? In order to describe KPN's current notification process, a distinction is made between the abuse feeds KPN acts upon, the notification mechanism KPN uses and the remediation process in Section 3.1 – 3.3. Thereafter, we evaluate the strengths and weaknesses of KPN's current notification mechanism pointed out during the expert interviews in Section 3.4. In Section 3.5, the different notification channels found in the literature are evaluated. We conclude this chapter, by answering the sub-questions. In this chapter we will discuss the information which is required to understand the research design. For the sake of the traceability of the data collection procedure, we elaborate on the systems KPN's abuse department uses in Appendix A.

3.1 ABUSE FEEDS

KPN's abuse department receives information about vulnerabilities and infections from external resources. KPN receives abuse reports from multiple abuse notifiers. Furthermore, every individual can report abuse caused by a subscriber of KPN via an email to the abuse team. The only data source that the abuse department consistently uses is Shadowserver. Established in 2004, the Shadowserver Foundation comprises volunteer security professionals from around the world that "gathers intelligence on the darker side of the Internet" (Shadowserver, 2018). Each day, Shadowserver sends reports to the abuse department containing all abuse logs from KPN's subscribers in the past 24 hours. One of Shadowserver's abuse feeds reports IoT malware infections: the drone-report contains information about Mirai infections. As this IoT abuse data can be used in this study, we discuss the format of this abuse feed in more detail in the research methodology in Section 4.1.1.1.

3.2 NOTIFICATION MECHANISM

Based on a predefined policy, the abuse department decides which of the infected and vulnerable customers to notify. First of all, a subset is created of the infection and vulnerability types KPN notifies (see Appendix B). In addition, the abuse department checks the timeliness of the abuse data. Since Shadowserver's abuse feeds report abuse events from the previous day, it should be checked whether the customer has solved the issue after the reported abuse event. Lastly, a distinction is made between four markets: the consumer, business and mobile market and the Telfort market. Based on the IP, it is determined to which IP range, and therefore market, the customer belongs. Since mobile customers cannot be identified based on an IP address, due to IP churn, no action is taken. Customers in the other markets have, in principle, static IP addresses. There are, however, exceptions. Firstly, a customer gets a new IP address after moving to a new address. Secondly, the IP address changes when changing from the copper to the fiber optic network. Moreover, it can be decided to change the IP address for other maintenance work. Although the IP address and contact details of a customer could change, the customerID a customer gets at subscription never changes. In the following, we describe the notification process for the consumer market, Telfort and the business market.

3.2.1 NOTIFICATION MECHANISM CONSUMER MARKET

In order to notify consumer market customers with a malware infection, KPN uses a so-called strict implementation of a "walled garden". The overall quarantine process is presented in Figure 6. While the user tries to browse the Web, the customer is redirected to a landing page, except for a small set of white-listed sites. The landing page provides information about the type of infection and how to clean it up. Moreover, an email is sent along with quarantine event. This email provides the same information as the landing page plus an email address to contact in case of questions or problems while solving the problem. In Appendix C and D we present the quarantine landing page and the email for a Mirai infection respectively.

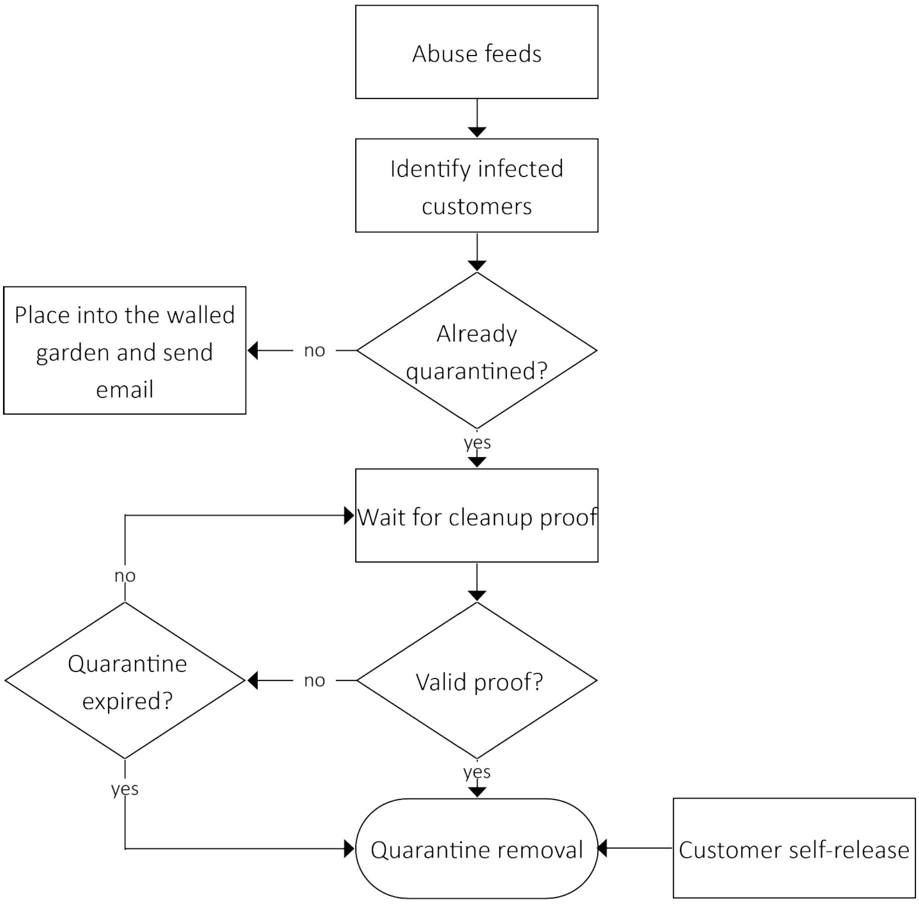


Figure 6 Quarantine flow diagram

There are three ways the customer can get out of the walled garden. Firstly, customers can release themselves from the quarantine environment, by filling out the contact form and reporting on how they have fixed the problem. This self-release option is revoked after two subsequent quarantine events within 30 days, to avoid customers using this route to restore their connection without making an effort at remediation. The second way out is when the ISP's abuse staff releases the customer connection. Customers might end up in assisted release because they no longer have the self-release option or because they have contacted KPN for help. Quarantined customers can contact abuse desk members via email and a walled-garden form. The third way of being released is when the expiration date passes. After 30 days, a customer is automatically released, even if they have not contacted the ISP.

The walled garden has a limited capacity. When all slots are taken, but the abuse department still wants to notify and remediate, it sends an email notification to the mail address that it has on record as the primary contact for that customer. The message contains the same information as the walled garden's landing page, plus an email address to contact in case of questions or problems while remediating the vulnerability. This email is presented in Appendix E.

3.2.2 NOTIFICATION MECHANISM BUSINESS MARKET

Business customers are, in principle, always notified by email. However, if the customer does not respond to the emails and the customer keeps returning in the abuse feeds, the customer is called. If the customer does not respond, the internet connection could be blocked to force a reaction from the customer.

3.2.3 NOTIFICATION MECHANISM TELFORT MARKET

The notification mechanism for Telfort customers is the same as for consumer market customers. Wherever KPN is mentioned in the messages, it is replaced by Telfort. In principle, both markets have the same priority. However, in reality, Telfort customers are often not notified due to the smaller size of this market, and as a result smaller number of infections, and the considerable amount of manual work required to notify these customers. Because of the limited capacity of the abuse team, the time to do such manual notifications is restricted.

3.3 REMEDIATION

In order to solve the problem, the customer can visit a limited number of whitelisted websites in the quarantine area. This list includes the websites of KPN, antivirus vendors, webmail services and banks. This way customers can take appropriate action to clean the system and contact KPN, even though they are quarantined. On the landing page, customers are stimulated to follow the link to the contact form (see Appendix F) and to send a completed form to the abuse team. In this form, which is the same for all infections and vulnerabilities, the customer is asked to describe how many laptops/computers are connected, to send log files of the executed virus scans, and to explain the measures taken to solve the issue. The targeted questions in the contact form enables the abuse team to assess to what extent the customer has taken the right actions.

In addition to the walled garden contact forms, customers can also contact KPN via other channels. The customers can send an email to the abuse department or contact the help desk via phone calls, store visits, chat or social media, for example to ask for additional help. When a customer contacts the KPN abuse team for additional help, the abuse team tries to help the customer to perform the requested steps via email. As all communication between customers and the abuse team is stored, contact forms and emails are in principal the only communication channels KPN uses. The KPN abuse team only calls a customer when they think it is beneficial to their time. When a customer contacts the help desk, the help desk employee advises the customer to send an email to the abuse team. Moreover the customer can get assistance from a technician. This is a paid service. This latter service is, however, not preferred by KPN as these technicians are not meant for support of the customer's devices. Nevertheless, when a help desk employee does not correctly identify a quarantine action from the abuse team, it might be thought that the customer cannot use its Internet because of a broken modem.

3.4 EVALUATION CURRENT NOTIFICATION MECHANISM MIRAI

In interviews with D. van Drunen and R. Teunissen, employees of the KPN abuse team, the strengths and weaknesses of the current notification mechanism for the consumer market have been discussed. A distinction is made between the notification channel (walled garden along with an email) and message content for Mirai infections. Moreover, it is discussed if there is a difference in terms of strengths and weaknesses between Mirai (IoT) and other infections.

3.4.1 NOTIFICATION CHANNEL

A first strength of using a walled garden notification along with an email is that the customer is notified and simultaneously, the communication between the bot and C&C server is blocked. Therefore, no commands can be issued to the bot. Secondly, in the quarantine area, customers can visit a limited number of websites, which could help to solve the issue. Lastly, the KPN abuse team knows for sure that the customer sees the walled garden notification. Since every non-whitelisted website, a customer tries to visit, will redirect the customer to the landing page, it is practically impossible for customers not to see the notification. Even when customers do not see the landing page, they will notice something is wrong and will therefore contact KPN. There is no difference between Mirai and other security problems with regard to the strengths of walled garden notifications.

The major weakness of the notification mechanism lies in the quarantine system design and the way KPN currently uses it. The system is not dynamic enough. The contact form (see Appendix F), which the customer gets when placed in the walled garden, is a static form. There is no ability to differentiate between infections. However, for a computer virus you want to ask vastly different questions than for a Mirai infection.

3.4.2 NOTIFICATION CONTENT

The notification content, as shown in Appendix C, describes a technical problem for customers that are often not technical. The abuse experts indicate that the message is too technical, but that it is also really hard to simplify it to a level where the customer understands what is meant. Most of the time the customers do not even know that they have a camera connected to the Internet. This problem will probably get worse, when more devices are (automatically) connected to the internet.

3.5 EVALUATION POSSIBLE NOTIFICATION CHANNELS

Livingood, Mody, and O'Reirdan (2012) published a document with recommendations on how ISPs can use various notification mechanisms. These notification mechanisms are discussed in an interview with employees of the KPN abuse team.

Email: A drawback of this notification mechanism is that it is uncertain whether the warning is received and if it is, you do not know whether the customer trusts the message. Customers could think the email is spam or phishing, considering that KPN has been a serious victim of phishing emails in the past years. Furthermore, the customer might use a different email address than the one provided to KPN.
Postal mail: KPN has never seriously considered this option, because it takes a while before the letter gets to the customer. Furthermore it is not an automated process, so it will take a lot of time, which makes it more expensive.

SMS: KPN is planning to use an SMS notification in the near future along with the walled garden system. Customers will receive a SMS, telling them that they are placed in walled garden. KPN thinks customers will see the notification sooner and are more likely to trust the notification.

Telephone call: KPN has tested this notification mechanism along with emails in the past. The abuse desk sent emails and there was a team that called the customer. The problem KPN ran into was that the people calling the customer did not have the technical ability to solve the issues of the customers. The only thing they did was telling the customer that they received an email from the abuse team, as a validation. If customers had questions, they had to email the abuse team. This led to frustration for the customers. This notification channel is something that KPN might want to investigate again in the future.

Instant messages: This notification mechanism has never been considered by the abuse department, as it would require the investment in a new notification system.

It is indicated that it is preferable to use multiple channels, as this increases the chance that customers will trust the notification.

3.6 CONCLUSIONS

The defined goal of this chapter is to answer sub-question 1 and 3: (1) What are possible IoT malware notification channels for an ISP?; and (3) What is KPN's current notification process for customers with infected IoT devices? In this section, an answer is provided to each of the sub-questions.

In Chapter 2, we describe the 6 different abuse notification channels which could be used by ISPs to notify infected customers that are described in the literature: (1) email; (2) telephone call; (3) postal mail; (4) instant message; (5) SMS; and (6) walled garden notification. During an interview with abuse experts from KPN, we evaluated the feasibility of the different notification channels for KPN's abuse department. KPN currently uses a combination of walled garden and email notification notifications. The major advantages of the walled garden notification are the high likelihood that the customer sees the notification and the fact that the communication between C&C and the bot is blocked. In contrast, the abuse department does not know whether a customer has received and read an email notification. The customer could think it is either spam or a phishing attempt. Therefore, the abuse experts indicated that it is preferable to use multiple channels, as this increases the credibility of the notifications. The abuse department considers to use SMS notifications together with another notification channel in the future. The abuse department does not use postal mail notifications due to the additional preparation time, delivery time, and additional cost to KPN. Moreover, telephone calls can be time-sensitive, but it requires even higher cost to keep it running. Instant message notifications have never been considered by the abuse department, as they would have to invest in a new notification system.

As stated in the previous paragraph, right now, when the abuse department receives information about IoT malware infections in their network from Shadowserver, the customer is placed in a so-called walled garden and simultaneously an email is send to the customer. The first two times a customer is quarantined within 30 days, the customer has a self-release option. After that, only KPN's abuse team can let the customer out of the walled garden, when they send valid proof of cleanup. After 30 days the quarantine event expires automatically. When a quarantined customer opens its web browser, a landing page appears with a malware-specific message. The customer is asked to perform certain steps and fill in and send a contact form to the abuse department. In this contact form questions are asked such as: "Which anti-virus software do you use?", "How many computers are connected?" and "Which measures did you take?". To solve problem, customers can contact the abuse team via email for additional information. Moreover, customers can contact the help desk via phone, chat or social media. During the expert interviews it was pointed out that this static contact form is a weakness of the current quarantine system. After all, questions about a virus scanner or the number of laptops are irrelevant for an IoT malware infection and could be misleading for customers. Additionally, the abuse experts indicated another weakness. The current Mirai notification message is too technical to be comprehended for customers.

4 RESEARCH METHODOLOGY

In this chapter the methodology that is employed to answer the main research question is presented. First, in Section 4.1, the available data sources are presented. Based on the available data, it is determined how the dependent and mediating variables are evaluated in Section 4.2 and 4.3 respectively.

The subsequent sections describe the different research methods used in this study. In Section 4.4 we describe the expert interview design aimed at mapping and evaluating KPN's abuse notification mechanisms with abuse experts from KPN. Secondly, in Section 4.5, we explain the design of the experiment. The purpose of this experiment is collecting information on how different notification mechanisms influence the cleanup rate and infection time of Mirai infections. In addition to the evaluation of the cleanup rates and infection times, the customer reactions to the different IoT malware notification mechanisms is investigated. To this end, the messages from the customers in the experiment to KPN are collected and analysed and customer interviews are carried out. In Section 4.6, the customer interview design is explained. Moreover, it is described how these communication logs and customer interviews are systematically analysed. In Section 4.7 the methods to statistically analyse the data are explained. Lastly, the ethical considerations are pointed out in Section 4.8.

The research is restricted to KPN's consumer market, because this is the only market for which the KPN abuse team has a uniform notification procedure. In contrast, customers are not consistently notified in the business and Telfort market. In case of reported abuse issues in these markets, the abuse team makes the decision whether or not to notify the affected customers. There are no set rules for this. Therefore, it is impossible to make a consistent analysis for these markets.

4.1 DATA

The data used in this research project was gathered from different sources: (1) abuse data; (2) notification logs; (3) abuse team communication logs; and (4) help desk communication logs.

4.1.1 ABUSE DATA

We use 2 different data sets in order to detect and track IoT infections in KPN's network: Shadowserver Botnet-Drone report and IoTPOT feed. Of these abuse feeds, the Shadowserver Botnet-Drone report is the one that specifies the malware infection. Therefore this abuse feed is used to detect the Mirai infections in KPN's network. After learning which customers are infected with Mirai, the Shadowserver and IoTPOT feed are both used to track the Mirai infections.

Shadowserver Botnet-Drone

As described in Section 3.1, the Shadowserver's Botnet-Drone report provides security incident data to KPN. This report contains a list of all the infected machines that Shadowserver detected from the monitoring of IRC Command and Controls, capturing IP connections to HTTP botnets, or the IPs of Spam relays. This report includes IoT malware Mirai. This is also the only reported IoT malware. An overview of the content of this report is presented in Appendix G. The reports, which are sent to KPN daily around 9 a.m. (UTC+1), represent the activity monitored in KPN's network during the 24 hours of the previous day. The logs are in UTC+0. Even though, there could be multiple events for an IP in a day, the reports only include the first event for each IP.

IoTPOT

The second information source regarding IoT compromises is a dataset from IoT honeypots (IoTPOTs). In computer sciences, a honeypot is a computer system which is intentionally vulnerable to viruses and other attacks. Generally, the honeypot consists of seemingly legitimate content that appears as valuable to the attacker. This acts however as bait, since the honeypot is actually isolated and monitored. An analysis of the gathered information is useful to prevent further spreading of a virus. IoTPOT is a novel honeypot, proposed by Pa et al. (2015), which mimics IoT devices and captures Telnet-based intrusions. Appendix H gives an overview of the IoTPOT. The gathered information, regarding IP addresses abusing IoT honeypots, can be downloaded from a public website[1]. The format of the hourly updated IoTPOT dataset is represented in Table 5. Which IoT malware it concerns is not mentioned. The logs are in UTC+9.

Table 5 Format IoTPOT data

Field	Description
ts	Timestamp the IP was seen in UTC+9
ip	The IP of the device in question
cc	The country location of the IP
asn	ASN of the IP
org	The organization to which the IP belongs
iot_type	The type of infected machine
manuf	The manufacturer of the infected machine

Preparatory analysis Shadowserver and IoTPOT data

In this section the preliminary analysis of the two abuse feeds is described. First it is analysed how many unique customers appear in the reports per day. This information is necessary in order to determine a feasible number of treatments with a certain sample size in the experiment. Secondly, the overlap between the abuse reports is investigated.

The abuse reports from April 11th, 2017 to October 10th, 2017 are used for this analysis. In order to analyse the data of these six months, preparatory steps are taken:

1. All abuse logs of organizations other than KPN are filtered from the IoTPOT data.
2. To determine how many unique IP addresses appear in the abuse reports in one day, the timestamps of all abuse logs are changed to the same timezone, UTC+1. Therefore 8 hours are subtracted from the timestamp of IoTPOT abuse events, and likewise one hour is added to each Shadowserver event.
3. Based on the IP ranges, a subset is created of IP addresses that belong to the consumer market.

Subsequently, a distinction is made between the total number of unique IP addresses on a day and the number of new IP addresses. An IP address is identified as new, if it is the first appearance with a specific infection in the dataset since April 11th, 2017. If an IP reappears with a different infection, the customer is indicated as new again. The number of total and new IP addresses is therefore equal on April 11th, 2017.

[1] http://pierogi.ip-eend.nl:108

Table 6 Number of unique consumer market IPs from April 11th, 2017 until October 10th, 2017

		Shadowserver Mirai	IoTPOT
# total unique IP addresses per day	Min.	1.0	1.0
	Median	6.0	3.0
	Mean	6.3	3.7
	Max.	18.0	14.0
# new unique IP addresses per day	Min.	0.0	0.0
	Median	1.0	1.0
	Mean	1.9	1.0
	Max.	14.0	8.0
	Complete period	339.0	144.0

Looking at the number of new unique consumer market IP addresses with a Mirai infection, the median is 1 (see Table 6). This information is used in Section 4.5.1, to determine a feasible number of treatment groups with a certain sample size.

As the IoTPOT data does not identify the infection type, it is interesting to see the overlap between both datasets in the six months period. It is found that 74 unique consumer market IP addresses that appear in Shadowserver with a Mirai infection, also appear in IoTPOT. An additional 21 unique consumer market IP addresses that appeared in IoTPOT, appeared in Shadowserver with a different infection indication. Of these, 1 was indicated as aaeh, 1 as unknown and 18 as sinkhole, where in all cases the dates of appearance differed by less than 2 weeks. The other IP address was indicated as wannacrypt. But here, the appearance in both datasets were 4 months apart.

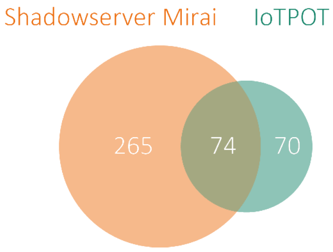


Figure 7 Visualisation overlap IoTPOT and Shadowserver consumer market data from April 11th, 2017 to October 10th, 2017

4.1.2 NOTIFICATION LOGS

The notification logs record details of quarantine events and email notifications in KPN's network. For each of the events the timestamp of the notification and the infection type are recorded. Additionally, for quarantine events the quarantine release mechanism, the quarantine removal timestamp, the quarantine event number and the self-release option are stored.

4.1.3 HELP AND ABUSE DESK LOGS

The help desk logs contain KPN's help desk communication with customers. Abuse team communication logs provide email exchange between abuse team employees and customers. Beside the email communication, the abuse desk logs contain the walled garden contact forms that customers can submit through the walled garden landing page (see Appendix F).

4.2 EVALUATING IOT MALWARE CLEANUP

Like in previous studies (Çetin et al., 2016; Vasek & Moore, 2012), the effectiveness of notification mechanisms in terms of IoT malware cleanup is evaluated quantifiably based on two metrics:

- *Median infection time:* to determine the infection time, the infection is tracked by means of the Shadowserver reports and the IoTPOT feeds in the 14 days after the first notification. The tracking procedure is described in more detail in Section 4.5.3. The median infection time is the "middle" value, when taking all infection times. For the customers that were still infected at the end of the experiment period, we take an infection time of 336 hours (14 days). The advantage of the median infection time compared to the mean, is that it gives a better idea of a "typical" infection time. The value is not skewed by extremely large or small values.
- *Cleanup rate:* the cleanup rate is defined as the percentage of customers that cleaned up within the 14 days experiment period.

4.3 EVALUATING THE REACTIONS OF THE CUSTOMERS

The reactions of the customers after an IoT malware notification are evaluated both quantitatively and qualitatively. The customer reactions are evaluated quantifiably based on two metrics:

- *Median reaction time:* the reaction time is defined as the time (in hours) between the first notification and the first reaction from the customer via an email, the contact form, or contact with the help desk. The median reaction time is based on the reaction times of the customers that actually reacted.
- *Reaction rate:* the reaction rate is the percentage of customers that respond to the notification within 14 days after the initial notification.

In addition, by means of an analysis of the help and abuse desk communication logs and telephone interviews, it is explored how the customer perceived the different notification mechanisms. As presented in the conceptual model in Section 2.6, there is a particular focus on certain aspects:

- *Trust:* whether customers trust the notification messages, or if they think of it as either spam or a phishing email.
- *Understanding:* to what extent customers understand the content of the notification message. Do they understand what to do? Are they able to identify the infected device? Do they need additional help?
- *Satisfaction:* how satisfied are customers with the notification mechanism? Are they angry, and if so, why?
- *Suggestions:* what do customers suggest in order to improve the current notification mechanism?

4.4 EXPERT INTERVIEWS

To understand what IoT malware notification channels are possible for an ISP (SQ 1) and what KPN's current notification process for customers with infected IoT devices is (SQ 3), 2 semi-structured interviews are carried out with employees of the KPN abuse team. Since the interviews serve different purposes, we discuss the design of both interviews separately in Section 4.4.1 and 4.4.2. It should be noted that the result of these interviews has already been presented in Chapter 3.

4.4.1 EXPERT INTERVIEW I: MAPPING NOTIFICATION MECHANISM

The goal of the first interview is to map KPN's current notification process for customers with IoT malware infections (SQ 3). The interview questions (see Appendix I) are formulated and structured based on the reference model from Pijpker & Vranken (2016). This reference model summarizes measures for botnet mitigation that ISPs currently take, structured according to the anti-botnet lifecycle: prevention, detection, notification, remediation and recovery (Online Trust Alliance, 2012). The interview questions concern 3 of these stages:

- *Detection:* we need to know which IoT abuse data KPN has and acts upon.
- *Notification:* we examine what KPN's IoT malware notification procedure looks like. We examine how the abuse department decides whether or not to notify a customer who is known to be infected with IoT malware. It is investigated which notification content and channel are used and whether this is different for subsequent notifications. In addition, the differences between IoT malware notifications and notifications for other malware types are mapped.
- *Remediation:* it is examined which additional help KPN provides to infected customers in order to solve the problem. Moreover, we discuss how customers react to IoT malware notifications and to what extent this is different for IoT malware infections compared to notifications for other security issues. Lastly, we investigate how KPN knows whether a customer has successfully removed the infection.

4.4.2 EXPERT INTERVIEW II: EVALUATING NOTIFICATION MECHANISM

The aim of the second expert interview, that is conducted after the first interviews, is to evaluate the findings from literature about possible notification channels for ISPs (SQ 1) and to evaluate KPN's current IoT malware notification channel and content (SQ 3). This interview is held with R. Teunissen and D. van Drunen, employees of KPN's abuse team, simultaneously. The interview questions (see Appendix J) are structured as follows:

- *Notification channel:* we examine the strengths and weaknesses of KPN's current notification channel. Moreover, we discuss the other possible notification channels we found in the literature.
- *Notification content:* the strengths and weakness of the notification content for customers with an IoT malware infection are evaluated. Additionally, possible improvements are discussed.

4.5 EXPERIMENT

What is the added value of quarantining (SQ 5) and what is the influence of the actionability of the walled garden notification content (SQ 6) on IoT malware cleanup? We designed an experiment at KPN's abuse department measuring infection times and cleanup rates as a result of notifications sent to customers via different channels and with different contents. In this section, we explain the design of the experiment. First, in Section 4.5.1, we describe how the sample size is calculated. Then, we outline the different treatment groups and discuss the experimental design that is used in Section 4.5.2 and 4.5.3 respectively. Subsequently, we describe the content of the notifications. Finally we discuss the limitations of the experiment in Section 4.5.5.

4.5.1 SAMPLE SIZE

In order to draw right conclusions from the experiment, a minimal sample size is needed. On the other hand, testing too many subjects is also undesirable. If a treatment turns out to be ineffective, too many subjects have been exposed to this ineffective intervention. Therefore, it is critical to calculate the appropriate sample size before the experiment.

To determine the sample size, a power analysis is done by means of the pwr.t.test in R. The values used in the power calculation are shown in Table 7. As the effect size is unknown, a medium effect size is chosen. Based on these values, the required sample size is 100. As the median number of new unique consumer market customer daily is 1, see Table 6, three months are required for a treatment group.

Table 7 Values used in power analysis

Parameter	Value
Desired power of the study (power)	80%
Desired significance level (sig.level)	5%
Effect (d)	40%
Desired test direction (alternative)	Two sided test
Test to be used in the statistical analysis (type)	Two sample t-test

4.5.2 TREATMENTS

In order to investigate the added value of quarantining in terms of IoT malware cleanup, two notification channels have been selected for the analysis: (1) walled garden together with email; and (2) email. These mechanisms have been chosen, because both are already used by KPN. Therefore the experiment could be executed on the short-term and without additional costs. Moreover, this way it is analysed whether the walled garden notification, that restricts the customer, actually improves the cleanup.

The influence of a more actionable walled garden notification content on IoT malware cleanup is analysed by comparing two notification messages. This first message is the message the KPN abuse team wrote for customers with a Mirai infection, referred to as old content. The second message, which is referred to as new content, is drafted for the experiment based on the guidelines from the literature (see Section 2.5). The reasoning behind this new message is described in Section 4.6.4.

In addition, the control group serves as a baseline to understand the natural cleanup rate of a compromise. This group receives no treatment.

This way four treatments groups are formed: (1) Control; (2) Walled garden with old content; (3) Walled garden with new content; and (4) Email with new content. First of all, these four treatment groups enable us to compare the treatment groups with the control group, to measure if any kind of IoT malware notification reduces the duration of the infections compared to not sending a notification (H1). Secondly, we can compare the email with new content and walled garden with new content treatment group, to examine whether quarantining reduces the duration of the IoT malware infection compared to an email notification alone (H2). Lastly, the walled garden with old content and walled garden with new content treatment groups can be compared to investigate if a more actionable walled garden notification message reduces the duration of the IoT malware infection (H3).

4.5.3 EXPERIMENTAL DESIGN

There are two sequential stages in the experiment, as illustrated in Figure 8. Data of the walled garden with old content treatment group is historical data provided by KPN. In contrast, data of the other treatment groups is collected during a randomized controlled experiment. First, we discuss how we collected the notification and abuse data in both stages of the experiment. Thereafter, in Section 4.5.3.3, it is explained how we tracked the infection and determined the infection time after the initial notifications for both the historical data and the randomized controlled experiment.

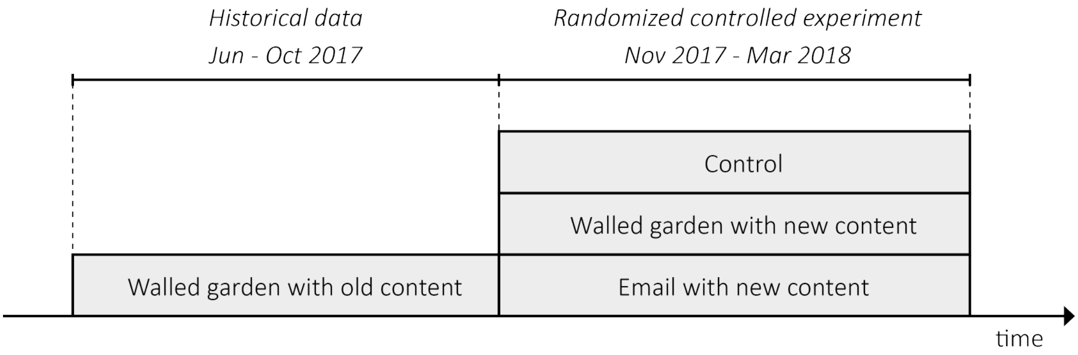

Historical data
Jun - Oct 2017
Randomized controlled experiment
Nov 2017 - Mar 2018
Control
Walled garden with new content
Walled garden with old content
Email with new content
time

Figure 8 Timeline experiment

Historical data collection

For the historical data collection, two of the data sources discussed in Section 4.1 are used: (1) abuse data; and (2) notification logs. In order to detect the Mirai infections in KPN's consumer market, all daily Shadowserver reports from April 11th, 2017 to October 10th, 2017 are collected. The notification logs are used to examine whether the Mirai infections have triggered a quarantine action. Then a subset is created of the last 100 unique customers that were placed in the walled garden before October 11th, 2017. These 100 customers form the walled garden with old content treatment group.

Randomized controlled experiment

The analysis and data collection for the other three treatment groups started on November 6th, 2017 and continued through March 1st, 2018. Figure 9 illustrates the rules we applied to get the experimental data from the abuse feeds provided by Shadowserver. During the experimental period, consumer market customers with a Mirai infection that appeared in the Shadowserver feed of the previous day were notified on working days. Before notifying, we checked whether the customer has been notified before. Since experience with both the notification procedure and remediation could influence the infection time, customers that have been notified for any infection or vulnerability after April 10th, 2017 are discarded. Consumer market customers that have not been notified before are randomly distributed to a treatment condition or to the control group. After the initial notification (or first detection for the control group), the customer is tracked for 14 days. When the customer is seen in the Shadowserver feeds within these 14 days, the same notification is repeated. Given that the Shadowserver feeds report abuse events of the previous day, two things were checked before repeating the notification: (1) has the customer been notified after the timestamp in the abuse feed?; and (2) has he customer sent proof of cleanup after the timestamp in the abuse feed? If so, no notification is sent.

For the control group, if the Mirai infection of a customer has not been cleaned up after the 14 days experimental period, the customer is randomly assigned to either the walled garden with new content or email treatment group. This is possible, because people in the control group are not aware of the Mirai infection since their IoT device kept working as intended. In contrast, if the Mirai infection of a customer in the email or walled garden with new content treatment group has not been cleaned up after the 14 days experimental period, KPN's regular notification procedure is executed: the customer is placed in quarantine and receives an email notification simultaneously.

To assign customers to a treatment group or to the control group, R function sample() was used. Each treatment group should contain 100 customers. Therefore, before the start of the experiment, a list of 300 entries has been created, in which all three experimental treatments are represented 100 times in random order.

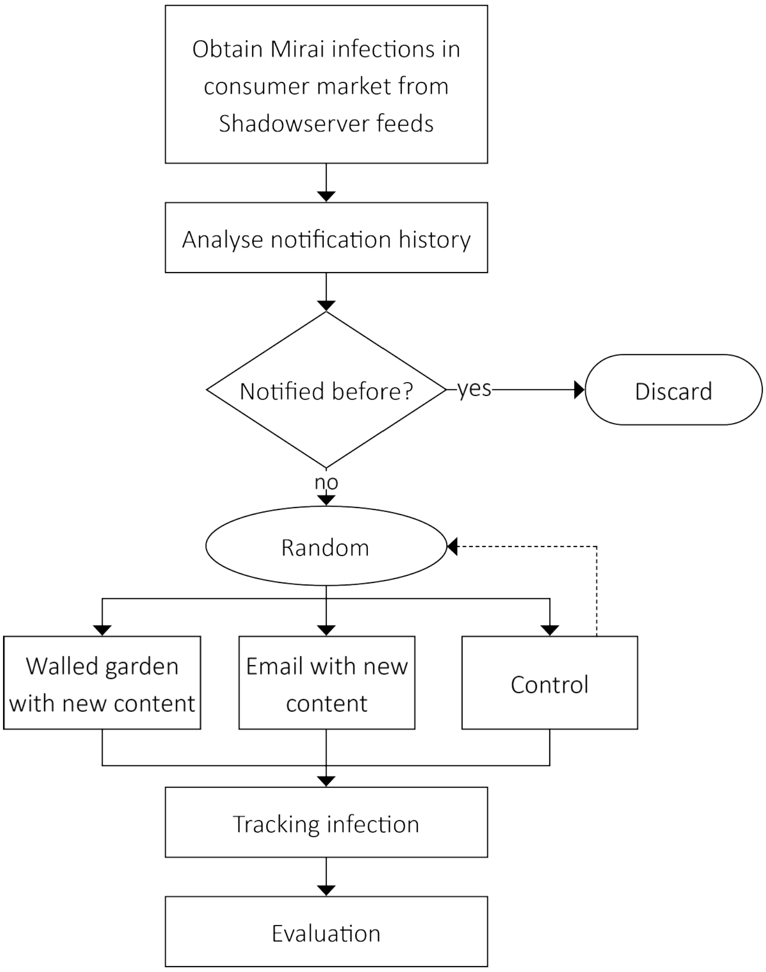

Obtain Mirai infections in consumer market from Shadowserver feeds
Analyse notification history
Notified before?
yes
Discard
no
Random
Walled garden with new content
Email with new content
Control
Tracking infection
Evaluation

Figure 9 Flow diagram of randomized controlled experiment

The role of KPN

During the experiment, the abuse team removed all consumer market customers with a Mirai infection from their regular notification procedure. Email and walled garden notifications for Mirai infections were all sent manually. Moreover, the abuse team did not notify customers during their experimental period of 14 days for other infection and vulnerability types either. For this purpose, a list containing all IP addresses, which were part of the experiment, was shared with all employees of the abuse team. This list was updated daily.

Tracking the presence of the infection

In the 14 days after the initial notification, the infection is tracked with the abuse feeds from Shadowserver and IoTPOT. These 14 days is referred to as "experimental period". Thereafter, the infection is tracked for an additional 30 days ("observation period") to prevent an underestimation of the infection time. Infected machines are not seen in the abuse feeds every day or even every few days. It depends on the malware behaviour, but also whether the user turns on the device. Therefore, we decided to count conservatively in terms of cleanup success and use a long period before considering a device clean.

In principal, in all treatment groups, the infection time is based on the period (in hours) between the initial appearance and the last appearance in either of the abuse feeds during the study period. However, due to the quarantine event of the walled garden treatment group and the lack of an initial notification in the control group, additional factors need to be taken into account. Figure 10 illustrates the rules we applied to determine the infection time for the different treatment groups. This procedure is described in more detail in the following sub-sections.

Walled garden with old content and new content

Since all communication between the C&C server and the bot is blocked while the customer is in the walled garden, the infected device will not appear in the abuse feeds during the quarantine event, even though the device may still be infected. It is reasonable to assume that a customer will not take action to remove the infection and subsequently remains in the walled garden area. Therefore, the timestamp of the last release from quarantine is taken as the end of the infection time. There is, however, an exception. If an IP address subsequently returns in the abuse feeds, but no notification is sent, the last appearance in the abuse feeds will be taken as the last day of the infection. There are two reasons why no notification is sent after an appearance in an abuse report. Firstly, because the abuse feed is sent during the weekend and secondly, when an IP address occurs only in the IoTPOT report.

In short, to evaluate cleanup for the walled garden treatment groups, we distinguish four different outcomes: (1) the customer successfully performed cleanup, was released from quarantine and then stays clean for the rest of the study period; (2) the user released from quarantine, but did not successfully cleanup the machine, as evidenced by seeing the customer reappear in the abuse feeds during the experiment period; (3) quarantine release after the end of the experiment period of 14 days; and (4) the customer released from quarantine but did not successfully cleanup the machine, as evidenced by the reappearance of the customer in the abuse feeds during the observation period. In the latter two cases the infection time is equal to the maximum: 336 hours (14 days).

Email with new content

For the email treatment group, the period between the initial notification and the last appearance in the abuse feeds is taken as the infection time. However, if the customer appears in the abuse feeds in the observation period, the infection time is equal to the maximum: 336 hours (14 days). When a customer is notified on day 0 based on a Mirai infection reported by Shadowserver and is not seen in either of the abuse feeds afterwards, the infection time is considered 0 hours.

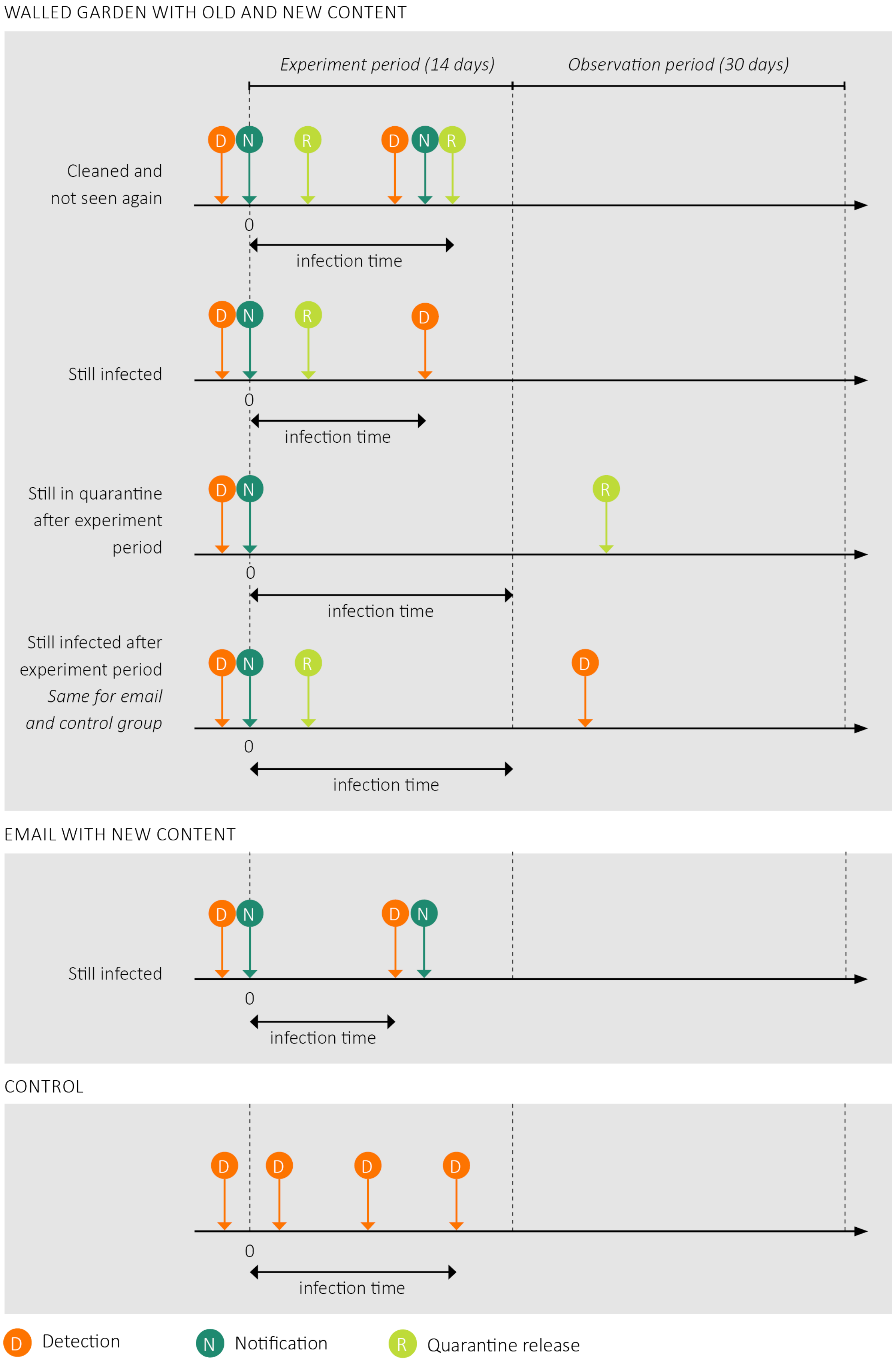

WALLED GARDEN WITH OLD AND NEW CONTENT
Experiment period (14 days)
Observation period (30 days)
Cleaned and not seen again
Still infected
Still in quarantine after experiment period
Still infected after experiment period
Same for email and control group
0
infection time
D
N
R
EMAIL WITH NEW CONTENT
Still infected
CONTROL
Detection
Notification
Quarantine release

Figure 10 Tracking the presence of the infection per treatment group

Control

There is no notification in the control group. Therefore the starting point of the experiment period for the control group is 9 a.m. (UTC +1) on the day after the reporting day of Shadowserver, as this is comparable with the timestamp of the initial notification in the treatment groups. The infection time is defined as the number of hours between the starting point and the last appearance in either of the abuse feeds. Again, when the customer appears in the abuse feeds in the observation period, the infection time is equal to the maximum: 336 hours (14 days).

4.5.4 COMPOSING A NEW NOTIFICATION MESSAGE

In order to be able to test the influence of a more actionable notification content on IoT malware cleanup, a new notification message for Mirai infections is composed based on the guidelines from literature (see Section 2.5). Appendix L and Appendix M show the new walled garden and email notification respectively. In this chapter, it is described how we arrived at the message. First, the content of KPN's current IoT malware notification content is evaluated in Section 4.5.4.1. Secondly, the rationale behind the new content is discussed in Section 4.5.4.2.

Notification message KPN

KPN has three notification messages for Mirai infections:

1. Message on landing page quarantine area (see Appendix C)
2. Email along with the walled garden notification (see Appendix D)
3. Email notification that is sent when daily limit is reached (see Appendix E)

The messages consist of three parts. Only the middle part, under heading "What is the problem and how can you solve it", can be adjusted. The other two parts are the same for all infections and vulnerabilities. Therefore, this chapter will focus on this adjustable part, referred to as body of the notification message.

In Table 4 guidelines and principles for writing abuse notification messages for resource owners from literature were presented. Based on these guidelines, KPN's notification message is evaluated and possible improvements are proposed (see Table 8).

Rationale behind content new notification message

As the effectiveness of the walled garden and email notifications is compared in the experiment, the body of the messages is the same in both notifications. For this purpose, an English translation is added to the email notification as well.

Restrictions

The content and layout of the notification message have various restrictions. Firstly, the variable body of the message has a limit of 2000 characters, including spaces. Therefore, it has been decided to focus on the action steps without describing the possible consequences of the Mirai infection. Next to that, no hyperlinks and italic, bold or underlined text can be used. Therefore, the URLs need to be completely written out and headers are difficult to recognize. This affects the readability of the message.

Moreover, there are two restrictions that might lower customer's trust. Firstly, the email message is plain text (see Appendix M). No KPN logo can be added. Furthermore, the message is static. No information, such as the customer's name or customerID can be added in the message. Therefore, the message does not meet the guideline to provide the ability to verify the authenticity of the message.

Choice of words

In the new notification message the amount of technical terms has been minimised as much as possible. Firstly, the wording Mirai virus is used to indicate the Mirai malware. Although this is technically incorrect, non-technical users are more likely to be familiar with the concept. Moreover, the words Internet of Things and IoT, DDoS, Telnet and SSH have been avoided. IoT devices are consistently referred to as Internet connected devices. Furthermore, as there is no certainty about the nature of the infection and the infected device, words like most likely, might and could have been used in describing the problem.

Table 8 Evaluation of KPN's Mirai notification content

Guideline	**Observation**	**Possible improvements**
Describe the risk comprehensively	+ The problem is stated - No description of the consequences of the problem - No actionable advice to avoid the consequences	• Describe the possible consequences • Give actionable advice to avoid the consequences
Be concise and accurate	+ The message is short + Sentences are short - Contains technical terms (Telnet, SSH) - No clear action steps - Not encouraging - Email is only in Dutch	• Explain with less technical terms and explain the technical terms if necessary • Explain action steps such that a less tech savvy customers also understands what to do • Explain what to do to remove malware • Explain what to do to prevent future infections • Encourage the customer to take action • Add English translation to email
Follow a consistent layout	+ Logical order + Consistent wording - No headers - No lists	• Chunk the information
Minimize the user's memory load	- Instructions are not easily recognisable	• Create action steps
Improve trust	- No ability to verify the authenticity of the notification	• Add ability to verify the authenticity of the notification

Rationale behind action steps

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.

For people who directly skip to the action steps, it is useful to repeat that the Mirai malware mainly infects IoT devices. The word ‘mainly’ is used, because researchers have indicated that Mirai malware can also be spread by infected Windows devices (Kaspersky Lab, 2017).

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.

Since the device has been accessed and infected by a brute force attack, using a built-in dictionary of common usernames and passwords, the password has to be changed to prevent reinfection. There might, however, be a limitation when a customer is in a walled garden environment. Customers need to know the current (default) credentials in order to change password. If a customer does not know the credentials and neither has the manual of the manufacturer nor access to mobile internet, the customer cannot change the password.

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.

It is advised to restart the device. As Mirai’s bot only exists in the memory, the bot disappears when the device is restarted (Cao et al., 2017).

The obvious difference between a walled garden and email notification is the internet blockade. When a customer is in a walled garden environment, devices cannot get reinfected after restarting the device. However, if a customer is notified by email, there is a risk of reinfection. Therefore, it is critical that the customer changes the password of the device before restarting it.

4. Reset your modem/router to the factory settings. On https://forum.kpn.com/internet-9/hoe-reset-ik-de-kpn-experia-box-modem-97446 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

Since, certain devices have hard coded credentials, which cannot be changed by the customers, it is important to take additional measures to prevent future infection (Cao et al., 2017). By resetting the modem/router to factory settings, all ports are closed and the demilitarized zone (DMZ) and universal plug and play (UPnP) option are disabled on the router. Below, the need for these steps are briefly discussed.

Closed ports

It is necessary to open ports if you want to remotely access devices on your home Internet connection. However, hackers could use ports that are open to the wider Internet to remotely install the malware on the device. The first version of Mirai scanned for an open port 23. Newer variants of Mirai scanned for other open ports as well.

DMZ

The DMZ option allows for a device to bypass the firewall setting of the router (see Figure 11). This is occasionally necessary when you are using a device that has its own firewall configuration. However, employees of the KPN abuse team indicate that some customers with IoT devices use this option out of convenience. A customer might use this option if he wants remote access to a device on their home internet connection, for example viewing his live security camera footage of his house at work. The DMZ option makes the device accessible from the Internet, without having to manually define port forwards. It is problematic when the device in the DMZ does not use its own firewall. Then, it is extremely vulnerable to attack as all ports are exposed to the wider Internet (Virgin Media, 2017).

UPnP

Just like the DMZ option, customers could use the UPnP out of convenience. The UPnP option enables the IoT devices, and other Internet connected devices, to open the ports in the router it needs in order to be accessible from the external network. The customer cannot control which ports are open.

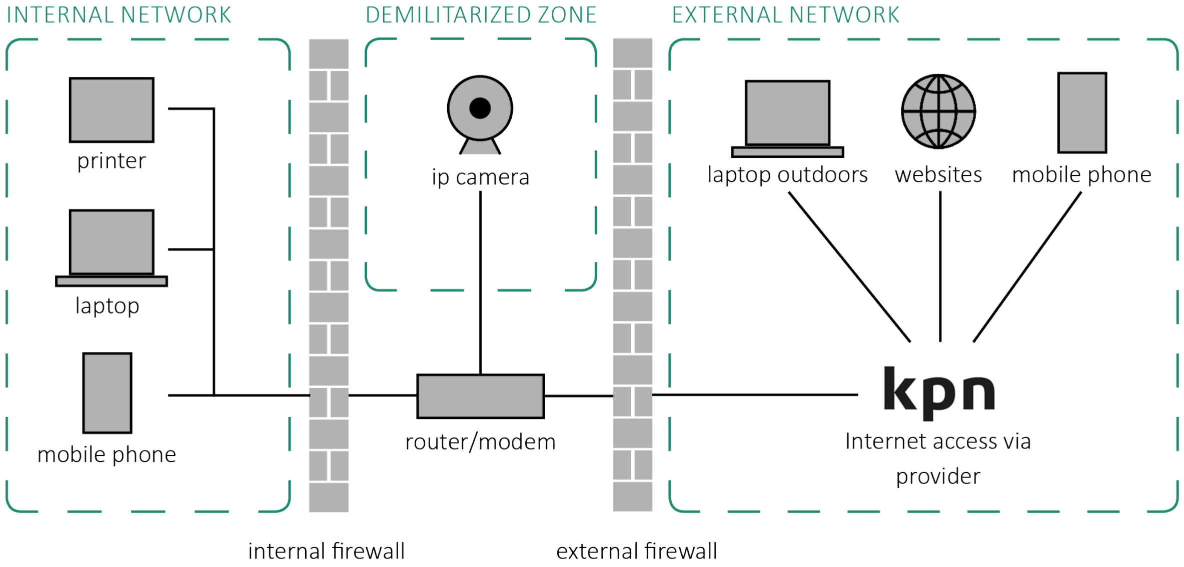


Figure 11 Schematic representation of internal and external network with DMZ

4.5.5 LIMITATIONS EXPERIMENTAL DESIGN

The experiment has several limitations. The influence of a more actionable walled garden notification content on IoT malware cleanup is determined by comparing the treatment groups with new and old content. The notifications are, however, sent during different periods and therefore the average infection duration in the treatment groups could be influenced by outside influences. The infection time could for example increase due to a holiday period, when customers might use their computers less, or because a C&C issues more commands to its bots in a certain period. Moreover, in periods when DDoS attacks that relied upon IoT devices are in the news, customers might be more aware of the urgency to take action. Moreover customers might better understand the content of the notification messages.

Secondly, the standard procedure of the KPN abuse team is followed. Besides the new more actionable notification message, the KPN Abuse team is responsible for the subsequent communication with the customer. They determine the content of the messages to customers. Furthermore, the abuse team works from Monday to Friday. Therefore, customers are only notified during working days and no emails from customers are answered outside working hours. Customer could contact the KPN help desk. The quality of the advice of the help desk could, however influence the duration of the infection. When a help desk employee does not immediately recognize that the customer is dealing with an abuse problem, this might extend the infection duration.

Thirdly, customers are notified only when they appear with a Mirai infection in the Shadowserver feeds. If the feed is incomplete, when Shadowserver could not identify the infection type as Mirai or when KPN does not receive any reports from Shadowserver, infected customers might not receive a notification. This absence may cause us to overestimate the cleanup rate. Moreover, the Shadowserver reports only include the first event for each IP address on a day, even though there could be multiple events. Therefore, we could underestimate the infection time. However, it affects the infection time of the customers in the different treatment groups in the same way.

4.6 ANALYSIS OF THE CUSTOMER REACTIONS

In order to investigate the reactions of customers to different IoT malware notification mechanisms (SQ 7) customer interviews are carried out and communication data of the customers in the different experiment groups is collected and analysed. We present the design of the customer interviews in Section 4.6.1. In Section 4.6.2, it is described how the customer interviews and communication logs are evaluated. In this chapter we do not discuss the data collection procedure of the communication logs, as this requires knowledge of the system the abuse team uses. Nevertheless, for the sake of the traceability of data collection, we discuss the procedure in Appendix N.

4.6.1 CUSTOMER INTERVIEWS

To investigate how customers experience the different IoT malware notification mechanisms (SQ 7) telephone interviews with customers are carried out. In this section, the interview design is presented.

Interview population

The interview population consists of all customers in the treatment groups of the experiment. To obtain the largest possible sample size, all customers that received a notification in the experiment are called. In case a customer does not answer, the customer is called again on another day. If the customer has not responded after three times, no more attempts are made. Customers who have terminated their contract are not interviewed. Moreover, customers can choose not to cooperate.

Rationale behind interview questions

The aim of the customer interviews is to investigate how the customers perceived the different IoT abuse notification mechanisms and to determine the perceived computing expertise of the customers. As described in Section 4.2, there is in particular a focus on certain aspects of the customer reaction: trust, understanding, satisfaction and suggestions for improvement. Therefore, a set of questions is composed, which are subdivided in related categories: acknowledgement, understanding, computing expertise and suggestions. A brief explanation is provided for the questions in each category. An overview of the complete interview protocol and flow chart is presented in Appendix K.

Acknowledgement

1. Do you remember receiving the message?

Before, questions are asked about the notification mechanism, it is important to know whether the customer actually remembers receiving the message. If the customer does not remember receiving the message, it is checked if the message was sent to the correct email address.

2. Do you remember reading the message?

When the customer has confirmed to remember receiving the notification, it is important to know whether the customer remembers reading the message and if not, why not. Here customer could point out a lack of trust.

Understanding

3. Did you take any action after reading the message, if so, can you please explain to me what you did?

To examine the level of understanding about what to do after the notification, the interviewee is asked who has taken action after the notification and what these actions were. Based on this answer, it can be determined if the customer was able to identify the infected device and subsequently took the right actions.

4. While cleaning the virus, have you needed any additional materials such as searching on Google, assistance of someone else or searching in the paper manual?

Additionally, the level of understanding about what to do is determined by whether or not additional help was needed.

Computing expertise

5. How confident are you about your technical ability to solve issues like this one?

Customer can answer on a four-point scale how confident they are about their ability to solve issues like this one. This question is adapted from the Special Eurobarometer 390 Cyber security (European Commission, 2012).

Suggestions

6. How could the communication to customers be improved when KPN see problems like this? What are your suggestions?

By means of this last question the customer has the ability to indicate what could be improved. These suggestions can be used as a recommendation for KPN. Furthermore, the answer to this question could reflect the customer's level of satisfaction with the notification mechanism.

Wording

A number of choices have been made regarding the formulation of the introduction and questions to ensure that customers answer as honestly and as completely as possible. It is assumed that it is important that the customer feels comfortable answering the questions, and certainly does not feel stupid. To this end, technical jargon is avoided. Technical terms are replaced by phrases or expressions that might be better understood by the interviewee. For example, Mirai and malware have been replaced by virus. Furthermore, the introduction states: "KPN notices that sometimes customers find it difficult to understand this message, because it includes technical information." The purpose of this sentence is to ensure that a customer, who did not understand the message, is more inclined to truthfully admit his misunderstanding of the message, without being ashamed. Lastly, to increase the number of respondents, it is stated in the introduction that the interview will take less than five minutes. This is considered a reasonable amount of time to interrupt a customer.

Phone calls

The interviews are conducted on working days between 2 and 7 p.m. The phone calls are made from a private number, as this is KPN's regular procedure when they call customers for marketing purposes.

Limitations customer interview design

The customer interviews described in this section have several limitations. Firstly, customers who have terminated their contract are not called. These customers might have a strong opinion regarding the notification mechanism. This opinion and their suggestions are, however, not included in the interview results.

Secondly, a customer in the email treatment group who was still infected after 14 days, is placed in quarantine. In the interview, questions are asked about the email notification specifically. Nevertheless, the customer might confuse the different notification mechanisms while answering the questions.

Third, the period between the notification and the interview differs for the different customers, as the end of the experiment almost coincides with the end of the research project. This could influence how well the customer remembers receiving and reading the notification message.

4.6.2 EVALUATING COMMUNICATION LOGS AND CUSTOMER INTERVIEWS

To analyse the customer reaction systematically, help and abuse desk logs and customer interviews are labelled. To this end, we examined 20% of the help and abuse desk logs and all customer interviews. Afterwards, various categories were created based on recurring themes. After defining these labels, we manually labelled all the communication logs and interviews. In this section, it is described what these categories entail.

Understanding

Understanding indicates whether or not a customer understands the content of the notification message. This is assessed based on three aspects, which are described below.

- *Actions:* this represents the actions the customer took after the notification.
- *Additional help:* this indicates whether or not the customer needed additional help to solve the issue. Here a distinction is made between a paid KPN technician, requesting to call with the abuse team, the help desk, general help, searching with Google, someone else's help, paper manual and the supplier of the device. General help refers to customers who indicate to have "no idea what to do".
- *Identify device:* specifically for Mirai, it is interesting to know if the customer understood that the notification concerned an IoT device. Therefore, it is investigated whether the customer is talking about devices, such as a computer, laptop or tablet, or refers to an IoT device.

Trust

Trust refers to customers who do not trust the notification messages, because they think it is either spam or a phishing email.

Satisfaction

This category indicates how satisfied customers are with the notification mechanism. Next to the customers who indicate to be very satisfied, other customers are angry. A distinction is made between the angry customers, which is explained below.

- *Satisfied:* customer that explicitly expressed their content with KPN's notification procedure.
- *Cannot work due to quarantine:* customers who indicate to make losses because of KPN's walled garden procedure, as they cannot use their pin device or cannot run their business without Internet.
- *Terminate contract:* customers who say they will switch to a different provider because of the walled garden procedure.
- *Complaints over disruption services:* customers indicating that they have lost their patience, are angry and cuss.

4.7 STATISTICAL ANALYSIS

In this section the statistical analysis methods are explained.

4.7.1 KAPLAN-MEIER SURVIVAL ANALYSIS

A survival analysis is performed to examine the time to a certain event, which is cleanup and reaction in this study. The Kaplan-Meier estimator is used to estimate the survival function that gives the probability that a customer will survive beyond any specified time. The survival probability at any particular time is calculated by:

$$S(t)=\frac{\textit{Number of subjects infected at the start-Number of subject cleaned up}}{\textit{Number of subjects infected at the start}}$$

$$=\frac{\textit{Number surviving}}{\textit{Number at risk}}$$

The total probability of survival till a time interval is calculated by multiplying all the probabilities of survival at all time intervals preceding that time.

The Kaplan-Meier analysis is a non-parametric approach, which means that the events times do not have to have a normal distribution. Moreover, an important advantage of the Kaplan-Meier curve is that the method can take into account some types of censored data, particularly right-censoring. This happens in this study, when the customer is still infected at the last day of the 14 day observation period, or in other words, when no event has occurred (Goel, Khanna, & Kishore, 2010).

Log-rank test

To test if the survival curves of two (or more) groups significantly differ, the log-rank test is used. This test compares the observed differences between the survival curves, with differences that could occur when there is no difference between the groups. For each event time, the number of observed events (O) in each group is calculated. Additionally, the number of expected events (E), if there would be no difference between the groups, is calculated. The same calculations are repeated each time an event occurs. The test statistic is:

$$\textit{Log-rank test statistic}=\frac{(O_1-E_1)^2}{E_1}+\frac{(O_2-E_2)^2}{E_2}$$

where, E and E2 are the expected number of events in each group, and O1 and O2 are the total number of observed events in each group.

The pairwise log-rank test is purely a test of significance, it does not provide an estimate of the size of the difference between the groups or a confidence interval (Bland & Altman, 2004).

4.8 ETHICAL CONSIDERATIONS

The research aims at improving the notification mechanisms for compromised IoT devices. The only valid method to analyse the effect of IoT malware notifications to resource owners, is testing it in a life-like setting. There might be objections. People in the control group get no treatment, while the infection is known. Furthermore, customers do not know that they are part of an experiment. Therefore, to minimize the potential negative impact of the IoT malware infection, the experiment period is limited to 14 days. If Shadowserver reports the IP address again after the observation period, the customer is notified via KPN's current IoT malware notification procedure.

The research approach described in this chapter is possible because of the cooperation with KPN. Unavoidably, the data used in this study contains confidential information. Therefore, the data was only stored on a laptop provided by KPN. Moreover, it is ensured that is not possible to identify a customer from any of our results.

5 RESULTS OF THE EXPERIMENT

The experiment described in Section 4.5 has been conducted at KPN's abuse department to answer sub-research questions 5 and 6: (5) what is the added value of quarantining in terms of IoT malware cleanup?; and (6) what is the influence of a more actionable walled garden notification content on IoT malware cleanup? In this chapter, we describe the course of the experiment in Section 5.1 and the results in Section 5.2. Based on these results we draw our conclusions and provide an answer to sub-questions 5 and 6 in Section 5.3.

5.1 COURSE OF THE EXPERIMENT

As explained in Section 4.5.3, the experiment consists of two stages. The first stage consists of the analysis of historical data for the treatment group walled garden with old content. The second element is a randomized controlled experiment for the other three treatment groups: (1) walled garden with new content; (2) email with new content; and (3) control group. In Section 5.1.1 and 5.1.2 the course of the experiment and the sample sizes of the different treatment groups are explained.

5.1.1 HISTORICAL DATA COLLECTION

The walled garden with old content treatment group consists of the last 100 unique customers who were placed into quarantine and received an email from the abuse team before October 11th, 2017. In this group, the first notification was sent on June 20th, 2017. Throughout this data collection period, 3 customers were released when the quarantine period expired after 30 days. However, customer interviews (described in Chapter 6) demonstrated that the quarantine events did not properly work. Therefore they were capable of surfing the Internet without any limitations during the quarantine period. As a result of this, these customers have been removed from the analysis. This resulted in a sample size of 97 customers in the walled garden with new content treatment group.

5.1.2 RANDOMIZED CONTROLLED EXPERIMENT

From November 5th, 2017 until February 14th, 2018, a total number of 117 consumer market customers with a Mirai infection were reported by Shadowserver. The number of customers per day fluctuated with a maximum of 12 customers per day, which is shown in Figure 12. The median number of unique customers per day was 4, the median number of new unique customers per day was 1.

Of these 117, 28 customers were not part of the experiment for various reasons:

- 19 customers only appeared during the weekend and were therefore not notified. Of these, 18 customers appeared only once during the weekend. Moreover, a single customer was observed in the Shadowserver reports for 3 consecutive weekends,
- 7 customers were removed from the analysis, as they were accidently placed in quarantine during the 14 days observation period due to another infection or vulnerability,
- 2 customers were in treatment group walled garden with old content.

4 customers in the control group were still infected after the observation period of 14 days. These customers were randomly assigned to either the walled garden or the email treatment group after the first experimental period. Therefore, these customers were part of another treatment group after being previously in the control

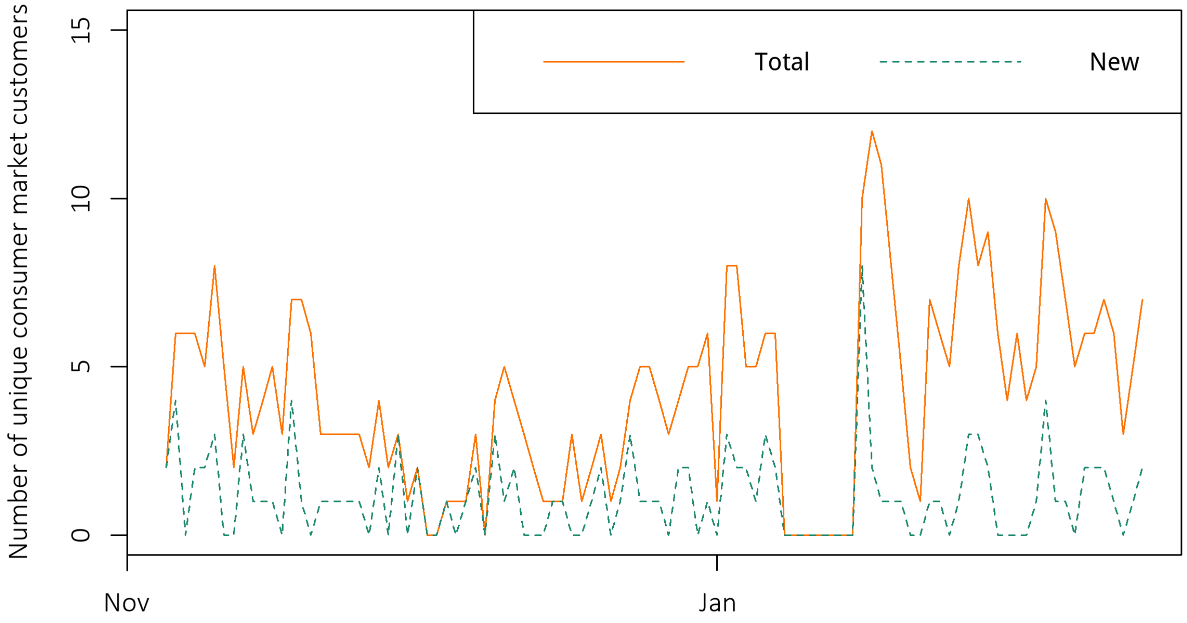


Figure 12 Number of unique consumer market customer with Mirai infection in Shadowserver feeds between November 5, 2017 –February 14, 2018

group. As a result of this, the sample sizes were: 33 customers in the control group, 30 customers in walled garden with new content and 30 customers in the email with new content treatment group.

From Figure 12 it can be seen that Shadowserver did not send a Botnet-Drone report for 8 consecutive days (from 8th to 15th of January). Therefore, customers could not be notified during this period. In order to have an experimental period of 14 days after the notification, the experiment was paused and resumed on January 16th, 2018.

This is illustrated with Figure 13. For example, a customer received its first email notification on January 3rd, 2018. Normally, the experimental period would last till January 17th, 2018. However, in this special case, the customer is tracked (and notified when still infected) for four days till January 7th, 2018. After resuming the experiment at January 16th, the infection is tracked for 10 additional days, till January 25th, 2017.

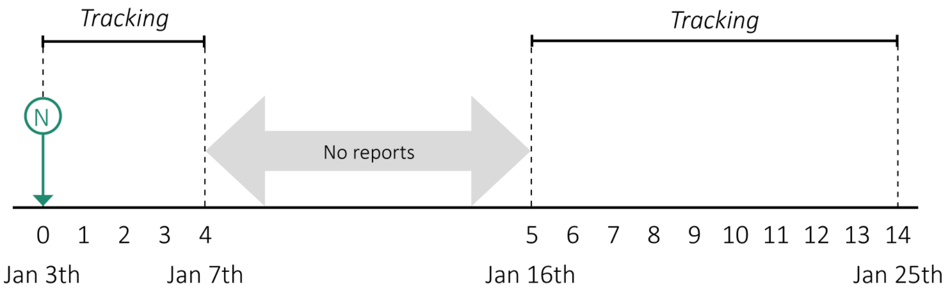


Figure 13 Example of exception tracking procedure

5.2 RESULTS

In this section, the results of the experiment are described and interpreted. First, we evaluate the impact of the experimental groups on cleanup by looking at the cleanup rates and the infection time in Section 5.2.1. As described in Section 4.5.3.3, we distinguished two outcomes to evaluate the infection time: (1) the customer successfully performed cleanup (is removed from quarantine) and stays clean for the rest of the study period; and (2) the user did not successfully address the problem, as evidenced by the repeated appearance of the customer's IP address in the Shadowserver or IoTPOT reports during the study period. The cleanup rate is defined as the percentage of IoT devices that were clean, 14 days after the initial notification.

In Section 2.6, 3 hypotheses were formulated regarding the influence of the IoT malware notification mechanisms on cleanup: (H1) any kind of IoT malware notification reduces the duration of the infection, (H2) walled garden notifications reduce the duration of IoT malware infections compared to email notifications and (H 3) a more actionable walled garden notification message reduces the duration of the IoT malware infection. In the following sections, we discuss these hypotheses.

5.2.1 HYPOTHESIS 1: THE IMPACT OF THE DIFFERENT IOT MALWARE NOTIFICATIONS ON CLEANUP

First, it is determined whether the different IoT malware notifications reduce the duration of the infection, by comparing the control group with the other treatment groups. Table 9 presents some summary statistics regarding the percentage of IoT devices that were no longer infected 14 days after the initial notification and the median infection time for each experimental group. Interestingly, the email treatment group has the lowest cleanup rate (76.7%) of all the treatment groups. The difference between the email treatment group and the control group (78.8%) is small. Furthermore, the median infection time of the control group (40 hours) is comparable with the email group (41 hours). The median infection time is shorter for the walled garden with old and new content treatment groups: 27 and 16.5 hours respectively. Based on this table, only the walled garden treatment groups seem to have an impact on the infection time. To further assess whether the differences are significant, the survival probabilities are computed for the different treatment groups (see Figure 14). Looking at the survival graph, we see similar cleanup rates 1 day after the initial notification for all 4 treatment groups. In contrast, 5 days after the initial notifications we see notable differences in the cleanup rates among the treatment groups. After 5 days, we see cleanup rates of 90% and 68% for the customers that were quarantined and received the new and the old content respectively. Of those who received the email notification and no notification, about 57% cleaned up. The log-rank test shows that the difference between the control group and the walled garden with new content treatment group is significant (χ^2 = 4.3, p = 0.0376). However, the differences between the control group and the other treatment groups are not significant (see Table 10). The Cox proportional hazard model is used to compute the hazard ratios (HRs) for the different treatment groups. Table 11 shows the HRs of each of the groups that received notifications versus the control group. From this table it can be derived that the group that received the walled garden with the new content treatment achieved a 1.7 times faster cleanup rate than the control group.

Contrary to what was expected based on the literature, only the walled garden notification with new content has a significant impact on the infection period. From this we can conclude that not all notifications make measurable differences. It should be noted that the notification campaigns for the old content and the campaign for the new content and control group took place in different periods. Therefore, there might be a behaviour change of the Mirai botnet between the different periods. As a result of this, this approach lacks diagnostic information about how exactly the malware behaviour influences infection time.

Table 9 Summary statistics cleanup rate and infection time per treatment group

Treatment group	# of customers	% clean after 1 day	% clean after 5 days	% clean after 14 days	Median infection time
Control	33	45.5%	57.6%	78.8%	40 hours
Email with new content	30	50.0%	56.7%	76.7%	41 hours
Walled garden with new content	30	60.0%	90.0%	96.7%	16.5 hours
Walled garden with old content	97	49.5%	68.0%	88.7%	27 hours

Table 10 Log-rank test results cleanup rates

Comparison		χ^2 (df =1)	p-value
Control	Email with new content	0.0	0.92
Control	Walled garden with new content	4.3	0.0376
Control	Walled garden with old content	0.6	0.437
Email with new content	Walled garden with new content	4.0	0.0453
Walled garden with old content	Walled garden with new content	4.2	0.0393

Table 11 Cox proportional hazard test results: hazard ratios per treatment group versus control group

Treatment group	Coef.	HR	95% CI
Email with new content	0.002	1.00	0.57 – 1.76
Walled garden with new content	0.529	1.70	0.99 – 2.92
Walled garden with old content	0.148	1.16	0.75 – 1.80

5.2.2 HYPOTHESIS 2: THE ADDED VALUE OF A WALLED GARDEN NOTIFICATION IN TERMS OF CLEANUP

In order to examine the added value of a walled garden notification, the email with the new content and walled garden with new content treatment groups are compared. Table 9 shows a higher cleanup rate and shorter median infection time for the walled garden with new content treatment group compared to the email group. Again, to assess whether the differences between the groups are significant, the survival probabilities are computed for the different notification channels (see Figure 14). A log-rank test (see Table 10) concluded that the treatments were significantly different (χ^2 = 4.0, p = 0.0453). Based on these findings, we can conclude that, when the contents are the same, walled garden notifications are more effective in reducing the infection period than email notifications.

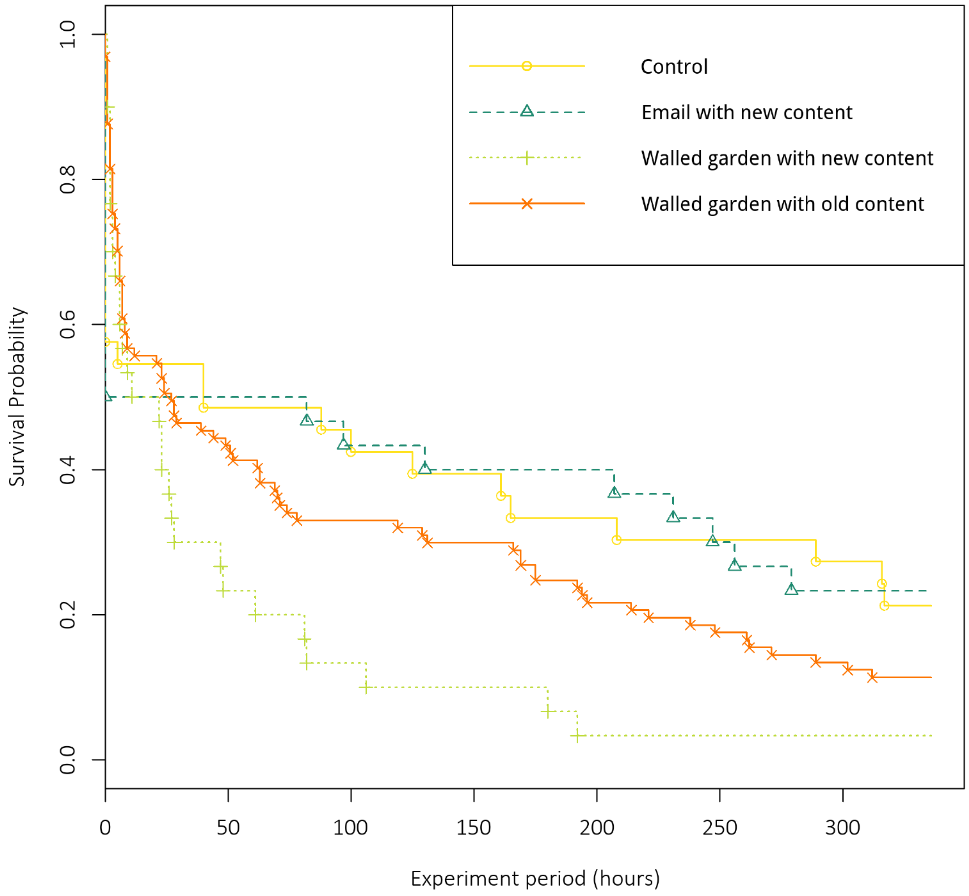


Figure 14 Survival probabilities cleanup per treatment group

5.2.3 HYPOTHESIS 3: THE IMPACT OF A MORE ACTIONABLE CONTENT ON CLEANUP

To investigate the influence of a more actionable notification content, walled garden treatment groups with the new and old content are compared. Table 9 shows a higher cleanup rate and a shorter the median infection time for the walled garden with new content treatment group compared to the walled garden with old content treatment group. This supports the hypothesis that a more actionable walled garden notification message reduces the duration of the IoT infection. Moreover, the log-rank test (see Table 10) corroborated that the treatments were significantly different (χ^2= 4.2, p = 0.0393). Based on these findings, we can conclude that a more actionable walled garden notification content reduces the duration of an IoT infection. Again, it should be noted that the walled garden notifications with the new and the old content were performed in different periods of time.

5.3 CONCLUSIONS

The defined goal of this chapter is to answer sub-question 5 and 6: (5) What is the added value of quarantining in terms of IoT malware cleanup?; and (6) what is the influence of a more actionable walled garden notification content on IoT malware cleanup? In this section, an answer is provided to each of the sub-questions.

In the experiment, we compared the influence of email notifications and a combination of email and walled garden notifications in terms of infection time and cleanup rates (SQ 5). The notification content of the both treatment groups was the same. We found that customers that were placed in quarantine had a significantly shorter infection time compared to customers that only received an email notification. Based on these results we can conclude that quarantining improves the IoT malware cleanup. In addition, the difference in infection time between the control group and each of the treatment groups was investigated. Based on literature one would expect that all notifications would reduce the infection time (see Section 2.4). However, in our experiment we only found a significant difference in terms of infection time between the group that received a walled garden notification with the new content and the control group. From this we can conclude that not all notifications make measurable differences.

We also investigated how a more actionable walled garden notification message influences the IoT malware cleanup (SQ 6). To this end, a new more actionable notification message was composed based on guidelines from literature (see Section 4.5.4). To measure the influence of a more actionable content, infection times of the customers that were placed in quarantine with the old and the newly composed notification messages were compared. As a result of this comparison, we found a significant difference, at a 0.05 significance level, in terms of the infection time between the different walled garden notification contents. The group that received the more actionable walled garden notification cleaned up faster. Based on these results we can conclude that a more actionable walled garden notification content improves the IoT malware cleanup.

6 ANALYSIS OF THE CUSTOMER REACTIONS

In order to evaluate the reactions of the customers to the different notification mechanisms for IoT malware infections (SQ 7), we used two data source: (1) interviews with the customers in the experiment; and (2) help and abuse desk communication logs of the customers in the experiment. First, we describe both data sources in more detail in Section 6.1 and 6.2. Thereafter, in Section 6.3 and 6.4, we discuss the results of the quantitative and qualitative analysis of the customer reactions after an IoT malware notification. Lastly, we provide an answer to sub-question 7 in the conclusions in Section 6.5.

6.1 CUSTOMER INTERVIEWS

All 157 customers that were notified during the experiment have been contacted for a telephone interview afterwards. Of them, 71 customers (45.2%) were interviewed. Other customers did not want to contribute (17, 10.8%), terminated their contract in the meantime (4, 2.5%) or were not available by phone (65, 41.4%). Among the customers that did not want to contribute, one customer did not trust the phone call. This customer made the comparison with the Microsoft scam. Of the customers that were not available, for 6 of them (3.8%) the phone number provided at subscription is currently not in use. The other 60 customers (37.6%) did not answer their phone after calling them 3 times. Table 12 shows the distribution of the interviewed customers over the treatment groups.

Table 12 Number of interviewed customers per treatment group

Treatment group	# of customers	# of customers interviewed
Email with new content	30	14 (46.7%)
Walled garden with new content	30	13 (43.3%)
Walled garden with old content	97	44 (45.4%)

6.2 ABUSE AND HELP DESK COMMUNICATION LOGS

All communication logs between the abuse and help desk and the customers have been collected. A total number of 158 contact forms, 362 emails and 103 help desk logs are investigated (see Table 13).

Table 13 Summary of messages and number of unique customers per communication channel

Communication channel	# of total logs	# of unique customers	# of logs per customer		
			Min	*Med*	*Max*
Contact form	158	82	1	1	10
Email	362	98	1	2	41
Help desk	103	60	1	1	6

It is explored which communication channel customers use to contact KPN after receiving a walled garden notification. It is found that only about a third of the customers who are placed in walled garden directly submits the walled garden form to use the self-release option (see Table 14). Moreover, among rest of the customers in the walled garden with new content treatment group, about 17% sent an email to the abuse team and about 43% contacted the help desk. Comparable results have been observed in the walled garden with old content treatment group. Of these, 29% of the customers emailed the abuse team and 38% contacted the help desk.

Each communication channel was used for different reasons. Generally, emails were sent to inform abuse desk employees about the cleanup efforts and possible causes of the infection. The content of the submitted walled garden forms often contained more specific information on the cleanup actions taken by the quarantined customers. On the other hand, customers contacted the help desk employees, mainly to ask for more information about the quarantine and how to resolve the situation. Moreover, customers contacted the help desk outside working hours of the abuse team.

In addition, we examined which communication channels have been used by the customers that received the email notification. We found that about a third of the customers replied to the email notification. Interestingly, while about 40% of the quarantined customers called the help desk, only 1 customers that received an email notification contacted the help desk.

Table 14 Communication channel used by customers per treatment group

Treatment group	# of customers	Contact form	Email	Help desk
Email with new content	30	-	33.3%	3.3%
Walled garden with new content	30	40.0%	16.7%	43.3%
Walled garden with old content	97	33.0%	28.9%	38.1%

6.3 QUANTITATIVE ANALYSIS OF THE CUSTOMER REACTIONS

As described in Section 4.3, the customer reactions are evaluated quantifiably based on two metrics: (1) the median reaction time; and (2) the reaction rate. The reaction time is defined as the time (in hours) between the first notification and the first reaction from the customer to KPN. The median reaction time is based on the reaction times of the customers that actually reacted. The reaction rate is the percentage of customers that respond to the notification within 14 days after the initial notification. For this analysis we only used the abuse and help desk logs.

In Section 2.6, two hypothesis were formulated regarding the influence of the IoT malware notification mechanisms on the reaction time of the customers: (H 4) Customers respond faster to a walled garden notification compared to an email notification; and (H 5) Customers respond faster to a more actionable walled garden notification message. In the following, we discuss these hypotheses.

6.3.1 HYPOTHESIS 4: THE ADDED VALUE OF WALLED GARDEN NOTIFICATIONS ON REACTION TIME

To examine whether customers react faster to a walled garden notification compared to an email notification, the walled garden treatment group and email treatment group with the same content are compared in terms of reaction time. Table 15 provides a summary statistics regarding the median reaction time and the percentage of customers that reacted to the notification within the 14 days after the initial notification. Table 15 shows a clear difference between the treatment groups in terms of reaction times and reaction rates. Customers that received an email had a median reaction time of 47 hours, whereas the customers that were also placed in quarantine had a median reaction time of 6 hours. When we investigate the survival curves in Figure 15, we see a big difference in the reaction rates across the 2 treatment groups. Almost 90% of the customers that were placed in quarantine reacted within 50 hours after the initial notification, while only 20% of the emailed customers reacted within this period. At the end of the experimental period, all quarantined customers reacted. In contrast, only 36.7% of the customers in the email group contacted KPN. A log-rank test (see Table 16) confirmed that the differences in reaction rate between the treatment groups were significant ($\chi^2 = 45.7$, $p = 1.39{\cdot}10^{-11}$). Based on these findings, we observe that customers who receive a walled garden notification reacted much faster than the ones that received an email notification. This can be explained by the fact that customers have to contact KPN to get out of the walled garden. After all, they have to send the contact form for self-release or ask KPN employees, via email or help desk, for assisted release. On the contrary, when a customer receives the email, the customer does not have to contact KPN. The additional value of the walled garden notification is that KPN knows if the customer has received and read the notification.

6.3.2 HYPOTHESIS 5: THE IMPACT OF AN ACTIONABLE CONTENT ON THE REACTION TIME

To investigate the influence of a more actionable walled garden notification content on the reaction time, the walled garden treatment groups with the new and old content are compared. Looking at Table 15, hardly any differences can be seen in terms of reaction rates and reaction times. Both groups have 100% reaction rates and a median reaction time of 3 and 6 hours. The log-rank test results are presented in Table 16. This test indeed confirms that there is no statistical difference between the groups receiving messages with different levels of actionability in terms of reaction time ($\chi^2 = 0.3$, $p = 0.561$). Based on these findings, we can conclude that a more actionable walled garden notification message does not make a measurable difference in getting customers to react faster.

Table 15 Summary statistics reaction rates and reaction times per treatment group

Treatment group	# of customers	% reaction after 1 day	% reaction after 5 days	% reaction after 14 days	Median reaction time
Email with new content	30	16.7%	33.3%	36.7%	47 hours
Walled garden with new content	30	83.3%	96.7%	100%	6 hours
Walled garden with old content	97	84.5%	96.9%	100%	3 hours

Table 16 Log-rank test results reaction times

Comparison		χ^2 (df =1)	p-value
Email with new content	Walled garden with new content	45.7	$1.39{\cdot}10^{-11}$
Walled garden with old content	Walled garden with new content	0.3	0.561

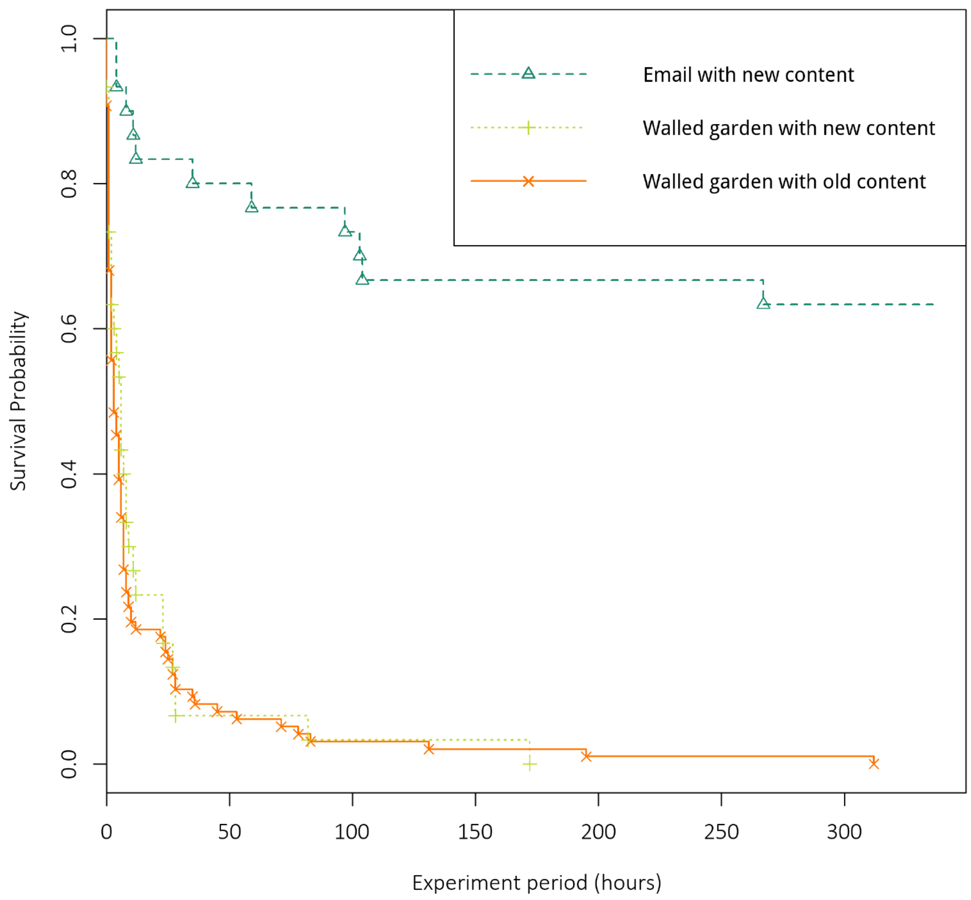


Figure 15 Survival probabilities reaction per treatment group

6.4 QUALITATIVE ANALYSIS OF THE CUSTOMER REACTIONS

To get a better sense of the actual experience of customers receiving an IoT malware notification, the communication of the customers with the abuse and help desk and the customer interviews are analysed qualitatively on the basis of recurring themes in the conversations that illustrate the customer's experiences: (1) acknowledgement of receiving the notification; (2) misunderstanding the notification; (3) disconnecting the device; (4) distrusting the notification; and (5) voicing complaints. For each customer all communication logs after the initial notification and the answers provided in the interview are merged in order to be able to examine the number of unique users associated with a certain topic.

6.4.1 ACKNOWLEDGEMENT RECEIVING AND READING NOTIFICATION

At the beginning of the interviews, customers were asked if they remembered receiving the notification. Customers who did not remember receiving the notifications were asked if the email was sent to the right email address. They all indicated that the email address used to contact the customer about the IoT malware problem was correct and currently in use. For 2 of the customers in the walled garden with old content treatment group who did not remember receiving the message the quarantine period expired. They say having used their Internet connection during the 30 days quarantine period. This could indicate that the walled garden system did not work as intended. Therefore, the customers for which the quarantine period expired have been removed from the

Table 17 Summary acknowledgement of receiving and reading the notification based on customer interviews

	Email with new content	**Walled garden with new content**	**Walled garden with old content**
	n=14	*n=13*	*n=44*
Remembers receiving the notification	7 (50%)	13 (100%)	40 (90.9%)
Remembers reading the notification	5 (35.7%)	13 (100%)	36 (81.8%)
Distrusts the notification	2 (14.3%)	0 (0%)	6 (13.6%)

experiment. Customers that received an email notification, who did not remember receiving the notification argue that they receive many emails and therefore only pay attention to emails from KPN regarding payments.

As can be seen in Table 17, not all customers that received the message, actually read the email. A customer, that received an email notification, argued that he did not read the message because he only reads emails from KPN about special offers. Other customers did not read the email due to a lack of trust in the message. Moreover, 6 customers (13.6%) in the walled garden with old content treatment group did not trust the email that was sent along with the quarantine action. The trust issue will be discussed in more detail in Section 6.4.4.

6.4.2 PERCEIVED COMPUTING EXPERTISE

Interviewees are asked how confident they are about their ability to solve issues like a Mirai malware infection. Out of 71 interviewed customers, 25 indicated to be very confident in their technical ability to solve issues like this one, because they work as an IT professional. On the contrary, 15 interviewees stated to know nothing about the matter and to have no confidence at all. This group indicated that they were also not willing to put effort into understanding the problem and therefore, always ask for help. Some of the interviewees who stated having no confidence at all argued being too old for such problems.

Table 18 Perceived computing expertise of the interviewees

Perceived computing expertise	**# of customers**
Don’t know	1 (1.4%)
Not at all confident	15 (21.1%)
Not very confident	10 (14.1%)
Fairly confident	20 (28.2%)
Very confident	25 (35.2%)

6.4.3 UNDERSTANDING

Different subjects in the communication logs and customer interviews give an idea of how well the customers understand the notification message. We analysed the customer’s understanding of the IoT malware notification on the basis of 5 different topics: (1) running a virus scanner; (2) identifying the IoT device; (3) requesting additional help; (4) requesting to call with the abuse team; and (5) requesting a paid technician. Notable findings, related to how well the customers understand the notification message are presented in this section.

Running a virus scanner

About a third of the customers that received a walled garden notification regardless of the content indicated having run a virus scanner to remove the IoT malware (see Table 19). These customers wanted to solve the problem, but they were unable to understand that running a virus scanner on a desktop computer or laptop, would not solve the malware infection on their IoT device. In some cases, customers believe that they have been unjustly quarantined because no infection has been found on their computer after running a virus scanner. Given that comparable percentages of the walled garden with the old and the new content treatment groups have run a virus scanner, we can conclude that the newly composed notification message did not improve the understanding that running a virus scanner does not remove the IoT malware. In contrast, out of the 14 interviewed customers in the email treatment group no one mentioned that they run a virus scanner. Among the 11 customers that reacted to the email notification, 6 customers indicated having scanned for viruses. This corresponds to 20% of the total number of customers that received an email notification. During an expert interview, it is argued that the static walled garden contact form might be misleading for customers with a Mirai infection, as in this form questions such as “How many computers/laptops are connected?” and “Which anti-virus software do you use?” are being asked. We indeed find that a smaller part of customers who received an email notification, and never saw the walled garden contact form, indicated having scanned for viruses. However, seeing that 6 of the 11 emailed customers that reacted (55%) mentioned having scanned for viruses, we cannot conclude that a misleading contact form caused more customers to run a virus scanner.

Identifying IoT device

By reading the notification, some customers were unable to identify the infected IoT device. Therefore, they have contacted KPN employees to gather more information about the infected device. It is found that more than half of the customers who were placed in quarantine and only about 23% of the customers that received an email notification were mentioning a type of IoT device in their message(s) to KPN or during the customer interview. These customers understood that Mirai targets IoT devices. There are, however, examples where customers initially state that they do not have any IoT device. It only becomes clear after additional examples of KPN, and in some cases a network scan, which of their devices are connected to the Internet. These customers have probably not realized when buying and using an IoT device that they were connecting it to the Internet and therefore also making it vulnerable to intruders. We noticed that customers do, in particular, not realize that their DVR or TV decoder is connected to the Internet.

Requesting additional help

About 42% of the customers in the walled garden with old content treatment group contacted KPN for additional help to clean up the IoT malware infection. These customers wanted to solve the problem, but did not understand the notification or were unable to follow the requested steps. Some of the customers stated “having no idea about what to do”, while others needed additional help to identify which of their many devices was causing the problem or wanted to know how to prevent future infections. The percentage of the customers that received the walled garden or email notification with the new content and requested additional help is considerably lower, 3% and 10% of the customers in the email and walled garden with new content treatment groups respectively. From this, it could be concluded that avoiding technical terms and providing guidance by means of action steps in the notification message reduces the customers’ need for additional help.

Possibility to call the abuse team

More than 20% of the walled garden with old content treatment group and about 7% of the email and walled garden with new content treatment groups requested to talk with the abuse team. These customers indicated that the emails from the abuse team were not sufficient to solve the problem. Furthermore, the help desk could not answer their questions in these situations. After all, the help desk employees are no trained to provide support for abuse issues. These customers think that the cleanup process could be accelerated when they would

Table 19 Summary of investigated issues regarding the customer reactions

	Email with new content	**Walled garden with new content**	**Walled garden with old content**
	n=30	*n=30*	*n=97*
Understanding			
Runs a virus scanner	6 (20.0%)	11 (36.7%)	33 (34%)
Identifies IoT device	7 (23.3%)	17 (56.7%)	70 (72.2%)
Requests additional help	1 (3.3%)	3 (10.0%)	41 (42.3%)
Wants possibility to call the abuse team	2 (6.7%)	2 (6.7%)	22 (22.7%)
Requests paid technician	1 (3.3%)	5 (16.7%)	18 (18.6%)
Disconnecting device			
Disconnects device	3 (10%)	15 (50.0%)	49 (50.5%)
Distrust			
Distrusts the notification	2 (6.7%)	0 (0%)	9 (9.3%)
Satisfaction			
Expresses satisfaction	2 (6.7%)	3 (10.0%)	8 (8.2%)
Cannot work due to quarantine	0 (0%)	3 (10.0%)	19 (19.6%)
Complaints over disruption services	0 (0%)	0 (0%)	16 (16.5%)
Threatens to terminate contract	0 (0%)	1 (3.3%)	5 (5.2%)
Suggestions			
More information in the notification	3 (10%)	5 (16.7%)	15 (15.5%)
No quarantine action outside working hours	0 (0%)	0 (0%)	4 (4.1%)
Notification before quarantine action	0 (0%)	2 (6.7%)	6 (6.2%)

have the possibility to call the abuse team. We found a difference in terms of requesting to talk with the abuse team between the treatment groups with the old and new content. About 23% of the customers in the walled garden with the old content treatment group requested to call with the abuse team compared to 7% of the quarantined customers that received the new content. From this, we can again conclude that avoiding technical terms and providing guidance by means of action steps in the notification message reduces the customers' need for additional help.

Requesting a paid technician

We found hardly any difference in the percentages of customers that request a paid technician in the walled garden treatment groups with different messages. Around 17% of the customers were not capable of removing the infection by themselves. They requested the help desk to send a paid technician to their houses or got help from an IT specialist, which they always contact in case of computer related problems. After the visit, customers reported the technician's findings to the abuse team. In one case, the customer communicated that the technician could not find the infection. However, the customer kept reappearing in the Shadowserver reports and was placed in quarantine again. We hardly see any difference between the percentage of customers that requests a paid technician after a walled notification with the new and the old content. This could be explained by the fact that some people are not willing to solve this type of problems by themselves regardless of how simplified (or actionable) the message is. This was pointed out during the interviews by customers with low confidence in their ability to solve problems like an IoT malware infection (see Section 6.4.2).

6.4.4 DISCONNECTING DEVICE

About 10% of the users that received an email notification said to have disconnected their device from the Internet. As shown in Table 19 this percentage is 5 times larger for customers that were additionally placed in quarantine. An explanation for the fact that 50% of the quarantined customers disconnected their device, could be the willingness to take more drastic measures in order to prevent Internet access restrictions in the future.

6.4.5 DISTRUST OF THE NOTIFICATION

About 9% the walled garden with old content treatment group did not trust the authenticity of the email notification they had received. These customers stated that the email looked like a phishing email, because their name was not mentioned and the KPN logo was not included. One customer argued that he did not trust the email address “abuse@kpn.com”, as he was unfamiliar with the term abuse. It was only after seeing the quarantine landing page that these customers trusted that the email notification was actually sent by KPN. They contacted the help desk to confirm the veracity. A customer responded to the email, asking whether the email was send with the right intentions. In addition, 2 customers (6.7%) from the email treatment group, also did not trust the message. One of them indicated having checked the authenticity of the email on the forum of KPN. The information he read there convinced him, so he subsequently took action. None of the customers in the walled garden with new content treatment group distrusted the notification.

6.4.6 SATISFACTION

Different subjects in the communication logs and customer interviews give an idea of the customers’ satisfaction with the notification mechanism. We analysed the customer’s satisfaction on the basis of 4 recurring topics: (1) expressing satisfaction; (2) cannot work due to quarantine; (3) complaints over disruption services; and (4) threatening to terminate contract.

Expressing satisfaction

About 8% of all the customers in the experiment explicitly stated that they were very satisfied with the IoT malware notification. These customers very much appreciate that their ISP has an abuse department that actively notifies its customers in case of Internet connection misuse.

Cannot work due to quarantine

Despite the fact that a walled garden notification reduces the duration of the infection and shortens the reaction time, it is a disruptive measure that leads to customer dissatisfaction. About 20% and 10% of the customers that received a walled garden notification with the old and the new content respectively complained because their business was disrupted due to having no Internet to work with. These were customers that run small businesses, like restaurants, with a consumer Internet subscription. They claimed making losses, because they could not provide services to their customers. A number of customers mentioned that their payment terminals did not work. As a result of this, the customers could not pay using the payment terminals. Others pointed out that their customers could not place orders in their web shop. Therefore, some of the quarantined customers asked for a compensation from KPN.

Complaints over disruption services

In addition to the customers that were angry because their business was disrupted, about 16% of the customers from the walled garden with old content treatment group clearly indicated being very unhappy with the quarantine event because they could not use their Internet connection. A customer complained because he could not monitor his security cameras while in a quarantine environment. This customer stated that he would charge KPN in case of burglary damage. In order to illustrate the level of discontent with the disruptive walled garden notification, a customer even shouted during the interview. Another interviewee described the walled garden notification as a trauma. Additionally, customers expressed their dissatisfaction with the fact that they were placed in quarantine on Friday and could not use their Internet connection for the entire weekend. These customers had to be removed from quarantine by the abuse team, because they did not have a self-release option. However, by the time they came home after work on Friday and saw the quarantine landing page, the abuse team was not available anymore.

Threatening to terminate contract

Around 3% and 5% of the customers in the walled garden treatment groups with the new and old content (see Table 19) were so dissatisfied with the Internet access restriction that they threatened to terminate their subscription and switch to another ISP. In half of these cases, the customers were quarantined multiple times, because they could not identify and remove the infection. Two of these customers have sent the remarkable number of 41 and 26 emails to the abuse team to solve the issue. For the other customers, the infection was quickly cleaned up. They were, however, displeased by the fact that KPN has the right to place a customer in a quarantine environment.

6.4.7 SUGGESTIONS

Some of the customers gave suggestions to improve the notification process. In the following, we describe the different suggestions that customers gave.

More information in the notification

Some customers would like to have more information in the notification message. They would like to know which of their devices is causing the problem. Multiple customers pointed out that it was not clear that the notification message concerned an IoT device. It was only after additional information from the abuse department that the customers realised that the IP camera was infected. These customers would like to know about the possibility of a network scan to identify which of their devices are connected to the Internet right away. Other customers said to have many IoT devices connected to the Internet. They tried to identify the infected machine by turning off devices one by one. When Shadowserver did not report the IP address, while a device was turned off, the infected device was identified. This process could obviously take a lot of time. Interviewees stated that the timestamp of detection and providing information about the information source of KPN would be valuable information for these customers. Furthermore, some customers stated that they would like to know how an infection like this could have happened.

Improve credibility

Customers gave several suggestions to improve the credibility of the notification. They pointed out that the email should include the customer’s name and the logo of KPN. Moreover, customers recommend to remove the links in the message, because this is similar to a phishing emails. One interviewee suggested to use another email address, as he did not know the meaning of ‘abuse’.

No quarantine action outside working hours

Several interviewees stated that customers should not be placed in quarantine on Fridays when the abuse team is not available during weekends. Once a customer has completed the requested steps, the customer must be released from quarantine in their opinion. However, when a customer sees the landing page on Friday outside the working hours of the abuse team, the customer cannot use his Internet connection throughout the entire weekend. That is why the interviewees advise the abuse team either not to quarantine customers on Friday or to be available outside the regular working hours.

Notification before quarantine action
About 6% of the quarantined customers wants to get a notification via email or telephone before being placed in quarantine. They find the partial blockage of an Internet connection a measure too severe, when the customer has not first been given the opportunity solve the problem. They would like to have at least one day after a first warning to solve the problem before being quarantined.

6.5 CONCLUSIONS

The defined goal of this chapter was to answer sub-question 7: what are the reactions of the customers to different notification mechanisms for IoT malware infections? To this end, we quantitatively and qualitatively analysed the communication logs between the abuse and help desk and the customers in the experiment. In addition, the interviews with customers in the experiment are evaluated qualitatively. In this section we describe our main findings.

We evaluated the customer reactions quantifiably by comparing the median reaction time and reaction rate of customers in the email with new content and walled garden with new content treatment groups in the experiment. We found that quarantining significantly improves the reaction rate compared to email notifications. All customers reacted to the walled garden notification, while only 36.7% of customers that received an email notification reacted at all. We did not find that a more actionable walled garden notification message results in a shorter reaction time.

Furthermore, with the communication logs between KPN and customers and the outcomes of the telephone interviews it is explored how the customers perceived the different notification mechanisms. Firstly, when comparing customers that were only notified by email with quarantined customers, we found that quarantined customers read the notification more often and more frequently disconnected their device. However, about 20% of the quarantined customers were dissatisfied with the quarantine event and some of them even threatened to terminate their contract. Secondly we compared the reactions of the customers that received a walled garden notification with the old and the new more actionable content. We observed that a more actionable notification content improved the understanding and reduced the need for additional help considerably. Moreover, it is found that a handful of customers that received notification with the old content did not trust the credibility of the IoT malware notification from KPN, while none of the customers in the walled garden with new content treatment group distrusted the notification.

From these results we can conclude that quarantining improves the reaction time and reaction rate after an IoT malware notification significantly compared to email notifications. In addition, walled garden notifications have a higher probability of being read and more often encourage people to disconnect their device from the Internet. However, in some cases the quarantine event leads to complaints over the disruption. Regarding the notification content, we can conclude that the more actionable content of a walled garden notification does not make a difference in the reaction time and reaction rate compared to a less actionable content of the notification. Though, the newly composed notification content improves the understanding and trust compared to the old notification content.

7 CORRELATING CUSTOMER UNDERSTANDING WITH CLEANUP

In Chapter 5 it is discussed how the different notification channels and messages influence the IoT malware cleanup (SQ 5 and 6). Thereafter, in Chapter 6, we elaborated on the reactions of the customers to the different notification channels and messages for IoT malware infections (SQ 7) . In this chapter, we address the relationship between a specific aspect of the reactions of the customers and IoT malware cleanup success. We investigate the impact of the customer's understanding of the notification on IoT malware cleanup (SQ 8) by measuring the differences in terms of infection time and cleanup rates among customers that do or do not showed an indication to understand the notification message. To this end, we use the experimental results and the outcomes of the qualitative analysis of the communication logs and customer interviews.

In the previous chapter, we analysed the customer's understanding of the IoT malware notification on the basis of 5 different topics: (1) running a virus scanner; (2) requesting a paid technician; (3) requesting additional help; (4) identifying the IoT device; and (5) requesting to call with the abuse team. In Section 7.1 – 7.5 we use these variables related to customers' understanding to correlate the customers' understanding with the customers' infection time and cleanup rates. The results of the analysis are summarized in Table 21 and Figure 16. Lastly, in Section 7.6, we provide the conclusions.

7.1 RUNNING A VIRUS SCANNER

Our analysis of the customer reactions (see Chapter 6) showed that about a third of the customers of both the walled garden and 20% of the email treatment groups wanted to solve the problem, but did not understand that running a virus scanner on a desktop computer or laptop, would not solve the malware infection on their IoT device. To analyse how this misunderstanding influences the infection time, we compared the infection time and cleanup rates of the customers that indicated having run a virus scanner with those who have not. We found that the median infection time for the customers that indicated having run a virus scanner was 68.5 hours, while for the other customers the median infection time was 9 hours. This difference in median infection time could indicate that misunderstanding the notification negatively influences the infection time. To further investigate the issue, we estimated the survival probabilities for customers in the groups (see Figure 17a). After 1 day, 33% of the customers that run a virus scanner has cleaned up the IoT infection, while about 62% of those who did not run a virus scanner cleaned up the infection in this period. However, even though the cleanup rate is notably different during the first days after the initial notification, the cleanup rate after 14 days is about the same. The log-rank test shows that the survival curves are almost but not quite significantly different (χ^2 = 3.1, p = 0.077). Based on these findings, we can conclude that misunderstanding the notification messages increases the infection time. However, given that the differences are almost significant, it should be noted that a larger study population might change this result.

7.2 REQUESTING ADDITIONAL HELP

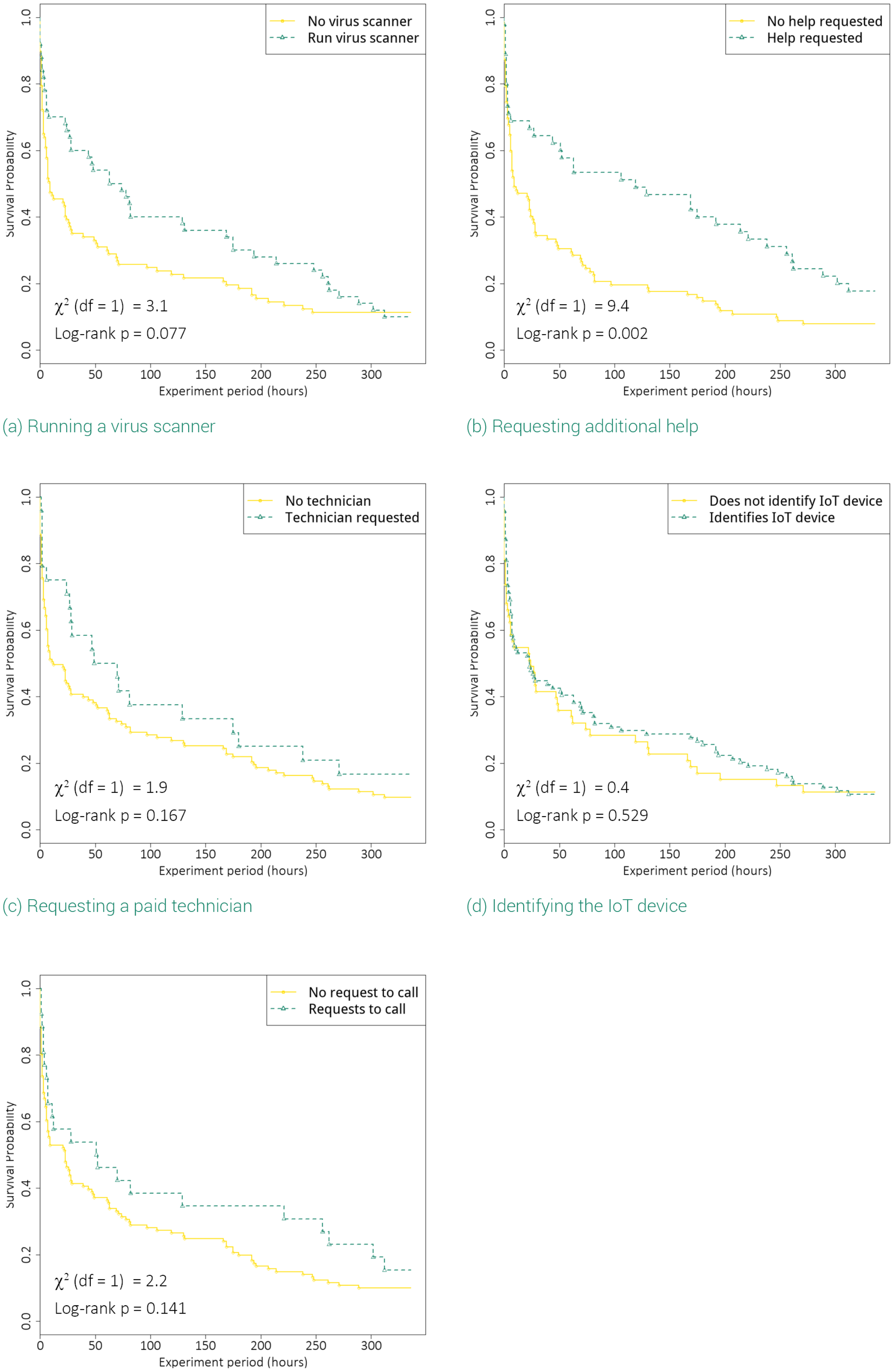
We found that more than 40% of the customers that received the old notification message contacted KPN to ask for additional help to clean up the infection. A notably smaller number of customers that received a walled garden or email notification with the new content requested additional help: 10% and 3% respectively. In order to examine whether the customers that request additional help also have a longer infection time and lower cleanup rates, we compared the infection time and cleanup rates of these customers with those who did not. Figure 17b shows the survival probabilities of both groups. After 1 day about 60% of the customers that did not need additional help cleaned up compared to about 30% of the other group. After 14 days, the cleanup rate differs 10%. About 82% of the customers that requested addition help cleaned up compared to 92% of the other group . The log-rank test results of this comparison, which are also presented in Figure 17b, indeed confirms that there is a statistical difference between the groups ($\chi^2 = 9.4$, $p = 0.002$). Based on these findings, we can conclude that customers that request additional help have a significantly longer infection time.

7.3 REQUESTING A PAID TECHNICIAN

In addition, we analysed if the infection time and cleanup rates also significantly differ between the customers that requested the help from a technician and those who did not. It is found that the latter group has a shorter median infection time: 12 hours compared to almost 60 hours for the group that needed a technician. In addition, the cleanup rate after 1, 5 and 14 days are higher for the group that did not request additional help. To assess whether the differences between the groups are significant, a log-rank test is conducted. However, the log-rank test shows that there is no significant difference between the groups ($\chi^2 = 1.9$, $p = 0.167$). From this, we can conclude that customers that need a technician to solve the IoT malware problem do not have a significantly longer infection time.

Table 20 Summary statistics cleanup rate and infection time for understanding vs. misunderstanding notification

	# of customers	**% clean after 1 day**	**% clean after 5 days**	**% clean after 14 days**	**Median infection period**
(a) Running a virus scanner					
No virus scanner	97	61.8%	77.3%	88.7%	9 hours
Run virus scanner	50	33.0%	60.0%	90.0%	68.5 hours
(b) Requesting additional help					
No help requested	102	59.8%	81.4%	92.2%	9 hours
Help requested	45	33.3%	51.1%	82.2%	119 hours
(c) Requesting a paid technician					
No technician	123	56.1%	73.2%	90.2%	12 hours
Technician	24	29.1%	62.5%	83.3%	59.5 hours
(d) Identifying the IoT device					
Does not identify IoT device	53	50.9%	73.6%	88.7%	24 hours
Identifies IoT device	94	52.1%	70.2%	89.4%	23 hours
(e) Requesting to call with abuse team					
No request to call	121	53.7%	73.6%	90.1%	23 hours
Requests to call	26	43.3%	61.5%	84.6%	51.5 hours



(a) Running a virus scanner

(b) Requesting additional help

(c) Requesting a paid technician

(d) Identifying the IoT device

(e) Requesting to call with abuse team

Figure 16 Survival probabilities cleanup for understanding vs. misunderstanding notification

7.4 IDENTIFYING THE IOT DEVICE

We examined the correlation between identifying of the IoT device and the cleanup rate. Looking at Table 20, hardly any differences can be found between the customers that identified the IoT device and those who did not in terms of median infection time and cleanup rates. This is confirmed by means of the log-rank test (χ^2 = 0.4, p = 0.529). From this, we can conclude that whether or not the customer identifies the IoT device in their communication to KPN does not make a measurable difference in terms of the infection time.

7.5 REQUESTING TO CALL WITH ABUSE TEAM

The last topic related to misunderstanding that we investigated is the request to call with the abuse team. As expected, in Table 21 and Figure 17, we can see that the median infection time of customers that wanted to call with the abuse team is higher than for those who did not need this additional help: 51.5 and 23 hours respectively. Even though the cleanup rates of the customers that wanted to have the ability to call the abuse team are lower, we did not find a significant different between the groups (χ^2 = 2.2, p = 0.141). From this, we can conclude that customers that request to call with the abuse team in order to help them solve the IoT malware problem do not have a significantly longer infection time.

7.6 CONCLUSIONS

The defined goal of this chapter was to answer sub-question 8: what is the impact of the customer's understanding of the notification on IoT malware cleanup? To this end, we correlated the customer's apparent misunderstanding of the notification with the infection time and cleanup rate based on 5 topics that illustrate the customer's understanding. We found that customers that requested additional help to clean up the infection had a significantly longer infection period than those who did not request this help. An investigation of the other 4 topics that illustrate the customer's understanding showed that customers that (1) ran a virus scanner on their computer; (2) requested the help from a technician; (3) requested to call with the abuse team; or (4) did not identify the IoT device do not have a significantly longer infection time or lower cleanup rate than those who did not. However, the differences in infection time and cleanup rates between the groups that ran a virus scanner and those who did not are almost significant. Therefore, it should be noted that a larger study population might change this outcome. From these results, we can conclude that not all topics that illustrate the customer's understanding of the notification make measurable differences in terms of infection time and cleanup rates. Only the customers that requested additional help from KPN have a significantly longer infection period than the customers who did not.

8 CONCLUSIONS, DISCUSSION AND FUTURE WORK

The defined objective of this research is to make recommendations to a Dutch ISP on what notification mechanism to adopt by providing insight into which notification channel and content are the most effective in terms of both of IoT infections cleanup and the customer reactions. In this final chapter, an answer is provided to the corresponding main research question.

In Section 8.1, we start with the main conclusion of this research project. Thereafter, we provide the conclusions for each of the sub-research questions that have been formulated in Chapter 1. This is followed by a discussion of the results in Section 8.3. Lastly, in Section 8.4, we provide our recommendations for future work.

8.1 MAIN CONCLUSION

In this study we have searched for an answer to the question: What notification mechanism is the most effective in terms of both IoT malware cleanup and improving the reactions of customers? To this end an experiment has been conducted at KPN's abuse department to measure the difference in cleanup among IoT malware notifications sent via different channels and with different messages. To explore the reactions of the customers to the different notification mechanisms, telephone interviews have been conducted and the communication logs between KPN and the customers in the experiment have been analysed. We have compared the influence of the notification channel on cleanup and the reactions of customers by comparing customers that received: (1) email notifications; and (2) a combination of walled garden and email notifications. The different notification messages that have been compared in this study include: (1) the walled garden notification content that KPN's abuse department uses to notify its customers with an IoT malware infection; and (2) a newly composed more actionable walled garden notification message which clearly defines the steps that need to be taken while avoiding technical terms.

The results of the experiment have shown that quarantined customers have a shorter infection time than customers that were only emailed. In addition, a more actionable walled garden notification content also improves the IoT malware cleanup compared to a less actionable content of the notification. We found no measurable differences in terms of infection time and cleanup rate when comparing email notifications and walled garden notifications with the old content to the control group.

The analysis of the customer interviews and the communication logs between KPN and their customers have shown that quarantining improves the reaction time and reaction rate after an IoT malware notification significantly compared to email notifications. In addition, walled garden notifications have a higher probability of being read and more often encourage people to disconnect their device from the Internet. However, in some cases the quarantine event leads to complaints over the disruption. Regarding the notification content, we can conclude that the more actionable content of a walled garden notification does not make a difference in the reaction time and reaction rate compared to a less actionable content of the notification. Though, the newly composed notification content improves the understanding and trust compared to the old notification content.

An analysis of the correlation between variables related to the customer's understanding of the notification and the cleanup has shown that customers' apparent misunderstanding of the IoT malware notification does not always correlate with a longer infection period. Only the customers that requested additional help to clean up the IoT malware infection have a significantly longer infection period than the customers who did not request additional help.

From these results, we can conclude that a combination of a walled garden and email notification with an actionable content is the most effective in terms of IoT malware cleanup. Furthermore, the walled garden notification is most effective in getting customers to read and react to the IoT malware notification, yet it sometimes results in customers having a low satisfaction with the service they receive. The more actionable notification content results in better understanding and trust from the customer compared to a less actionable content of the notification.

8.2 ANSWERING SUB-QUESTIONS

In this section, a conclusion for each of the six research questions will be provided. Based on these conclusions, we derived our answer to the main research question of this study.

SQ 1. What are possible IoT malware notification channels for an ISP?

In Chapter 2, we described different abuse notification channels which could be used by ISPs to notify infected customers by conducting a literature review. In the literature, 6 notification channels were described: (1) email; (2) telephone call; (3) postal mail; (4) instant message; (5) SMS; and (6) walled garden notification. During an interview with abuse experts from KPN, we evaluated the feasibility of the different notification channels for KPN (see Section 3.5).

First, they evaluated the current methods used by KPN to notify. Currently, KPN uses combination of walled garden and email notifications. Abuse experts emphasized the strength of a walled garden notification, as the customer is notified and simultaneously, the communication between the bot and C&C server is blocked. In addition, there is a high likelihood that the customer sees the walled garden notification and acknowledges it in order to leave the quarantine environment. In contrast, for an email notification, it is uncertain whether a notification is received and read. Moreover, customers might not expect that KPN would notify them about an infection and consider the notifications as spam or phishing attempt. To this end, KPN's abuse experts indicated that the abuse department prefers the use of multiple channels, as this increases the credibility of the notifications.

Furthermore, they mentioned why other mechanisms are not in use. Major drawbacks of the postal mail notifications are the additional preparation time, delivery time, and additional cost to KPN. Meanwhile telephone calls can be time-sensitive, but it requires even higher cost to keep it running. Moreover, instant message notifications have never been considered by the abuse department, as they would have to invest in a new notification system.

SQ 2. Which factors influence the actionability of an IoT malware notification content?

To determine which factors influence the actionability of an IoT malware notification content, a literature review has been conducted (see Section 2.5) . In the past years, a range of best practices and guidelines for ISPs around the content of malware notifications has been published by leading industry associations (Livingood et al., 2012; Messaging Anti-Abuse Working Group, 2007; Online Trust Alliance, 2012). However, there is no prior research into the content of malware notifications to end users based on communication and persuasion theory. Therefore we investigated a closely related area of work: the design of security warnings. We combined the C-HIP model with PMT. Following the C-HIP model, we identified content characteristics that attract the attention of the receiver and improve the understanding. Moreover, factors were identified that drive the protection motivation. These factors were translated into guidelines that can be used to compose an actionable walled garden notification message (see Table 4). In short, the guidelines include: (1) clearly specify the underlying risk; (2) write the message for the least technical user, therefore avoid technical terms; (3) provide clear and easily recognisable action steps; (4) write the message in the primary language of the reader; and (5) include the ability to verify the authenticity of the notification.

SQ 3. What is KPN's current notification process for customers with infected IoT devices?

In an interview with an abuse expert from KPN, the notification procedure of KPN's abuse team was described (see Chapter 3). In subsequent expert interviews, the strengths and weaknesses of the current notification mechanisms were discussed.

The IoT malware notifications to consumer market customers are triggered by IoT malware infections in KPN's network reported by Shadowserver. IP addresses with a Mirai infection, which is the only IoT malware infection Shadowserver currently reports, are quarantined in a so-called walled garden and receive an email notification simultaneously. The walled garden allows access to a landing page (see Appendix C) and a set of white-listed websites, including cleanup tools, anti-virus solutions, Microsoft updates, webmail providers and online banking. Customers are given only two options to self-release from the quarantine environment within a period of 30 days. With the third quarantine action, intervention of the KPN's abuse team is required for assisted release. After a period of 30 consecutive days in quarantine, the walled garden automatically releases those quarantined customers who did not self-release or contact the abuse staff. When using the self-release option, customers have to submit a contact form through the walled garden landing page (Appendix F). In this form, customers can describe which measures they took, as well as additional comments they might have. This quarantine procedure is the same for all reported abuse events in KPN's consumer market network. However, the customer is shown a malware-specific notification message on the quarantine landings and in the email. In order to solve the problem, customers can contact the abuse team via email for additional information. Moreover, customers can contact the help desk via phone, chat or social media. The help desk employees advise the customer to send an email to the abuse team.

During the expert interviews, it is pointed out that the quarantine system has some weaknesses which make notifying customers with a Mirai infection more difficult. The walled garden contact form is the same for all abuse problems. Therefore, questions are asked to customers who are quarantined because of a Mirai infection which are unrelated to an IoT infection. Moreover, the abuse team employees stated that the Mirai notification message used by the team is too technical to comprehend.

SQ 4. How can the effectiveness of notification mechanisms on IoT malware cleanup be measured quantifiably?

Like in previous studies, the malware cleanup is evaluated quantifiably based on two metrics: (1) the median infection time; and (2) the cleanup rate. To determine the infection time, Mirai infections are tracked for 44 days after the initial notification by using the Shadowserver and IoTPOT abuse feeds. The cleanup rate is defined as the percentage of customers that cleaned up within 14 days after the initial notification.

SQ 5. What is the influence of the notification channel on IoT malware cleanup?

From November 6th, 2017 until March 1st, 2018, an experiment has been conducted at the abuse department to investigate the influence of different notification channels on IoT malware cleanup (SQ 5). Moreover we investigated how a more actionable notification message influences the IoT malware cleanup (SQ 6). For this purpose a new, more actionable notification message was composed based on guidelines from the literature (see section 4.5.4). The experiment, which is described in Chapter 5, had 4 different treatment groups: (1) email with new content; (2) walled garden with new content; (3) control group. Moreover, historical data from between June 20th, 2017 and October 11th, 2017 about customers that were quarantined by the abuse department for a Mirai infection content was collected. This constitutes the fourth treatment group: (4) walled garden with old content.

We investigated the influence of the combination of email and walled garden notifications (walled garden with new content) compared to email notifications (email with new content) in terms of the median infection time and the cleanup rates. The notification content of both treatment groups was the same. As a result of this comparison, we found a significant difference, at a 0.05 significance level, in terms of infection time between the walled garden and email notifications. The group that received the email notification cleaned up slower than the customers that were also placed in quarantine. Moreover, we evaluated the difference in terms of infection time between the control group and each of the treatment groups. Contrary to what was expected based on the literature, we only found a significant difference between the group that received a walled garden notification with the new content and the control group in terms of infection time. The infection time after an email notification or a walled garden notification with the old content was not significantly different from the infection time of a Mirai infection for which no notification was sent.

SQ 6. What is the influence of a more actionable walled garden notification content on IoT malware cleanup?

In this experiment, we also investigated how a more actionable walled garden notification message influences the IoT malware cleanup. For this purpose a new more actionable notification message was composed based on guidelines from the literature (see section 4.5.4). In the experiment from November 6th, 2017 until March 1st, 2018, a treatment group received a walled garden notification with this new content. Moreover, historical data from between June 20th, 2017 and October 11th, 2017 about customers that were quarantined for a Mirai infection with the old notification content was collected.

To measure the influence of a more actionable content, infection times of the customers that were placed in quarantine with these different messages were compared. As a result of this comparison, we found a significant difference, at a 0.05 significance level, in terms of infection time between the different walled garden notification contents. The group that received the more actionable walled garden notification cleaned up the fastest.

SQ 7. What are the reactions of the customers to different notification mechanisms for IoT abuse?

After the notification campaigns in the experiment, communication data from the customers in the study was collected and customer interviews were conducted to investigate the reactions of the customers to the different notification mechanisms for IoT abuse. We evaluated the customer reactions quantifiably based on two metrics: (1) the median reaction time; and (2) the reaction rate. The reaction time is defined as the time (in days) between the first notification and the first reaction from the customer via an email, the contact form, or contact with the help desk. In addition, by analysing the outcomes of the telephone interviews and the communication logs between KPN and customers, it is explored how the customers perceived the different notification mechanisms.

We found that walled garden notifications significantly improve the reaction rate compared to email notifications. All customers reacted to the walled garden notification, more than 80% of the customers even reacted within one day of sending the notification. In contrast, only 37% of customers that received an email notification reacted at all. We did not find that a more actionable walled garden notification message results in a shorter reaction time compared to a less actionable content of the walled garden notification.

In a qualitative analysis of the customers reactions, based on data collected during the customer interviews and communication logs between the customer and KPN's abuse team and help desk, we saw recurring themes that speak to the customers' experiences of the notification mechanisms: (1) misunderstanding the notification; (2) disconnecting the infected device; (3) distrusting the notification; (4) voicing complaints; and (5) acknowledgement of receiving the notification.

It is observed that some customers wanted to solve the problem, but were unable to understand the notification. We found that about one third of the customers that received a IoT malware notification did not understand that running a virus scanner on a desktop computer or laptop would not remove the malware infection on their IoT device(s). In addition, some of the customers did not realise which of their devices were connected to the Internet. A hand full of customers only spoke about devices such as PCs, tablets and mobile phones in all their messages to KPN. Some customers become familiar with the problem after additional examples from the abuse department about which of their devices might be infected. We analysed how many customers requested additional help in their messages to KPN and found that when sending a more actionable notification message the need for additional help was considerably reduced. About 42% of the customers that received a walled garden notification with the old content requested additional help, while this percentage was 10% for the customers that received a walled garden notification with the new notification content. Moreover, 3.3% of the customers that received an email notification with the new notification content requested additional help. We observed, by means of the interviews, that the customers with a low confidence in their ability to solve the issue often do not try to understand the notification content. These customers always request additional help from a paid technician.

Moreover, we noticed that customers who are placed in quarantine indicated 5 times more often having disconnected their IoT device from the Internet than customers that only received an email notification.

Additionally, it is found that a handful of customers that received the notification with the old content did not trust the credibility of the IoT malware notification from KPN, while none of the customers in the walled garden with new content treatment group distrusted the notification.

By means of the customer interviews, we observed that 50% of the customers that received the email notification remembered receiving the message, even though in all cases the message was sent to the right email address.

Furthermore, surprisingly, only about 10% of the customers in the experiment very much appreciated that their ISP has a abuse department that actively notifies its customers in case of abuse problems. On the other hand, about 20% of the customers, who were placed in quarantine, voiced complaints over the disruption and some even threatened to terminate their contract. We noticed that most of these dissatisfied customers were running small businesses with a consumer market Internet subscription.

SQ 8. What is the impact of the customer's understanding of the notification on IoT malware cleanup?

We investigate the impact of the customer's understanding on IoT malware cleanup (SQ 8) by measuring the differences in terms of infection time and cleanup rates among customers that do and do not understand the notification message. To this end, we used the results of the experiment and the outcomes of the qualitative analysis of the communication logs and customer interviews. We correlated the customer's apparent misunderstanding of the notification with the infection time and cleanup rate based on 5 topics that illustrate the customer's understanding: (1) running a virus scanner; (2) requesting additional help; (3) requesting additional help from a technician; (4) identifying the IoT device; and (5) requesting to call with the abuse team. We found that customers that requested additional help to clean up the infection had a significantly longer infection period than those who did not request this help. An investigation of the other 4 topics that illustrate the customer's understanding showed that customers that ran a virus scanner on their computer, requested the help from a technician, requested to call with the abuse team or did not identify the IoT device did not have a significantly longer infection time or lower cleanup rate than those who did not. From this, we can conclude that not all topics that illustrate the customer's understanding of the notification make measurable differences in terms of infection time and cleanup rates. Only the customers that requested additional help from KPN have a significantly longer infection period than the customers who did not.

8.3 DISCUSSION

8.3.1 IMPLICATIONS OF THE EMPIRICAL FINDINGS

Our findings demonstrate that walled garden notifications with a more actionable content significantly reduce the IoT malware infection times. Improving both the customers' understanding of the notification content and the customers' satisfaction with a the quarantine event remains a challenge. In the next sections, we will reflect upon our findings and create recommendations for the KPN's abuse team.

Notification channel

Our study found that the walled garden notification is a great tool in terms of making notification acknowledgement visible. KPN knows if a customer is notified, as the customer has to either submit a walled garden contact form for self-release or contact a KPN employee for assisted release. In contrast, when an email notification is send, reacting is optional. As a result of this, KPN does not know whether the customer has received the notification. Based on the customer interviews, we observed that 50% of the customers that received the email notifications, remembered receiving the notification. In all these cases, email notifications were sent to the actively used email addresses. On the other hand, email notifications can be removed by the spam filters and might be sent to unused email addresses. In such cases, the customers might not be informed at all. Therefore, walled garden notification can be preferable to avoid situations such as this one.

On the other hand, maintaining a walled garden system is a significant investment for an ISP. Furthermore, providing support to users in their attempts to clean up also imposes a significant cost. During this study, many quarantined customers contacted KPN for help. In addition to the costs, ISPs could decide not to invest in a walled garden system because it could influence the customer satisfaction and retention as illustrated by the customers that voiced complaints over the disruption and in some cases even threatened to terminate their contract.

In order to reduce the costs of the quarantine environment, KPN could decide to allow self-release more broadly. Based on 2 reasons of dissatisfaction shown by the customer reactions analysis, we identified possible improvements of KPN's walled garden notification procedure. Firstly, we observed that the dissatisfaction with the quarantine events was partly caused by quarantine events on Fridays. These customers needed abuse team employees for assisted release. They were, however, not available during the weekend. Therefore, these customers stayed in quarantine throughout the weekend. The KPN abuse team might consider always giving a self-release option, when quarantining customers on Fridays. Secondly, we noticed that most of the customers that were dissatisfied with the disruptive nature of the walled garden notification were running small businesses with a consumer market Internet subscription. KPN could prevent them from being affected in the future by providing an easy transition to a comparatively-priced business subscription, which would take them out of the consumer market and thus keep them away from the walled garden.

It would be interesting to investigate how other notification channels, which have not been analysed in this study, further improve the reachability of customers and IoT malware cleanup. For example, an SMS could be send together with the email or walled garden notification. Moreover, KPN could consider calling the customers.

Notification content

We observed that the walled garden notifications are only effective in terms of cleanup when they are used properly. The user should be provided with an actionable content, which includes clear action steps and avoids technical terms. Walled garden notifications that lack such actionable steps about how to clean up the infection and how to prevent future infections appear to have no distinguishable impact compared to not sending any notification at all. Customers need clear explanations about the problem and the solution in order to perform cleanup. As the analysis of communication logs and interviews showed, one third of the customers do not understand that running a virus scanner on a desktop computer or laptop would not help them remove the malware infection on their IoT device(s). Moreover, some of the customers do not immediately realise which of their devices are connected to the Internet or compromised after a Mirai notification. These customers need additional information in order to solve the problem. Our findings demonstrate that the need for additional help reduced when customers are provided with an actionable notification content.

As communicating the technical problem to consumer market customer remains a challenge, KPN's abuse team might consider using other means. They could, for example, create a small video in which the problem is described and the actions steps are shown.

Lastly, the analysis of the communication logs and interviews have showed that some customers did not trust the credibility of the email notifications. Therefore, we suggest to include the customer's name, customerID and KPN's logo in the notification content to improve credibility of the notifications.

8.3.2 RESEARCH QUALITY

In this research project we have strived for the highest possible internal validity. In this section we describe which choices have been made for this purpose. In addition, we describe the limitations relevant to the findings of our study and discuss the generalizability of our findings to other Dutch ISPs.

Internal validity

All consumer market customers that were known to be infected with IoT malware Mirai have been included in the experiment. No selection have been made. In addition, the treatment groups are equivalent, because the infected customers have been distributed over the treatment groups randomly. Furthermore, none of the customers in this research project knew that they were part of a study.

Limitations

There are, however, several limitations that could negatively influence the internal validity of this study. First of all, the influence of a more actionable walled garden notification content on IoT malware cleanup is determined by comparing the treatment groups with new and old content. The notifications are, however, sent during different periods and therefore the average infection duration in the treatment groups could be influenced by outside influences. The infection time could for example increase due to a holiday period, when customers might use their computers less, or because a C&C issues more commands to its bots in a certain period. Moreover, in periods when DDoS attacks that relied upon IoT devices are in the news, customers might be more aware of the urgency to take action. Furthermore customers might better understand the content of the notification messages. Secondly, the standard procedure of the KPN abuse team is followed. Besides the new more actionable notification message, the KPN Abuse is responsible for the subsequent communication with the customer. They determine the content of the messages to customers. Furthermore, the abuse team works from Monday to Friday. Therefore, customers are only notified during working days and no emails from customers are answered outside working hours. Customer could contact the KPN help desk. The quality of the advice of the help desk could, however influence the duration of the infection. When a help desk employee does not immediately recognize that the customer is dealing with an abuse problem, this might extend the infection duration. Thirdly, only customers are notified when they appear with a Mirai infection in the Shadowserver feeds. If the feed is incomplete, when Shadowserver could not identify the infection type as Mirai or when KPN does not receive any reports from Shadowserver, infected customers might not receive a notification. This absence may cause us to overestimate the cleanup rate. Moreover, the Shadowserver reports only include the first event for each IP address on a day, even though there could be multiple events. Therefore, we could underestimate the infection time. However, it affects the infection time of the customers in the different treatment groups in the same way. Fourth, customers who have terminated their contract have not been called for a telephone interview, while these customers might have a strong opinion regarding the notification mechanism. Fifth, a customer in the email treatment group who was still infected after 14 days, is placed in quarantine. In the interview, questions are asked about the email notification specifically. Nevertheless, the customer might confuse the different notification mechanisms while answering the questions. Sixth, the period between the notification and the customers interview differed for the different customers, as the end of the experiment almost coincides with the end of the research project. This might have influenced how well the customer remembers receiving and reading the notification message. Lastly, this study has a relatively small study population. While the sample size was sufficient to reach statistically significant results, it would be nice to carry out an experiment on larger samples.

Generalisability

The external validity (generalisability) of this research project is debatable. On the one hand, the study is conducted in a real-world setting. Customers did not know that they were part of a research project. Therefore this knowledge could not have influenced their behaviour. In addition, KPN's consumer market customer represent a wide variety of people in terms of demographics. Therefore, one could argue that the findings can also be used by other Dutch ISPs that notify their retail customers which are known to be infected with IoT malware. On the other hand, the study is based on a single ISP. KPN is a relatively expensive ISP in the Dutch market and therefore KPN's customers might not be representative for all Dutch Internet users well. Moreover, the study is based on a single type of IoT malware and a specific implementation of a walled garden system to notify and quarantine end users. Therefore, the generalizability of our results is a matter for further studies.

8.3.3 COMPARISON TO RELATED WORK

In this study we found high cleanup rates for quarantined users, 89% and 97% for the old and new content respectively. Of those who received the email notification with the new content, 77% cleaned up. Surprisingly, 79% of the customers who were not notified also cleaned up within 14 days. There is no clear point of reference for this success rate. Prior work in the field of abuse notifications found lower cleanup rates, mostly around 50% (Çetin et al., 2016; Li, Ho, et al., 2016; Vasek & Moore, 2012). However, comparison is difficult as the recipient of the notifications in these studies is a server admin or webmaster, not a home user.

Vasek and Moore (2012) found that only abuse reports with detailed information result in higher cleanup rates of compromised websites compared to those not receiving a notice. They found no difference in cleanup rates for websites receiving a minimal notice and those not receiving any notice at all. In this study, we also observed no significant difference in cleanup rates between sending a walled garden notification that lacks actionable steps and not sending any notification at all. Thus, we corroborate the finding of Vasek and Moore (2012) for a different type of malware and recipient.

Çetin et al. (2017) suggested to move away from email as the main notification channel due to the high bounce rate. Moreover, Livingood et al. (2012) indicated that it is not assured whether the email notification is read in a timely manner, because of the recipient's distrust or an incorrect email address. In this study, we found that 50% of the customers that received an email-only notification did not remember receiving the notification even though the email was in all cases sent to the right email address. This finding illustrates that email bounces are not a problem for ISPs. The problem, however, is that customers do not pay attention to an email-only notifications. As stated during the customer interviews, customers only watch emails from KPN regarding invoices.

8.3.4 CONTRIBUTIONS

This research project has made several contributions, of both scientific and practical value. Previous studies have investigated which factors influence the effectiveness of abuse and vulnerability notifications to affected parties. But this project is the first that investigated the factors influencing the effectiveness of IoT malware notifications. Moreover, this is the first study that provides insight into the context of an ISP's abuse department which sends abuse notifications to home users. The collaboration with KPN made it possible study the effectiveness of a walled garden system to notify and quarantine end users with malware infected machines. Moreover, we provided insight into the experiences of users by customer interviews and analysing their communications with KPN employees. In addition, the collaboration with KPN made it possible to correlate the understanding of the customer with the infection time. Furthermore, this thesis makes contributions to KPN and society. For KPN, it presents guidelines for writing an actionable notification content for IoT malware infected customers. Furthermore, this thesis provides recommendations for improving IoT malware cleanup and suggestions for improving customer satisfaction.

8.4 RECOMMENDATIONS FOR FUTURE WORK

This research can be further extended in several ways to answer different research questions. Following topics can be explored to further investigate the topic.

8.4.1 GENERALISING THE RESULTS OF THE EXPERIMENT

As mentioned in the Section 8.3.2, the results of the experiment are tied to the data of a single ISP in the Netherlands. To assess the generalisability of our results, follow up studies can be conducted with ISPs in the Netherlands and other countries. These studies can use similar metrics to compare the reactions of the customers and cleanup rate of the IoT malware infection. This way we can understand how reproducible these results are in different networks.

Moreover, results are based on one type of IoT malware. This is mainly because of the feeds that we used in the experiment. KPN's abuse desk does not have any other reliable IoT malware feed that could be used in the study. In the future, different IoT malware feeds could be purchased or requested to conduct similar experiments to assess the generalisability of the results.

Lastly, to evaluate this experiment, we have used IoTPOT and Shadowserver feeds. These were the only available data sources for us during the experiment. Follow up studies can leverage additional data sources to improve the tracking of IoT malware infections. This way more accurate results can be obtained to measure the influence of the experimental variables.

8.4.2 INCLUDING DEVICE INFORMATION IN THE NOTIFICATIONS

During the customer interviews several customers stated that they would have liked to know which of their many devices was causing the problem. It would be interesting to test empirically whether providing the name of the compromised IoT device for a proportion of the reported infections, affects the infection time. Follow up studies can focus on the identification of the infected device type to include in the experiment. To what extent does providing device information influence IoT infection rate and speed?

8.4.3 INFLUENCE OF DEVICE TYPE ON THE CLEANUP

During the experiment period of 14 days, most of the experimental groups achieved high cleanup rates. We would like to know how the device type influences the cleanup rates and speed. This would help us to understand why the control group also has a relatively high cleanup rate while their owners did not receive any notifications.

8.4.4 EFFECTIVENESS OF OTHER COMMUNICATION CHANNELS

In this work, we measured the effectiveness of two commonly used notification channels in terms of IoT infection cleanup and customer reactions. In future work, the influence of other communication channels on IoT malware cleanup and customer reactions can be measured to find out the most effective communication channel at getting customers to act upon the IoT infection. For example, SMS messages or telephone calls can be used as a treatment, compared to walled garden notifications to see which one of these improve the customer satisfaction and IoT malware cleanup. How do other communication channels influence customer satisfaction and IoT malware cleanup rate and speed?

8.4.5 REINFECTION

In this study, we did not measure the impact of the reinfection rate due to urgent need for remediation for the email treatment group. In future work, it would be interesting to see which percentage of the customers did not address the underlying problem and became victim of IoT abuse again.

9 REFERENCES

Almuhimedi, H., Felt, A. P., Reeder, R. W., & Consolvo, S. (2014). Your Reputation Precedes You: History, Reputation, and the Chrome Malware Warning. In *Proceedings of the Tenth Symposium On Usable Privacy and Security (SOUPS)* (pp. 113–128). Menlo Park, CA: USENIX. Retrieved from https://www.usenix.org/system/files/conference/soups2014/soups14-paper-almuhimedi.pdf

Barcena, M. B., & Wueest, C. (2015). Insecurity in the Internet of Things.

Bauer, L., Bravo-Lillo, C., Cranor, L., & Fragkaki, E. (2013). Warning Design Guidelines. *CyLab*. Retrieved from http://repository.cmu.edu/cylab/113

Bertino, E. (2016). Data Security and Privacy in the IoT. In *EDBT* (pp. 1–3). https://doi.org/10.5441/002/edbt.2016.02

Bertino, E., & Islam, N. (2017). Botnets and Internet of Things Security. *Computer*, *50*(2), 76–79. https://doi.org/10.1109/MC.2017.62

Bland, J. M., & Altman, D. G. (2004). The logrank test. *BMJ (Clinical Research Ed.)*, *328*(7447), 1073. https://doi.org/10.1136/bmj.328.7447.1073

Bravo-Lillo, C., Cranor, L., Komanduri, S., Schechter, S., & Sleeper, M. (2014). Harder to Ignore? In *Proceedings of the Tenth Symposium On Usable Privacy and Security (SOUPS)* (pp. 105–111). Menlo Park, CA: USENIX. Retrieved from https://www.usenix.org/system/files/soups14-paper-bravo-lillo.pdf

Bravo-Lillo, C., Komanduri, S., Cranor, L. F., Reeder, R. W., Sleeper, M., Downs, J., & Schechter, S. (2013). Your attention please: Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. In *Proceedings of the Ninth Symposium on Usable Privacy and Security (SOUPS)*. Newcastle, UK: ACM Press. https://doi.org/10.1145/2501604.2501610

Canali, D., Balzarotti, D., & Francillon, A. (2013). The role of web hosting providers in detecting compromised websites. In *Proceedings of the 22nd international conference on World Wide Web - WWW '13* (pp. 177–188). New York, New York, USA: ACM Press. https://doi.org/10.1145/2488388.2488405

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

CDC. (2010). Simply put; a guide for creating easy-to-understand materials. Retrieved from https://stacks.cdc.gov/view/cdc/11938

Çetin, O., Gañán, C., Korczyski, M., & Van Eeten, M. (2017). Make Notifications Great Again: Learning How to Notify in the Age of Large-Scale Vulnerability Scanning. In *WEIS 2017*. Retrieved from http://weis2017.econinfosec.org/wp-content/uploads/sites/3/2017/05/WEIS_2017_paper_17.pdf

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

Desaulniers, D. R. (1987). Layout, Organization, and the Effectiveness of Consumer Product Warnings. *Proceedings of the Human Factors Society Annual Meeting*, *31*(1), 56–60. https://doi.org/10.1177/154193128703100112

Durumeric, Z., Payer, M., Paxson, V., Kasten, J., Adrian, D., Halderman, J. A., … Beekman, J. (2014). The Matter of Heartbleed. In *Proceedings of the 2014 Conference on Internet Measurement Conference - IMC '14* (pp. 475–488). New York, New York, USA: ACM Press. https://doi.org/10.1145/2663716.2663755

Eeten, M. J. G. van, & Bauer, J. M. (2008). *Economics of Malware*. OECD Publishing. https://doi.org/10.1787/241440230621

Egelman, S., Cranor, L. F., & Hong, J. (2008). You've been warned. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08* (p. 1065). New York, New York, USA: ACM Press. https://doi.org/10.1145/1357054.1357219

Egelman, S., & Schechter, S. (2013). The Importance of Being Earnest [In Security Warnings]. In *Proceedings of Financial Cryptography and Data Security (FC)* (pp. 52–59). Okinawa, Japan: Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39884-1_5

European Commission. (2012). *Special Eurobarometer 390 Cyber security*. Retrieved from http://ec.europa.eu/public_opinion/index_en.htm

Evans, D. (2011). *The Internet of Things How the Next Evolution of the Internet Is Changing Everything*. Retrieved from http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

Fagan, M., Maifi, M., & Khan, H. (2016). Why Do They Do What They Do?: A Study of What Motivates Users to (Not) Follow Computer Security Advice Why Do They Do What They Do? In *Proceedings of the Twelfth Symposium On Usable Privacy and Security (SOUPS)* (pp. 59–75). Denver, CO, USA: USENIX. Retrieved from https://www.usenix.org/conference/soups2016/technical-sessions/presentation/fagan

Federal Trade Commission. (2015). *Internet of Things - Privacy & Security in a Connected World*. *FTC Staff Report*. Retrieved from https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf

Forget, A., Pearman, S., Thomas, J., Acquisti, A., Christin, N., Cranor, L. F., … Telang, R. (2016). Do or Do Not, There Is No Try: User Engagement May Not Improve Security Outcomes. In *Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS)* (pp. 97–111). Denver, CO, USA: USENIX. Retrieved from https://www.usenix.org/conference/soups2016/technical-sessions/presentation/forget

Gallagher, S. (2016). How one rent-a-botnet army of cameras, DVRs caused Internet chaos | Ars Technica. Retrieved March 8, 2017, from https://arstechnica.com/information-technology/2016/10/inside-the-machine-uprising-how-cameras-dvrs-took-down-parts-of-the-internet/

Gañán, C., Çetin, O., & van Eeten, M. (2015). An Empirical Analysis of ZeuS C&C Lifetime. In *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security - ASIA CCS '15* (pp. 97–108). New York, New York, USA: ACM Press. https://doi.org/10.1145/2714576.2714579

Gartner. (n.d.). IT Glossary- Internet of Things. Retrieved January 6, 2017, from http://www.gartner.com/it-glossary/internet-of-things/

Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research*, *1*(4), 274–8. https://doi.org/10.4103/0974-7788.76794

Hewlett-Packard Development Company. (2014). *Internet of Things Research Study*. Retrieved from http://go.saas.hpe.com/fod/internet-of-things

Hofmeyr, S., Moore, T., Forrest, S., Edwards, B., & Stelle, G. (2013). Modeling Internet-Scale Policies for Cleaning up Malware. In *Economics of Information Security and Privacy III* (pp. 149–170). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-1981-5_7

Jacobs, B. (2016). Aftercare for the Internet of Things. *CSR Magazine*, *2*(2). Retrieved from https://www.cybersecurityraad.nl/binaries/csr_magazine_2_2016_web_tcm56-28305.pdf

Jhaveri, M. H., Cetin, O., Gañán, C., Moore, T., & Eeten, M. Van. (2017). Abuse Reporting and the Fight Against Cybercrime. *ACM Computing Surveys*, *49*(4), 1–27. https://doi.org/10.1145/3003147

Kaspersky Lab. (2017). Skilled Attacker Develops Advanced Windows Botnet to Spread Infamous Mirai Malware [Press release]. Retrieved from www.kaspersky.com/about/press-releases/2017_skilled-attacker-develops-advanced-windows-botnet-to-spread-infamous-mirai-malware

Kim, S., & Wogalter, M. S. (2009). Habituation, Dishabituation, and Recovery Effects in Visual Warnings. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*(20), 1612–1616. https://doi.org/10.1177/154193120905302015

Krol, K., Moroz, M., & Sasse, M. A. (2012). Don’t work. Can’t work? Why it’s time to rethink security warnings. In *2012 7th International Conference on Risks and Security of Internet and Systems (CRiSIS)* (pp. 1–8). IEEE. https://doi.org/10.1109/CRISIS.2012.6378951

Kührer, M., Hupperich, T., Rossow, C., & Holz, T. (2014). Exit from Hell? Reducing the Impact of Amplification DDoS Attacks. In *Proceedings of the 23rd USENIX Security Symposium (USENIX Security 14)* (pp. 111–125). San Diego. Retrieved from https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-kuhrer.pdf

Level 3 communications. (n.d.). Attack of Things! Retrieved March 28, 2017, from http://netformation.com/level-3-pov/attack-of-things-2

Li, F., Durumeric, Z., Czyz, J., Karami, Mohammad Bailey, M., McCoy, D., Savage, S., & Paxson, V. (2016). You’ve Got Vulnerability: Exploring Effective Vulnerability Notifications. In *Proceedings of 25th USENIX Security Symposium (USENIX Security 16)* (pp. 1033–1050). Austin, TX. Retrieved from https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_li.pdf

Li, F., Ho, G., Kuan, E., Niu, Y., Ballard, L., Thomas, K., … Paxson, V. (2016). Remedying Web Hijacking. In *Proceedings of the 25th International Conference on World Wide Web - WWW ’16* (pp. 1009–1019). New York, New York, USA: ACM Press. https://doi.org/10.1145/2872427.2883039

Livingood, J., Mody, N., & O’Reirdan, M. (2012). Recommendations for the Remediation of Bots in ISP Networks. RFC 6561 (Informational). Retrieved from http://www.ietf.org/rfc/rfc6561.txt

Markoff, J. (2016, November 3). Why Light Bulbs May Be the Next Hacker Target. *The New York Times*. Retrieved from http://www.nytimes.com/2016/11/03/technology/why-light-bulbs-may-be-the-next-hacker-target.html

Mathur, A., Engel, J., Sobti, S., Chang, V., & Chetty, M. (2016). &quot;They Keep Coming Back Like Zombies&quot;: Improving Software Updating Interfaces. In *Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS)* (pp. 43–58). Denver, CO, USA: USENIX. Retrieved from https://www.usenix.org/conference/soups2016/technical-sessions/presentation/mathur

Messaging Anti-Abuse Working Group. (2007). Best Common Practices for the Use of a Walled Garden. Retrieved from https://www.m3aawg.org/sites/default/files/document/M3AAWG_Walled_Garden_BCP_Ver2_2015-03_0.pdf

Modic, D., & Anderson, R. (2014). Reading this may harm your computer: The psychology of malware warnings. *Computers in Human Behavior*, *41*, 71–79. https://doi.org/10.1016/j.chb.2014.09.014

Nappa, A., Zubair Rafique, M., & Caballero, J. (2013). Driving in the Cloud: An Analysis of Drive-by Download Operations and Abuse Reporting. In *Proceedings of the 10th Conference on Detection of Intrusions and Malware & Vulnerability Assessment* (pp. 1–20). Berlin: Springer. Retrieved from https://link.springer.com/content/pdf/10.1007%252F978-3-642-39235-1_1.pdf

Neupane, A., Saxena, N., Kuruvilla, K., Georgescu, M., & Kana, R. (2014). Neural Signatures of User-Centered Security: An fMRI Study of Phishing, and Malware Warnings. In *Proceedings 2014 Network and Distributed System Security Symposium*. Reston, VA: Internet Society. https://doi.org/10.14722/ndss.2014.23056

Online Trust Alliance. (2012). *Combatting Botnets Through User Notification Across the Ecosystem: A View of Emerging Practices*. Retrieved from https://otalliance.org/system/files/files/best-practices/documents/ota_botnet_notification_whitepaper12-7.pdf

Pa, Y. M. P., Suzuki, S., Yoshioka, K., Matsumoto, T., Kasama, T., & Rossow, C. (2015). IoTPOT: Analysing the Rise of IoT Compromises. In *Proceedings of 9th USENIX Workshop on Offensive Technologies (WOOT 15)*. Washington, D.C. Retrieved from https://www.usenix.org/system/files/conference/woot15/woot15-paper-pa.pdf

Pa, Y. M. P., Suzuki, S., Yoshioka, K., Matsumoto, T., Kasama, T., & Rossow, C. (2016). IoTPOT: A Novel Honeypot for Revealing Current IoT Threats. *Journal of Information Processing*, *24*(3), 522–533. https://doi.org/10.2197/ipsjjip.24.522

Pijpker, J., & Vranken, H. (2016). The Role of Internet Service Providers in Botnet Mitigation. In *2016 European Intelligence and Security Informatics Conference (EISIC)* (pp. 24–31). IEEE. https://doi.org/10.1109/EISIC.2016.013

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: a revised theory of protection motivation. In *Social Psychophysiology: A Sourcebook* (pp. 153–176). New York: Guilford Press. Retrieved from https://www.scienceopen.com/document?vid=a182a645-fc12-4d21-9e22-15c96a792275

Schneier, B. (2014). The Internet of Things Is Wildly Insecure — And Often Unpatchable. *WIRED*. Retrieved from https://www.wired.com/2014/01/theres-no-good-way-to-patch-the-internet-of-things-and-thats-a-huge-problem/

Schneier, B. (2016). Your WiFi-connected thermostat can take down the whole Internet. We need new regulations. *The Washington Post*. Retrieved from https://www.washingtonpost.com/posteverything/wp/2016/11/03/your-wifi-connected-thermostat-can-take-down-the-whole-internet-we-need-new-regulations/?utm_term=.b680990356b5

Scott, J., & Spaniel, D. (2016). *Rise of the Machines: The Dyn Attack Was Just a Practice Run*. Retrieved from http://icitech.org/wp-content/uploads/2016/12/ICIT-Brief-Rise-of-the-Machines.pdf

Seiders, K., Flynn, A. G., Berry, L. L., & Haws, K. L. (2015). Motivating Customers to Adhere to Expert Advice in Professional Services. *Journal of Service Research*, *18*(1), 39–58. https://doi.org/10.1177/1094670514539567

Shadowserver. (2018). Home page. Retrieved February 27, 2018, from https://www.shadowserver.org/wiki/

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

Silic, M., Barlow, J., & Ormond, D. (2015). Warning! A Comprehensive Model of the Effects of Digital Information Security Warning Messages. Retrieved from https://www.alexandria.unisg.ch/244531/

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

Stock, B., Pellegrino, G., Rossow, C., Johns, M., & Backes, M. (2016). Hey, You Have a Problem: On the Feasibility of Large-Scale Web Vulnerability Notification Hey, You Have a Problem: On the Feasibility of Large-Scale Web Vulnerability Notification. In *Proceedings of the 25th USENIX Security Symposium* (pp. 1015–1032). Austin, TX. Retrieved from https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/stock

Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., & Cranor, L. F. (2009). Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *USENIX security symposium* (pp. 399–416). Retrieved from https://www.usenix.org/legacy/event/sec09/tech/full_papers/sunshine.pdf

Symantec. (2016). IoT devices being increasingly used for DDoS attacks. Retrieved March 23, 2017, from https://www.symantec.com/connect/blogs/iot-devices-being-increasingly-used-ddos-attacks

Telecommunicatiewet. (1998). Retrieved May 10, 2017, from http://wetten.overheid.nl/BWBR0009950/2017-03-10

Vasek, M., & Moore, T. (2012). Do malware reports expedite cleanup? An experimental study. *CSET*. Retrieved from https://www.usenix.org/system/files/conference/cset12/cset12-final20.pdf

Virgin Media. (2017). Mirai malware alert. Retrieved September 27, 2017, from https://help.virginmedia.com/system/templates/selfservice/vm/help/customer/locale/en-GB/portal/200300000001000/article/HELP-2578/Mirai-malware-alert

Wash, R., Rader, E., Vaniea, K., & Rizor, M. (2014). Out of the Loop: How Automated Software Updates Cause Unintended Security Consequences. In *Proceedings of the Tenth Symposium On Usable Privacy and Security (SOUPS)* (pp. 89–104). Menlo Park, CA: USENIX. Retrieved from https://www.usenix.org/system/files/soups14-paper-wash.pdf

Wogalter, M. S. (2006). Communication-Human Information Processing (C-HIP) Model. In *Handbook of Warnings* (pp. 51–61). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from http://www.safetyhumanfactors.org/wp-content/uploads/2011/12/271Wogalter2006Chap5.pdf

Wogalter, M. S., & Laughery, K. R. (1996). WARNING! Sign and Label Effectiveness. *Current Directions in Psychological Science*. Sage Publications, Inc.Association for Psychological Science. https://doi.org/10.2307/20182386

Zhang, J., Duan, H., Liu, W., & Yao, X. (2017). How to Notify a Vulnerability to the Right Person? Case Study: In an ISP Scope. In *Proceedings of IEEE GLOBECOM 2017* (pp. 1–7). IEEE. https://doi.org/10.1109/GLOCOM.2017.8253993

APPENDIX A

SYSTEMS OF KPN'S ABUSE TEAM

CONFIDENTIAL

APPENDIX B

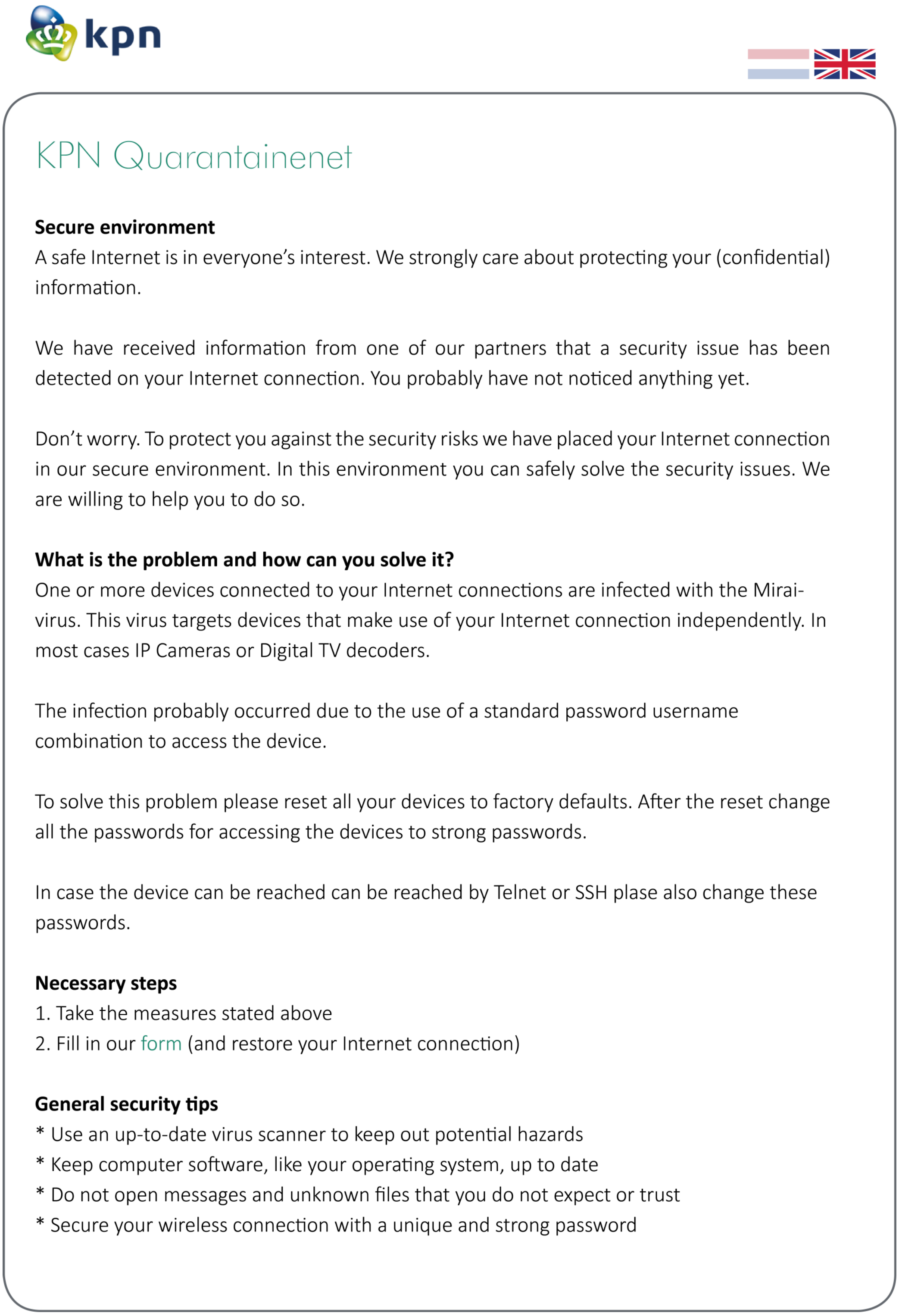
ABUSE AND VULNERABILITY TYPES KPN NOTIFIES

CONFIDENTIAL

APPENDIX C

LANDING PAGE QUARANTINE AREA FOR MIRAI INFECTION

C.1 ENGLISH VERSION



KPN Quarantainenet

Secure environment

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

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

Don't worry. To protect you against the security risks we have placed your Internet connection in our secure environment. In this environment you can safely solve the security issues. We are willing to help you to do so.

What is the problem and how can you solve it?

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

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

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

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

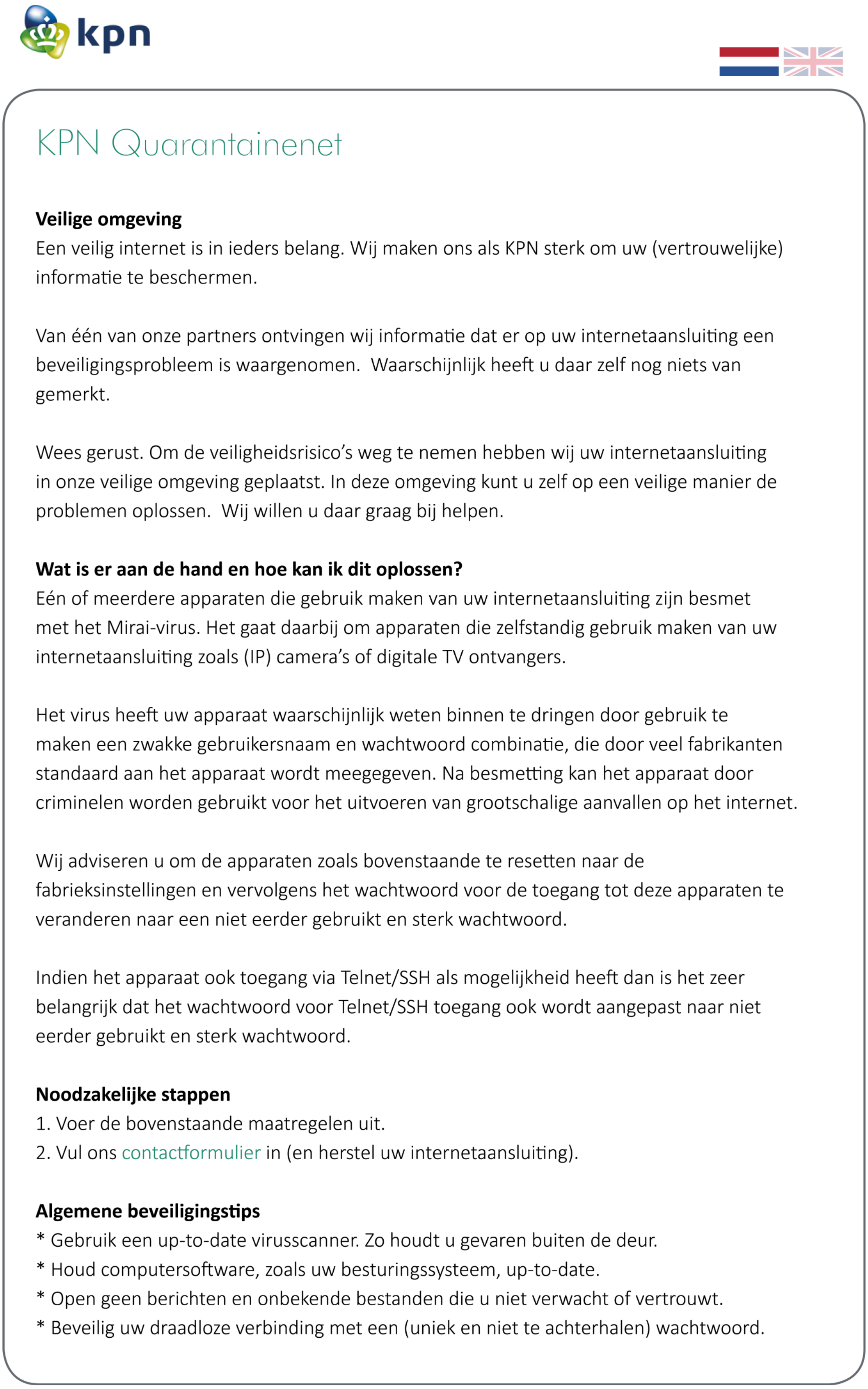
Necessary steps

1. Take the measures stated above
2. Fill in our form (and restore your Internet connection)

General security tips

* Use an up-to-date virus scanner to keep out potential hazards
* Keep computer software, like your operating system, up to date
* Do not open messages and unknown files that you do not expect or trust
* Secure your wireless connection with a unique and strong password

C.2 DUTCH VERSION



KPN Quarantainenet

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

Van één van onze partners ontvingen wij informatie dat er op uw internetaansluiting een beveiligingsprobleem is waargenomen. Waarschijnlijk heeft u daar zelf nog niets van gemerkt.

Wees gerust. Om de veiligheidsrisico's weg te nemen hebben wij uw internetaansluiting in onze veilige omgeving geplaatst. In deze omgeving kunt u zelf op een veilige manier de problemen oplossen. Wij willen u daar graag bij helpen.

Wat is er aan de hand en hoe kan ik dit oplossen?
Eén of meerdere apparaten die gebruik maken van uw internetaansluiting zijn besmet met het Mirai-virus. Het gaat daarbij om apparaten die zelfstandig gebruik maken van uw internetaansluiting zoals (IP) camera's of digitale TV ontvangers.

Het virus heeft uw apparaat waarschijnlijk weten binnen te dringen door gebruik te maken een zwakke gebruikersnaam en wachtwoord combinatie, die door veel fabrikanten standaard aan het apparaat wordt meegegeven. Na besmetting kan het apparaat door criminelen worden gebruikt voor het uitvoeren van grootschalige aanvallen op het internet.

Wij adviseren u om de apparaten zoals bovenstaande te resetten naar de fabrieksinstellingen en vervolgens het wachtwoord voor de toegang tot deze apparaten te veranderen naar een niet eerder gebruikt en sterk wachtwoord.

Indien het apparaat ook toegang via Telnet/SSH als mogelijkheid heeft dan is het zeer belangrijk dat het wachtwoord voor Telnet/SSH toegang ook wordt aangepast naar niet eerder gebruikt en sterk wachtwoord.

Noodzakelijke stappen
1. Voer de bovenstaande maatregelen uit.
2. Vul ons contactformulier in (en herstel uw internetaansluiting).

Algemene beveiligingstips
* Gebruik een up-to-date virusscanner. Zo houdt u gevaren buiten de deur.
* Houd computersoftware, zoals uw besturingssysteem, up-to-date.
* Open geen berichten en onbekende bestanden die u niet verwacht of vertrouwt.
* Beveilig uw draadloze verbinding met een (uniek en niet te achterhalen) wachtwoord.

APPENDIX D

EMAIL ALONG WITH QUARANTINE ACTION FOR MIRAI INFECTION

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?
Eén of meerdere apparaten die gebruik maken van uw internetaansluiting zijn besmet met het Mirai-virus. Het gaat daarbij om apparaten die zelfstandig gebruik maken van uw internetaansluiting zoals (IP) camera's of digitale TV ontvangers.

Het virus heeft uw apparaat waarschijnlijk weten binnen te dringen door gebruik te maken een zwakke gebruikersnaam en wachtwoord combinatie, die door veel fabrikanten standaard aan het apparaat wordt meegegeven. Na besmetting kan het apparaat door criminelen worden gebruikt voor het uitvoeren van grootschalige aanvallen op het internet.

Wij adviseren u om de apparaten zoals bovenstaande te resetten naar de fabrieksinstellingen en vervolgens het wachtwoord voor de toegang tot deze apparaten te veranderen naar een niet eerder gebruikt en sterk wachtwoord.

Indien het apparaat ook toegang via Telnet/SSH als mogelijkheid heeft dan is het zeer belangrijk dat het wachtwoord voor Telnet/SSH toegang ook wordt aangepast naar niet eerder gebruikt en sterk wachtwoord.

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.

Veilige omgeving
Aangezien het beveiligingsprobleem een groot gevaar vormt voor de veiligheid op internet hebben wij uw internetaansluiting in onze veilige omgeving (quarantaine) geplaatst. U kunt tijdelijk beperkt gebruik maken van uw internetaansluiting. Een dergelijke maatregel nemen wij ook om uw vertrouwelijke gegevens en bestanden te beschermen.

Met vriendelijke groet,

KPN Abuse Team
abuse@kpn.com

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

APPENDIX E

EMAIL NOTIFICATION FOR MIRAI INFECTION

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?
Eén of meerdere apparaten die gebruik maken van uw internetaansluiting zijn besmet met het Mirai-virus. Het gaat daarbij om apparaten die zelfstandig gebruik maken van uw internetaansluiting zoals (IP) camera's of digitale TV ontvangers.

Het virus heeft uw apparaat waarschijnlijk weten binnen te dringen door gebruik te maken een zwakke gebruikersnaam en wachtwoord combinatie, die door veel fabrikanten standaard aan het apparaat wordt meegegeven. Na besmetting kan het apparaat door criminelen worden gebruikt voor het uitvoeren van grootschalige aanvallen op het internet.

Wij adviseren u om de apparaten zoals bovenstaande te resetten naar de fabrieksinstellingen en vervolgens het wachtwoord voor de toegang tot deze apparaten te veranderen naar een niet eerder gebruikt en sterk wachtwoord.

Indien het apparaat ook toegang via Telnet/SSH als mogelijkheid heeft dan is het zeer belangrijk dat het wachtwoord voor Telnet/SSH toegang ook wordt aangepast naar niet eerder gebruikt en sterk wachtwoord.

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.

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.

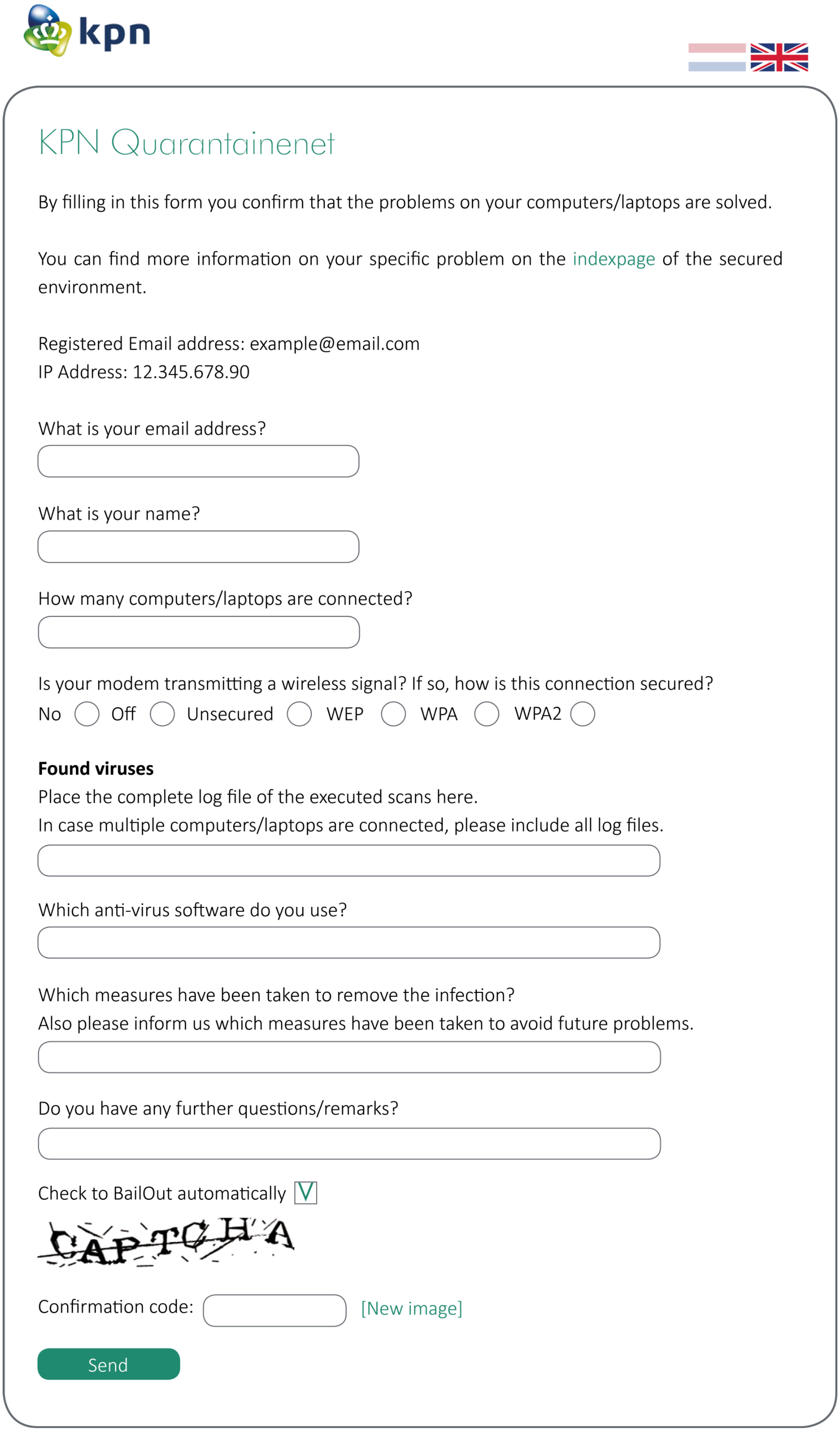
Met vriendelijke groet,

KPN Abuse Team
abuse@kpn.com

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

APPENDIX F

CONTACT FORM QUARANTINE AREA IN ENGLISH



KPN Quarantainenet

By filling in this form you confirm that the problems on your computers/laptops are solved.

You can find more information on your specific problem on the indexpage of the secured environment.

Registered Email address: example@email.com
IP Address: 12.345.678.90

What is your email address?

What is your name?

How many computers/laptops are connected?

Is your modem transmitting a wireless signal? If so, how is this connection secured?

No ○ Off ○ Unsecured ○ WEP ○ WPA ○ WPA2 ○

Found viruses

Place the complete log file of the executed scans here.
In case multiple computers/laptops are connected, please include all log files.

Which anti-virus software do you use?

Which measures have been taken to remove the infection?
Also please inform us which measures have been taken to avoid future problems.

Do you have any further questions/remarks?

Check to BailOut automatically ☑

Confirmation code: [New image]

APPENDIX G

FORMAT SHADOWSERVER BOTNET-DRONE REPORT

Table 21 Format Shadowserver Botnet-Drone report

Field	Description
Timestamp	Timestamp in UTC+0 the IP was seen/accessed the sinkhole system
ip	The IP of the device in question
port	Source port of the IP connection
asn	ASN of the IP
geo	Country location of the IP
region	State or province from the Geo
city	City from the Geo
hostname	Reverse DNS of the IP
type	Packet type of the connection traffic (udp/tcp)
	Drone type (if known)
infection	Infection name if known
url	Connection URL if applicable
agent	HTTP connection agent if applicable
cc	The Command and Control that is managing this IP / destination IP that the device in question is observed connecting to
cc_port	Server side port that the IP connected to
cc_asn	ASN of the C&C
cc_geo	Country of the C&C
cc_dns	For HTTP traffic, the content of the HTTP Host: header. Normally the fully qualified domain name of the C&C
count	Number of connections from this drone IP
proxy	If the connection went through a known proxy system
application	Application name / Layer 7 protocol
p0f_genre	Operating System family
p0f_detail	Operating System version
machine_name	Name of the compromised machine
id	Bot ID

APPENDIX H

OVERVIEW OF IOTPOT

Figure 17 shows an overview of the IoTPOT. The central part of the IoTPOT is the Front end Responder, which behaves as different types of IoT devices by handling incoming TCP connection requests, banner interactions, authentication, and command interactions with a set of device profiles. A device profile contains a banner profile, an authentication profile, and a command interaction profile. The banner profile determines how the honeypot responds in banner interactions, such as Telnet options, welcome message, and login prompt. The authentication profile determines how to respond to incoming authentication challenges. Command interaction profiles determine the responses to incoming commands. It includes a set of commands and their matching responses.

In the case of unknown commands, the Front end Responder establishes a Telnet connection with a back end IoTBOX and forwards the command to it. The IoTBOX is a set of sandbox environments that run Linux operating system for embedded devices with different CPU architectures. Front end Responder forwards a response from IoTBOX to the client. Since the incoming commands, which are forwarded to IoTBOX may cause malware infections or system alteration, the OS image is reset occasionally.

Profiler analyses the interaction between the Front end Responder and IoTBOX. The incoming command and corresponding response are extracted. Subsequently, this information is added to the command interaction profile, such that in the future the Front end Responder can handle the same command without interacting with IoTBOX. Furthermore, Profiler collects banners from devices in the internet.

The Downloader component analyses the interactions for download triggers of remote files, such as malware binaries.

Finally, the Manager handles configuration of IoTPOT. Namely, it links IP addresses to specific Device Profiles (Pa et al., 2016).

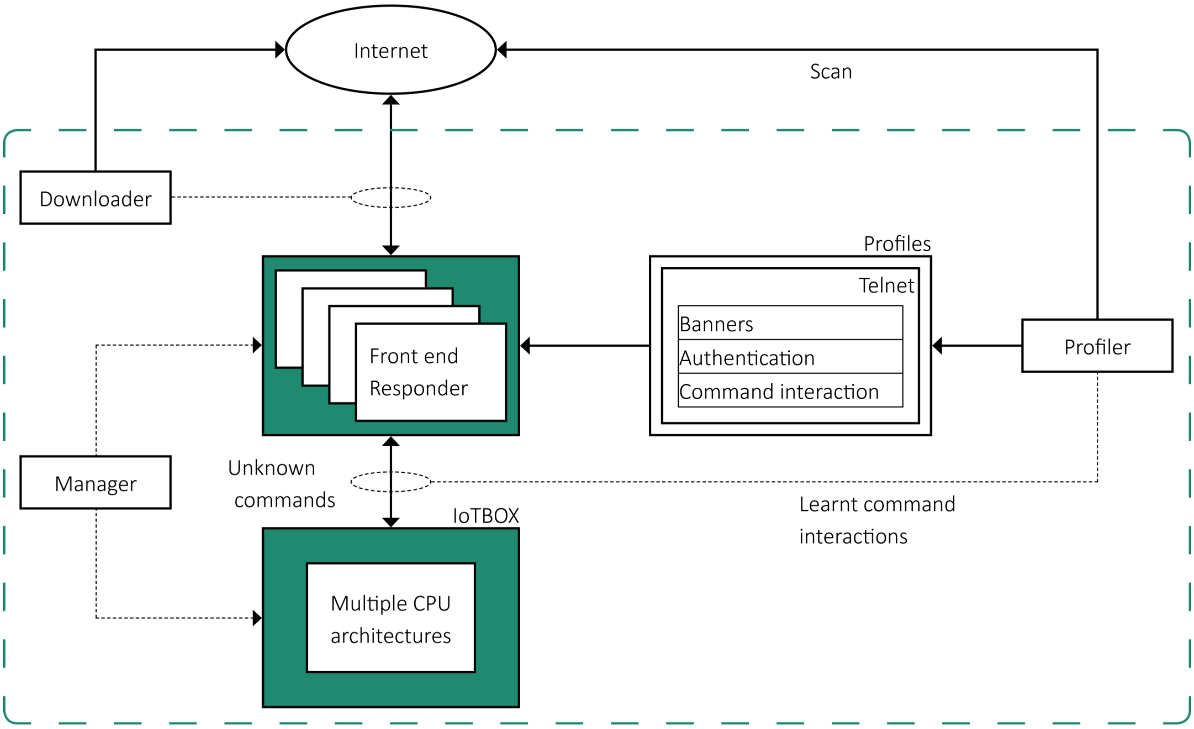


Figure 17 Overview of IoTPOT. Adapted from Pa et al. (2015)

APPENDIX I

QUESTIONS EXPERT INTERVIEW I

Abuse reports

1. Which IoT abuse data does KPN receive?
2. How does KPN distinguish IoT infections from other types infection types (e.g. malware types)?
3. Which other mechanisms does KPN use to detect IoT infections (e.g. honeypot)?
4. Which IoT abuse data does KPN use?
5. How many KPN customers infected with IoT malware does KPN see in the dataset?

Notification and remediation

6. If a customers is known to be infected with IoT malware, how does KPN decide whether or not to notify the customer?
7. Which communication channel is used to notify the customer (e.g. email, walled garden)?
8. What is the content of the notifications?
9. To what extent is the content different for different types of devices or infections?
10. What kind of additional help does your organisation provide, such as providing links to infected device owners to get professional help in case of infection?
11. When does KPN decide to repeat a notification if an infection is not removed?
12. To what extent is the content and communication channel different from the first notification?
13. Does KPN ever advices to disconnect a device from the internet? And if yes, when?
14. How do customers respond to the notifications? Are customers able to remove the infections?
15. Which communication channel do customers use to contact KPN?
16. How does KPN know whether a customer has successfully removed the infection?
17. How does KPN track the presence of IoT malware infections?
18. In case of walled garden notification, when is a customer allowed to leave the walled garden?
19. Which additional problems does KPN face in terms of notifying customers and remediating infected IoT devices?

Ideas

20. Which ideas does KPN have in terms of improving or changing the notification content or the communication channel?

APPENDIX J
QUESTIONS EXPERT INTERVIEW II

Notification channel

1. Strengths and weaknesses of the current channel: currently, KPN's abuse team uses a walled garden notification for the consumer market (with a daily limit of 100) and sends an email notification simultaneously in case of infection in the consumer market. What are the strengths and weaknesses of these communication channels?
2. Evaluation other possible channels: other communication channels found in the literature include: (1) postal mail; (2) telephone calls; (3) SMS; and (4) instant message notifications. What are the strengths and weaknesses of these notification channels? Has it ever been considered to use any of these channels?

Notification content

1. Strengths and weaknesses of the current notification content: what are the strengths and weaknesses of the current notification message for customers with infected IoT devices?
2. Brainstorm improvements: do you have any ideas in terms of improvements? What information would customers need to be able to do cleanup themselves?

APPENDIX K

INTERVIEW PROTOCOL CUSTOMER INTERVIEWS

K.1 ENGLISH VERSION

Good morning/good afternoon/good evening, this is Lisette from KPN.
On [date], KPN has send you an email (and partially blocked your Internet connection), because a virus was found on your Internet connected devices.

KPN notices that sometimes customers find it difficult to understand this message, because it includes technical information.

In order to improve the process where possible, I would like to ask you some questions about how you experienced the message from KPN. The conversation will take less than five minutes.
Do you have time?

Acknowledgement

1. Do you remember receiving the message?
   a. *Answer is no:* The message was sent to [email address]. Is this the correct email address?
2. Do you remember reading the message?
   a. *Answer is no:* Why not? Didn't you trust that KPN was the sender of the message, did you think it was spam, or something else?

Understanding

3. Did you take any action after reading the message, if so, can you please explain to me what you did?
   a. *Answer is nothing:* Did someone else take any action? If so, can you please explain to me what he/she did?

Coding scheme
3.1. Determine which device(s)
3.2. Change password device(s)
3.3. Restart device(s)
3.4. Reset modem/router
3.5. Set password modem/router
3.6. Other
3.7. DK
3.8. Nothing

4. While cleaning the virus, have you needed any additional materials such as searching on Google, assistance of someone else or searching in the paper manual?

Coding scheme
4.1. Search on Google (Internet)
4.2. Assistance of someone else
4.3. The paper manual
4.4. DK

Computing expertise

5. How confident are you about your technical ability to solve issues like this one?

Coding scheme
5.1. Very confident
5.2. Fairly confident
5.3. Not very confident
5.4. Not at all confident
5.5. DK

Suggestions

6. How could the communication to customers be improved when KPN see problems like this? What are your suggestions?

Okay, thank you for your time. Your opinion and experiences will be used to improve the process where possible.
I wish you a nice day/evening.

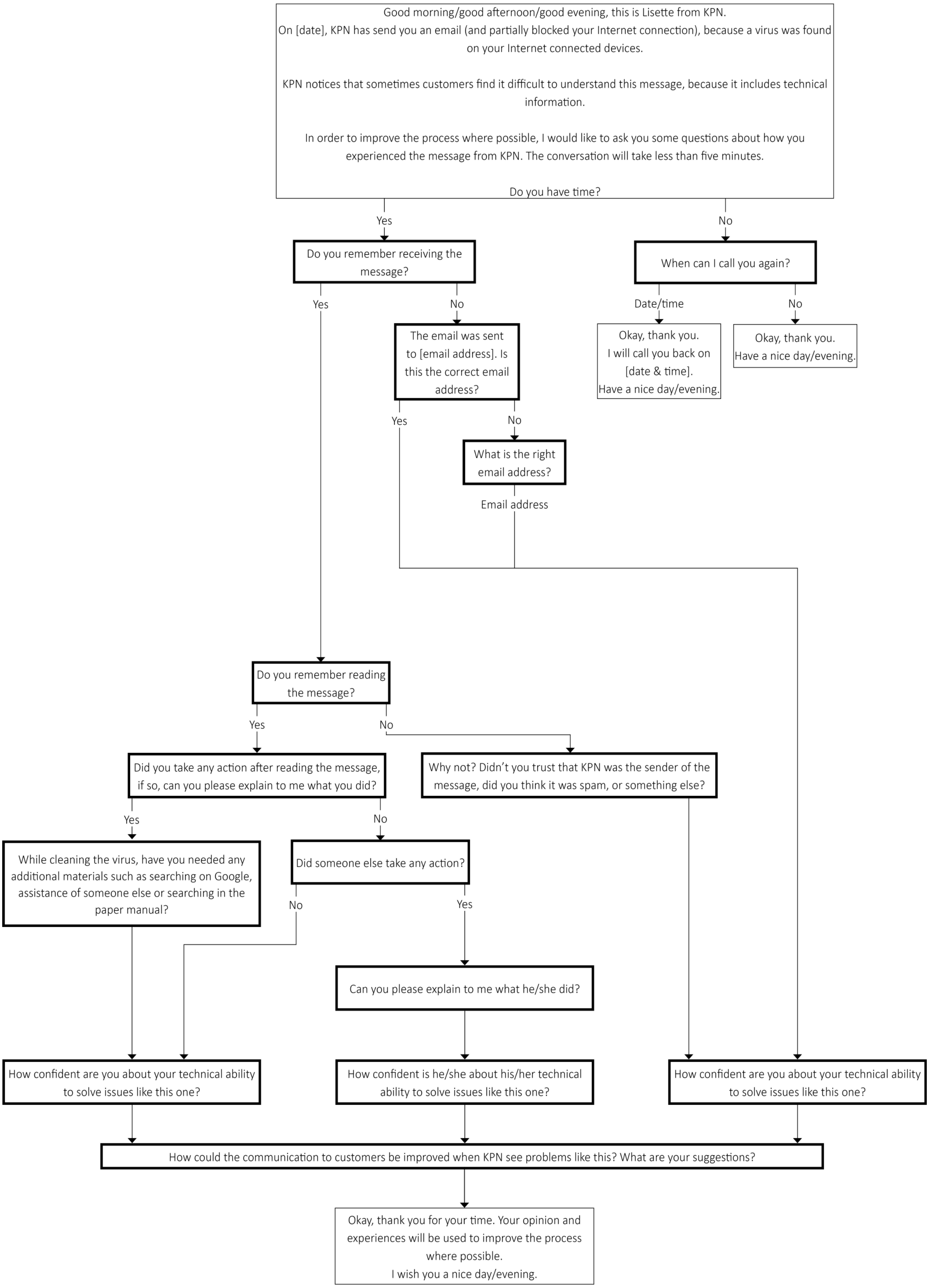

Good morning/good afternoon/good evening, this is Lisette from KPN.
On [date], KPN has send you an email (and partially blocked your Internet connection), because a virus was found on your Internet connected devices.
KPN notices that sometimes customers find it difficult to understand this message, because it includes technical information.
In order to improve the process where possible, I would like to ask you some questions about how you experienced the message from KPN. The conversation will take less than five minutes.
Do you have time?
Yes
No
Do you remember receiving the message?
When can I call you again?
Yes
No
Date/time
No
The email was sent to [email address]. Is this the correct email address?
Okay, thank you. I will call you back on [date & time]. Have a nice day/evening.
Okay, thank you. Have a nice day/evening.
Yes
No
What is the right email address?
Email address
Do you remember reading the message?
Yes
No
Did you take any action after reading the message, if so, can you please explain to me what you did?
Why not? Didn't you trust that KPN was the sender of the message, did you think it was spam, or something else?
Yes
No
While cleaning the virus, have you needed any additional materials such as searching on Google, assistance of someone else or searching in the paper manual?
Did someone else take any action?
No
Yes
Can you please explain to me what he/she did?
How confident are you about your technical ability to solve issues like this one?
How confident is he/she about his/her technical ability to solve issues like this one?
How confident are you about your technical ability to solve issues like this one?
How could the communication to customers be improved when KPN see problems like this? What are your suggestions?
Okay, thank you for your time. Your opinion and experiences will be used to improve the process where possible.
I wish you a nice day/evening.

Figure 18 Flow diagram customer interviews

K.2 DUTCH VERSION

Goedemorgen/goedemiddag/goedenavond, u spreekt met Lisette van KPN.
Op [datum], heeft KPN u een e-mail gestuurd (en u gedeeltelijk uw internet verbinding geblokkeerd), omdat er op uw op het internet aangesloten apparaten een virus gevonden is.

KPN merkt dat klanten soms moeite hebben met het begrijpen van het bericht, omdat het technische informatie bevat.

Om het proces te verbeteren, zou ik u graag een aantal vragen willen stellen over hoe u het bericht van KPN heeft ervaren. Het gesprek zal minder dan vijf minuten in beslag nemen.

Heeft u tijd?

Bevestiging

1. Herinnert u zich dat u het bericht heeft ontvangen?
   a. *Antwoord is nee:* Het bericht is verzonden naar [e-mailadres]. Is dat het juiste e-mailadres?
2. Herinnert u zich dat u het bericht heeft gelezen?
   a. *Antwoord is nee:* Waarom niet? Vertrouwde u niet dat KPN de afzender van het bericht was, dacht u dat het spam was of iets anders?

Actiegerichtheid van de inhoud van de notificatie

3. Heeft u actie ondernomen na het lezen van het bericht, en zo ja, kunt u mij uitleggen wat u heeft gedaan?
   a. *Antwoord is niets:* Heeft iemand anders actie ondernomen? Zo ja, kunt u mij uitleggen wat hij/zij heeft gedaan?

Coderingsschema
3.1. Bepalen welke apparaten
3.2. Veranderen wachtwoorden apparaten
3.3. Herstarten apparaten
3.4. Resetten modem/router
3.5. Instellen wachtwoord modem/router
3.6. Anders
3.7. Weet niet
3.8. Niets

4. Heeft u aanvullende informatie nodig gehad om het virus te verwijderen, zoals bijvoorbeeld opzoeken op Google, de hulp van iemand anders of het opzoeken in de papieren handleiding?

Coderingsschema
4.1. Zoeken met Google
4.2. De hulp van iemand anders
4.3. De papieren handleiding
4.4. Weet niet

Technische kennis

5. Hoeveel vertrouwen heeft u in uw technische vermogen om problemen zoals deze op te lossen?

Coderingsschema
5.1. Zeer zelfverzekerd
5.2. Vrij zelfverzekend
5.3. Niet erg zelfverzekerd
5.4. Helemaal niet zelfverzekerd
5.5 Weet niet

Suggesties

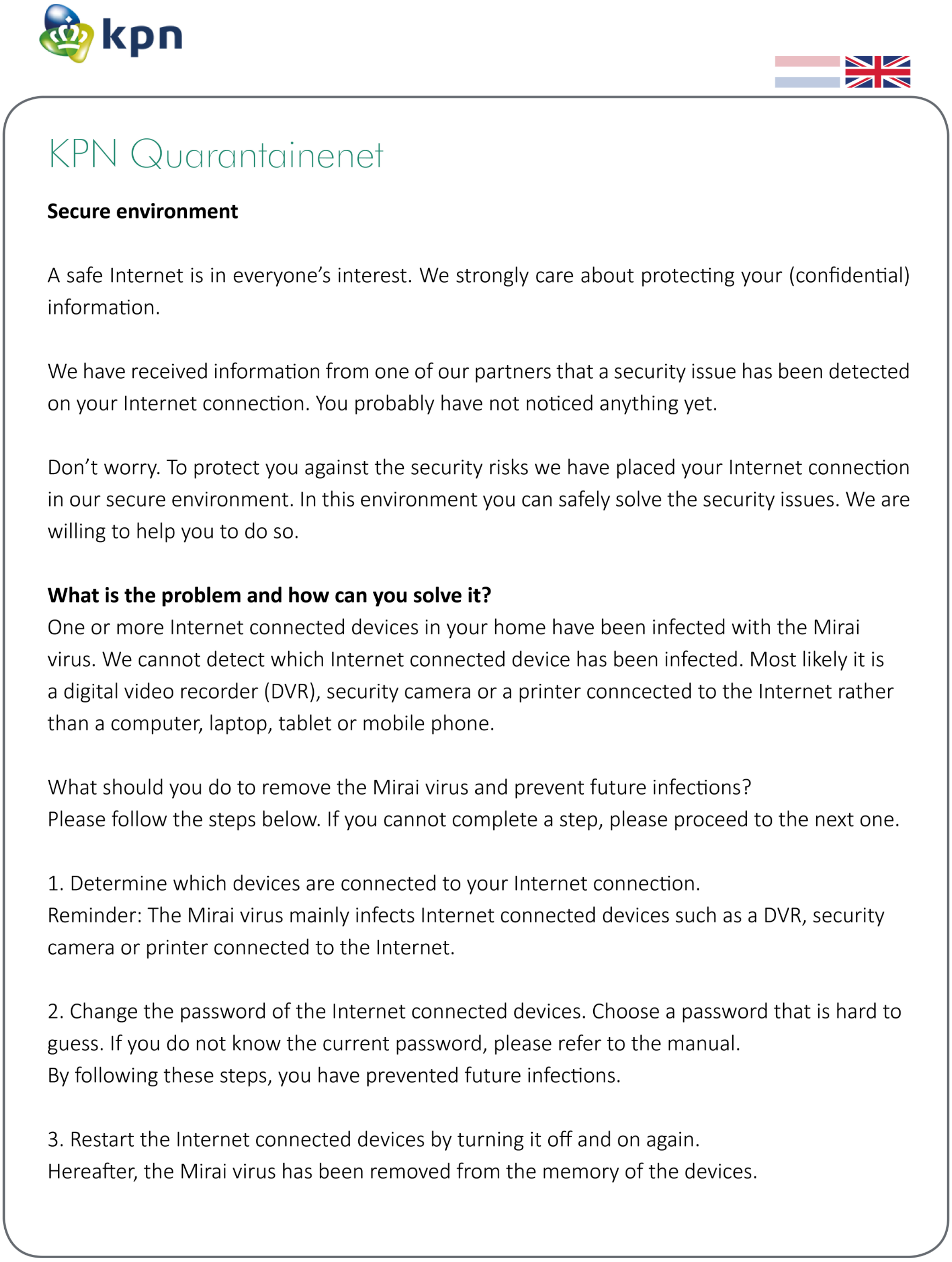
6. Hoe kan de communicatie naar klanten worden verbeterd wanneer KPN dit soort probleem ziet? Wat zijn uw suggesties?

Oké, bedankt voor uw tijd. Uw mening en ervaringen zullen gebruikt worden om het proces waar mogelijk te verbeteren.
Ik wens u een fijne dag/avond.

APPENDIX L

NEW LANDING PAGE QUARANTINE AREA FOR MIRAI INFECTION

L.1 ENGLISH VERSION



KPN Quarantainenet

Secure environment

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

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

Don't worry. To protect you against the security risks we have placed your Internet connection in our secure environment. In this environment you can safely solve the security issues. We are willing to help you to do so.

What is the problem and how can you solve it?
One or more Internet connected devices in your home have been infected with the Mirai virus. We cannot detect which Internet connected device has been infected. Most likely it is a digital video recorder (DVR), security camera or a printer conncected 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.

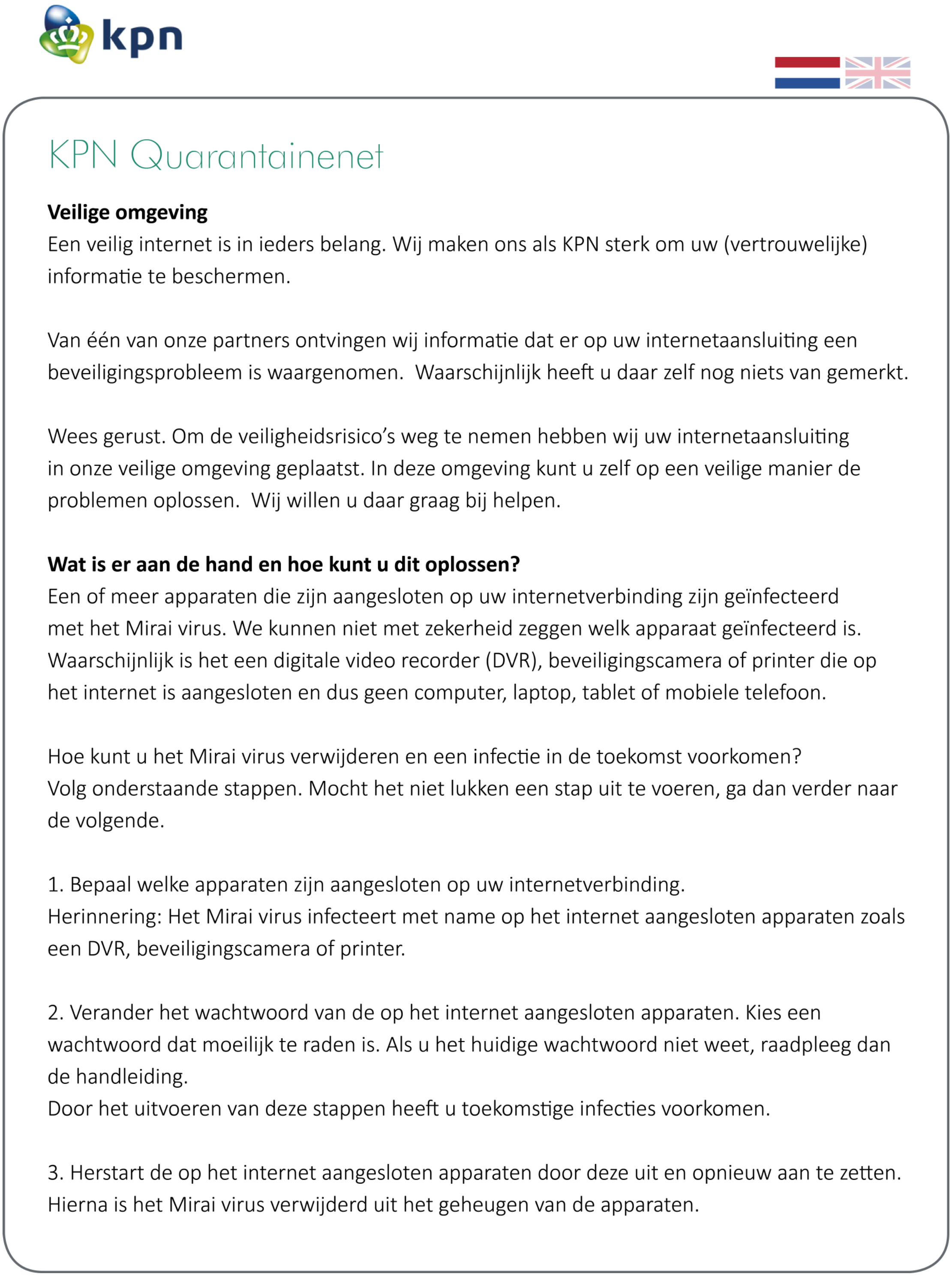
Now that your Internet connected devices are safe, the last steps are to protect your router/ modem.
4. Reset your modem/router to the factory settings. On https://forum.kpn.com/internet-9/hoe-reset-ik-de-kpn-experia-box-modem-97446 it is described how you do this for an Experia Box.
5. Set the password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

Necessary steps
1. Take the measures stated above
2. Fill in our form (and restore your Internet connection)

General security tips
* Use an up-to-date virus scanner to keep out potential hazards
* Keep computer software, like your operating system, up to date
* Do not open messages and unknown files that you do not expect or trust
* Secure your wireless connection with a unique and strong password

L.2 DUTCH VERSION



KPN Quarantainenet

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

Van één van onze partners ontvingen wij informatie dat er op uw internetaansluiting een beveiligingsprobleem is waargenomen. Waarschijnlijk heeft u daar zelf nog niets van gemerkt.

Wees gerust. Om de veiligheidsrisico's weg te nemen hebben wij uw internetaansluiting in onze veilige omgeving geplaatst. In deze omgeving kunt u zelf op een veilige manier de problemen oplossen. Wij willen u daar graag bij helpen.

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://forum.kpn.com/internet-9/reset-de-kpn-experia-box-modem-97446#M8199 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.

Noodzakelijke stappen

1. Voer de bovenstaande maatregelen uit.
2. Vul ons contactformulier in (en herstel uw internetaansluiting).

Algemene beveiligingstips

* Gebruik een up-to-date virusscanner. Zo houdt u gevaren buiten de deur.
* Houd computersoftware, zoals uw besturingssysteem, up-to-date.
* Open geen berichten en onbekende bestanden die u niet verwacht of vertrouwt.
* Beveilig uw draadloze verbinding met een (uniek en niet te achterhalen) wachtwoord.

APPENDIX M

NEW EMAIL NOTIFICATION MIRAI INFECTION

********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://forum.kpn.com/internet-9/reset-de-kpn-experia-box-modem-97446#M8199 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.

Met vriendelijke groet,

KPN Abuse Team
abuse@kpn.com

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

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

Dear Sir/Madam,

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.

Kind regards,

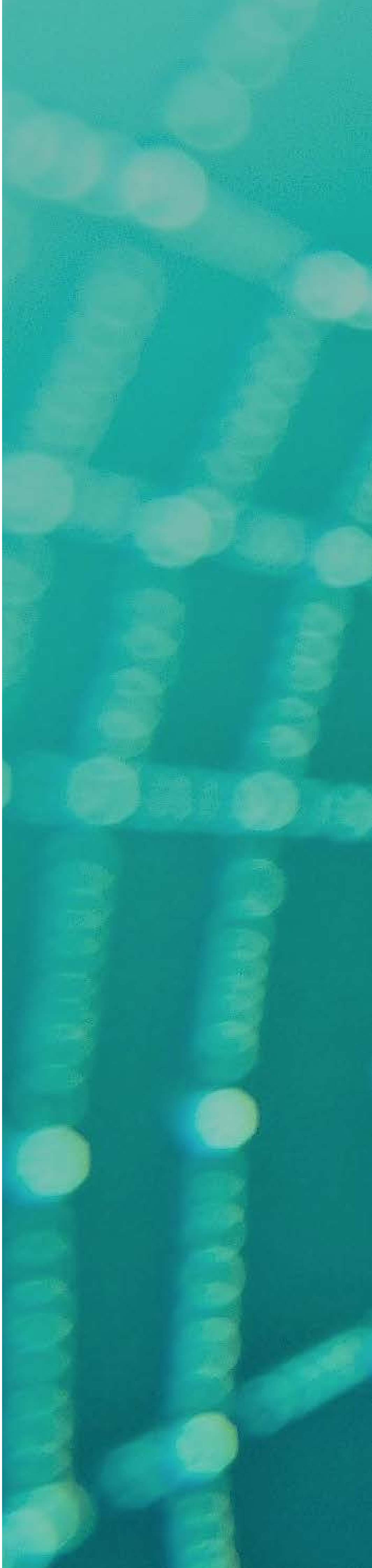
KPN Abuse Team
abuse@kpn.com

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

APPENDIX N

COMMUNICATION LOGS COLLECTION PROCEDURE

CONFIDENTIAL



TU Delft

kpn