Problem Analysis

Robert Nedergaard Nielsen

Magnus Stensli

Daniel Britze

Peter Møller

Jacob Vejlin Jensen

Aalborg University

1 Abstract

This project is a second semester scientific dissertation with vision to combine modern machine learning technologies and human interaction to develop an innovative network security solution.

Part of the EU ERASMUS + funded program EPIC

2 Introduction

By discussing the current cyber-threat and conducting initial research of the subject at the introductory seminar, an initializing problem definition is formed:

"Android smartphones are vulnerable to malware attacks on a corporate level as well as a personal level. It leads a constant market need for improving already existing ideas as well as defining new ones, keeping up to date with modern technology."

To engage on a supplementary analysis of the defined problem, malicious software is firstly studied with regard to different malware types. Additionally, market factors are carefully considered and combined with active use of analytical models, to definitively develop a conceptual model of the envisioned system and a detailed requirement specification.

Table of contents

3 Malware

Malware is a hard to define topic, but this chapter will explain how we in this project defines it. Based on the definition, examples will be given in order to give insight on what different kinds of malware there exists. In the end, it is decided upon what type of malware will be the main focus for this project.

3.1 What is malware?

Malware is a somewhat hard to define category of software. A popular definition is that malware is any kind of software that acts against the users will. But in that definition, begs the question: Are ads malware? Many people suddenly don´t want ads on their computer, hence the invention of ad blockers (LINK). Whilst there remains a grey line it is still possible to derive software types that suddenly can be defined under malware. For example, a piece of software that completely locks a computer until the user has paid a ransom to the person who created the malware.

Whilst the definition of malware leaves a grey zone, many software types can still be defined as malware with some insurance.
Which types of malware can we define then? Taking a starting point on the previous definition, it´s possible to find several types of malware. To describe but a few, there is: Spyware, Miners, Botnets, Ransomware and a lot other types. So, what does this malware do?

3.2 Spyware

Spyware just like all types of malware is difficult to define. As a report from the FTC in 2005 noted:

> "Spyware has evolved to have a variety of meanings. Panelists generally agreed that reaching an industry consensus on one definition has been elusive because of the technical complexity and dynamic nature of software. Several panelists observed that it is also difficult to define spyware because consumers and the business community may differ

on what they believe is appropriate behavior in distributing software and because harmful software may cause a wide variety of problems"[1]

However, in a broad context spyware can still be defined as a type of malware that collects and send data from a user's computer without their permission. This can range from malware that tries to collect a user's password to malware that tries to find incriminating use cases for extortion[2]. From the user's perspective, this can lead to identity theft, corporate espionage to the loss of ones' online systems such as email and social media.

3.3 Miners

Thanks to the rise of cryptocurrencies such as Bitcoin or Ethereum a new form of malware has appeared; mining malware. Cryptocurrencies are "mined" by a computer performing ever more complex math to get said currency. Said currency is then transferred to an anonymous "wallet". This malware is then designed to infect other PC´s and use their computing resources to "mine" the cryptocurrency and then send it to the wallet of the person responsible for the malware. According to anti-malware firm "Checkpoint" the mining malware "Coinhive" is one of the biggest current malware types infected up to "23%" of organizations in January 2018.[3]

---

[1] https://www.ftc.gov/sites/default/files/documents/reports/spyware-workshop-monitoring-software-your-personal-computer-spyware-adware-and-other-software-report/050307spywarerpt.pdf

[2] https://www.wired.com/2010/04/spam-extortion/

[3] https://blog.checkpoint.com/2018/02/15/januarys-most-wanted-malware-cryptomining-malware-continues-to-cripple-enterprise-cpu-power/

3.4 Current trends in malware attacks

Figure 1Malware trends timeline

The anti-virus company Barkly releases reports every year describing what trends are currently shown inside malware development as well as what types of malware are becoming more prevalent (https://www.barkly.com). In 2017 they released the descriptive timeline shown on *figure 1*.

From this timeline and the report, it´s pretty clear that ransomware is on the rise. With Wannacry and NotPetya hitting a huge amount of computer systems. Both of these are types on ransomware. Intermedia (a cloud hosting service) also reports that "Nearly half of ransomware attacks now infect at least 20 employees in an organization" (https://www.intermedia.net/report/ransomware).

The same article published by intermedia also sees ransomware as being in growth with no sign of stopping. Furthermore, bigger businesses are being targeted as intermedia describes it:

> "The criminals behind ransomware are going after businesses of all sizes. 89% of the businesses hit by ransomware had 10 employees or more, while 60% had more than 100 employees. And ransomware tends to hit multiple users at once; 75% of outbreaks affected three or more people, and 47% of outbreaks spread to at least 20 people."

From these reports, it seems of the upmost importance that something is done to prevent ransomware attacks. Therefore it´s necessary to go deeper into exactly what ransomware is.

4 Ransomware

Ransomware is one of the most well-known types of malware, this section will do a short walk-through of how ransomware works and who the main target of this type of malware is. Ransomware is a type of malicious software, were an attacker is blocking a user's access to their system, whether is a computer or a phone. This can happen if the attacker has had access to the user's system and encrypted some of the system files. The user is here after forced to pay a ransom to get the decryption key. The ransom is often asked to be paid in cryptocurrencies like Bitcoins because it is almost impossible to trace the transactions.

There is several ways ransomware can infect a system. This can happen through infected links, attached files in emails and other messages. A system can also be infected if the user visits an infected website. This is some reasons why it is important to be proactive when it comes to malware. One of the best ways to defend yourself against ransomware is backing up your computer so you can restore it to an earlier point of time, and minimize the loss of data.

Ransomware is both targeting private individuals and companies. The main goal of ransomware is often to earn a lot of money, therefore is attackers often targeting companies and authorities. One example on ransomware targeting windows systems is the well know attack called PetYa, this attack uses an exploit called EternalBlue. This exploit is using a weakness in the way Microsoft is implementing the Server Message Block(SMB) Protocol. The way SMB is implemented is making Windows mishandle some of the packets and allow attackers to run malicious code on a user's system. This exploit was also used in the WannaCry attack.

Petya is assumed to target the Ukrainian authorities, because of the day the attack was their constitution day. Attacks of same type is traced back Germany, England, Russia and Denmark. One of the most know Danish companies that was hit is A.P. Møller-Mærsk.

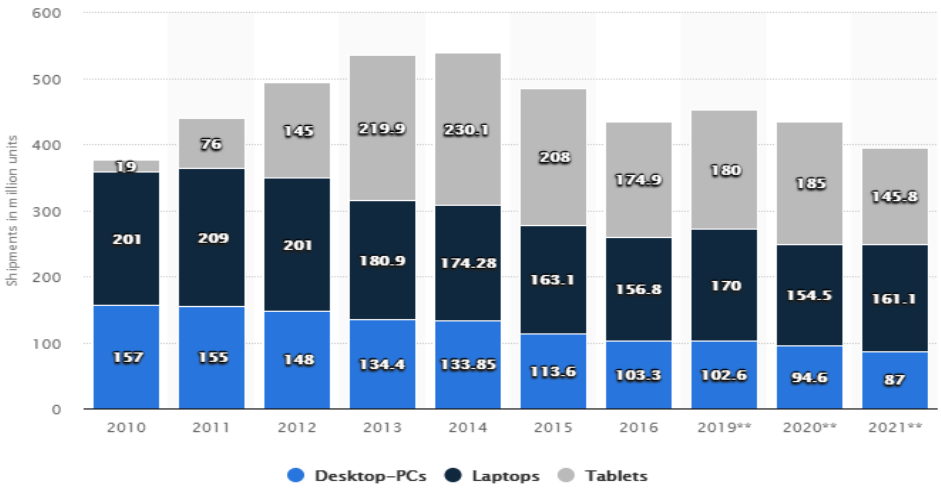


Figure 2

Ransomware is not only targeting Windows systems but also systems like android phones. And with the rising sale of android devices, which can be seen on *figure 2,* will this become a problem very quick.

Now there isn't recorded that many ransomware attacks on android devices, but based on the popularity of ransomware on PC's, it would be expected to also turn into a problem on android devices.

One example on ransomware on android devices is an attack named Android/FakeAV.E. This virus spreads by pretending to be an app to the adult video website PornHub. When the user wants to connect to the sites it gives an error message as seen on *figure 3*.

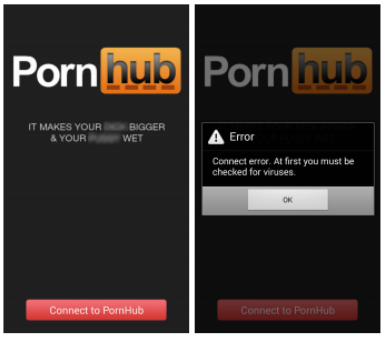


Figure 3 Infected website

After this error message it demands that you download "Avast antivirus". After the fake Avast GUI is done running its scam scan it shows a message saying “device is in danger and is now blocked for security reasons” and states that you must buy the pro version, this can be progress can be seen on *figure 4*.

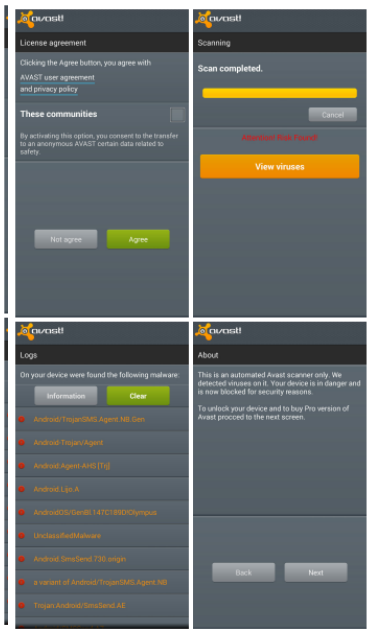


Figure 4Fake Avast GUI

After wards the payment message seen on *figure 5* , will be shown, this message is a direct copy of an earlier ransom message with the same typographic errors.

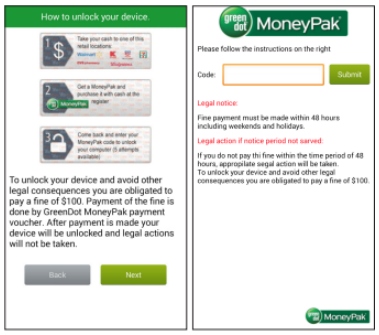


Figure 5 Ransom payment

5 Stakeholders

In chapter 3 Malware the global risk of cyberattacks is elaborated upon. By researching the current cyber-threats, their rates of occurrence and the extent of their affiliated consequences, it is concluded that malware of the type Ransomware is to be considered a global cyber-threat. An immediate response from the counterworking field of cybersecurity is a necessity, to effectively build defensive structures securing network systems from being hit by destructive ransomware attacks.

From supplementary research of the technology behind ransomware and previous attacks in chapter 4 Ransomware, it is concluded that ransomware attacks are of common occurrence and often associated with massive consequences. Furthermore, looking at market statistics illustrating the progress of smartphone and pc sales over the past decade gives reason for the establishment of another demarcation.

5.1 Market statistics

Companies as well as regular consumers have a lot of sensitive data stored on their smartphones, for example (but not limited to):

- Credit card data
- Compromising photos

- Passwords
- E-mails
- Customer data

This in itself is worrisome, however the popularity of smartphones is rising as well. At this point in time smartphones outsell PC's. Computers have been susceptible to malware since they were first introduced to the market, however it seems that Android phones, while vastly gaining popularity over PC's, are still relatively under protected to Malware attacks. This is represented in the graph below, which is derived from the AAU HoneyPot 1st semester dissertation.

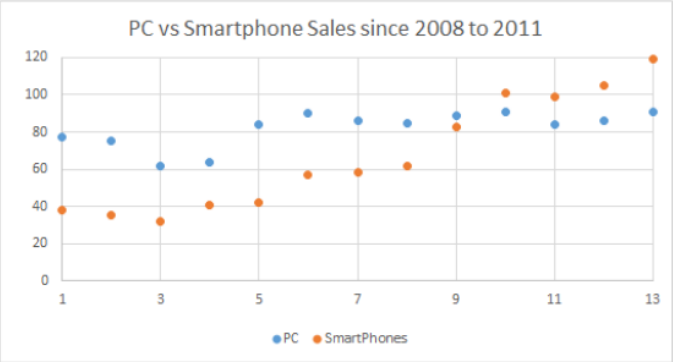


Figure 6 Smartphone vs Pc sales statistic

A security threat report published by Symantec released April 2017, exemplifies the vulnerability of the Android operating system in the current threat landscape.

> *"The overall volume of the malicious Android apps increased significantly In 2016, growing by 105 percent."*[4]

The combination of a quickly evolving market of Android devices and a malware attack threat level rising drastically every year, it is of growing concern to develop new innovative technologically advanced cybersecurity systems. Adding the assumption that ransomware, due to extensive occurrence and destructive consequences on PC systems, is expected to play an increasingly immense role in malware attacks targeting Android devices. From this hypothesis, the project is demarcated to focus on creating a security

[4] *https://www.symantec.com/content/dam/symantec/docs/reports/istr-22-2017- en.pdf*

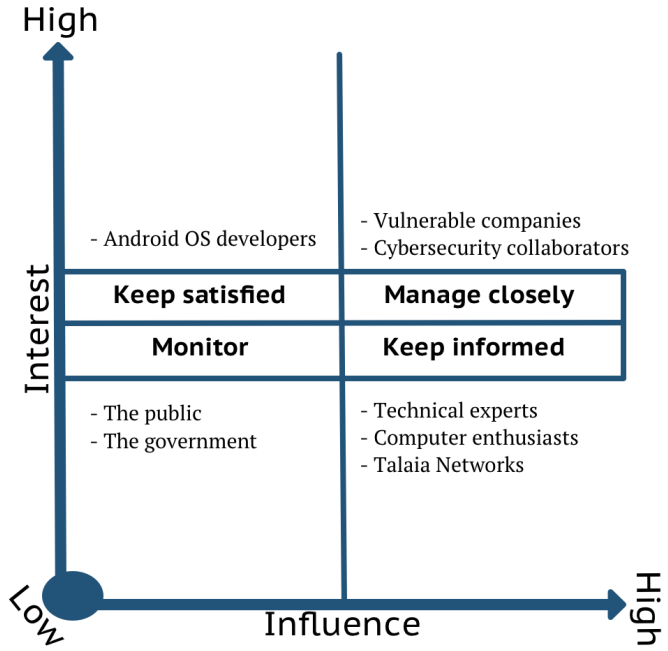
system with the ability to detect/predict ransomware attacks on Android devices and thereby improving overall security on the operating system.

5.2 Stakeholder identification

To identify relevant stakeholders a structured stakeholder analysis is conducted. The general purpose of this analysis is to clarify what potential parties that either has personal interest in a project ultimately focused on improving data security on Android devices, or is able to affect/influence it.

Firstly, stakeholders are identified by carrying out a brainstorming process, and categorized into four categories by rating each stakeholder on two parameters:
1) The stakeholders possible interest in the end product this project has envisioned to develop. (Interest)
2) The stakeholders possible influence on the development process of the envisioned product. (Influence)

Each stakeholder is evaluated on both parameters as either "low" or "high", creating the four categories as shown on the illustrative figure below.



5.3 Categorization

Monitor (Low Interest, Low Influence):

- Stakeholders in this category are not that important. They don't bring a lot to the table, but it's worth monitoring them as they might influence the product in the future. We're not going to ask them anything.

Manage closely (High interest, High influence):

- Stakeholders in this category, are the most important. They bring knowledge to us, regarding what they expect from our product which helps develop something more relevant. Because these stakeholders are also interested in the final product, they want to tell us exactly what they need so we can adapt. We have to manage closely with these stakeholders to keep them interested and to keep getting information. (e.g. a company could tell us they need a design that makes it possible to turn the system on/off remotely, so we know our product has to have this as a feature.)

Keep satisfied (High interest, Low Influence):

- Stakeholders in this category also bring knowledge. These stakeholders function as supporters. They help with every part of the project we're having trouble with. Doesn't matter if it's regarding machine learning, business approach, overall collaboration or help with the server. These stakeholders don't really have an interest in the final product, so we should try to keep them satisfied so they can continue guiding us.

Keep informed (Low interest, High Influence):

- These stakeholders don't supply anything to our project. They can however have an influence on the project. We need to keep close contact, informing and getting informed from these stakeholders, so we make sure no major issues are arising.

5.4 Elaborated explanation of stakeholders

Cybersecurity collaborators

- Other companies also trying to fight malware, have high interest and high influence. They could have an effect on the product, because we could learn a lot from them (like Talaia) but they are also interested in our product because our algorithm could help them as well. Having other Cybersecurity Companies as a

collaborator would be greatly beneficial if they would be willing to share their findings with us in order to improve our algorithm at a more rapid pace.

Vulnerable Companies

- Vulnerable companies (hospitals, banks or just large companies etc.) have high interest and high influence. High value companies such as banks or hospitals are in focus do to the large amount of money that are in play on an everyday basis. This makes them a very profitable target to infect with ex ransomware.

Technical Experts

- Experts have low interest and high influence. They are going to support us on developing our product. They can bring knowledge and teach us how to setup more technical issues of the project. They won't get affected by our product.

Computer enthusiasts

- Same as experts they have low interest and high influence. They won't get affected by our product, but can bring knowledge, and supply with ideas and methods.

Android OS developers

- Android developers have high interest and low influence. We need to keep the satisfied and have communication, without boring them. It will have relevancy for developers of android devises as they will be able to show improvement in their security when they try to sell their product to customers.

Talaia Networks

- Talaia have low interest and high influence. They function as experts and will bring knowledge and help us understand how to go to the market with a product like this.

The public

- The public have low interest and low influence. They will get affected by our product, but on a small scale. It's worth monitoring these stakeholders to make sure, we're making a product, that the public actually wants.

The government

- The government have low interest and low influence. It's worth monitoring them, as they COULD influence the project later on through the law.

5.5 Conclusion

The categorization identifies two stakeholders to be of high interest and high influence; Vulnerable companies and cybersecurity collaborators. Each of these stakeholders must be manages closely throughout the project. Identifying the most valuable stakeholders leads to a further demarcation of the project.

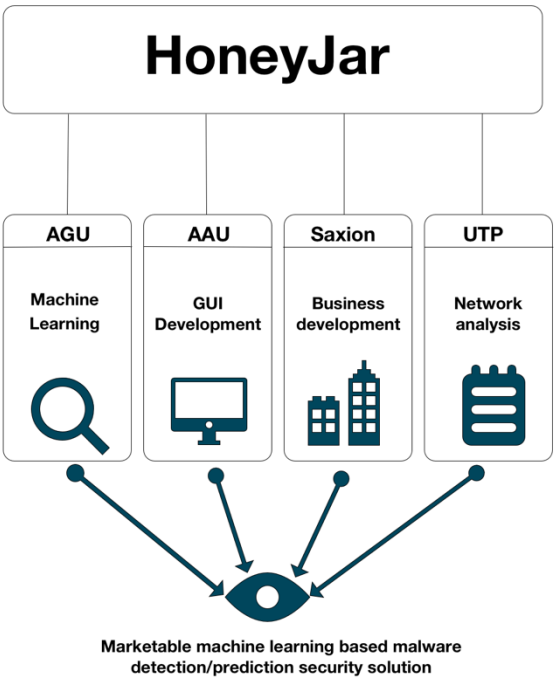
6 Problem demarcation

Combining the needs of the identified most valuable stakeholders, the research concerning ransomware and the market statistics, the following demarcations are established:
- Companies could use a solution allowing less technical experienced operators to interact with a dynamic cybersecurity system.
- It is necessary to create a database able to collect large datasets from multiple collaborators.

To meet these expectations, a conceptual model of the HoneyJar project and the division between EPIC partners is developed.

6.1 HoneyJar

This section will explain how the HoneyJar works. How it can help us prevent malware attacks by examining network traffic. This section will also explain how the HoneyJar concept can be split into smaller projects and a short description of those. This will also be a good occasion to talk about how the different tasks is split between the different universities in this project.



6.2 HoneyPot

In the last section, the concept of HoneyJar was explained, this section will now focus on the smaller part called a HoneyPot, it will be explained how it is connected to the HoneyJar and what it contributes with.

6.3 GUI

This section will examine the pros and cons of using a graphical user interface(GUI) instead off the command line interface(CLI). It will also examine what making a GUI can do for the product from a sales perspective.

This section showed us that making a GUI can make it a lot easier for a normal user to use a product, because they don't have to remember commands, and a lot easier to navigate through file systems.

7 Problem definition

Under work.

8 Requirement specifications

This section will state the requirements that we got from our problem analysis. It will be split in to two parts a part focusing on our GUI and a part focusing on the HoneyPot system that the GUI is going to control.

The requirements for the HoneyPot is essential for it to be a useful tool to fight cybercrime. The focus will be to generate a useful dataset that will show how malware spreads through the network and communicate with a command and control server. The requirements are listed below:

- The HoneyPot can generate a PCAP-file for each machine running.
- The HoneyPot can generate a PCAP-file with the traffic from the whole system.
- The HoneyPot can store the traffic data in a database the is easily accessible.

The requirements for the GUI, will make sure that it will as easy as possible for a user to run a HoneyPot without a lot of knowledge about configuring a system through the CLI and pulling out basic data from a PCAP-file. The requirements are listed below:

- The GUI is able to control the basic functionalities of the HoneyPot.
- The GUI should allow for non-technical operators to interact with the HoneyPot.
- The GUI is able give the user alters about problems in the HoneyPot.
- The GUI will give a graphical overview of the utilization.

With the GUI requirements, out of the way focus can switch to being on how this will be made into a product. For that a marketing analysis will be conducted with the aim being how to sell the product and too who.