



POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Informatica

Tesi di Laurea

Automatic Malware Signature Generation

Relatori

prof. Antonio Lioy
ing. Andrea Atzeni

Michele CREPALDI

ANNO ACCADEMICO 2020-2021

Thanks...

Summary

Summary...

Acknowledgements

Aknowledgments...

Contents

1	Introduction	7
2	Background	8
2.1	Malware	8
2.1.1	Why is Malware used	9
2.1.2	Malware types	9
2.1.3	Malware History	16
2.2	Obfuscation and detection evasion	19
2.2.1	Reverse-Engineering	19
2.2.2	Anti-reversing	19
2.2.3	Anti-disassembly	19
2.2.4	Anti-debugging	23
2.2.5	Anti-virtual machine	27
2.2.6	Packers and unpacking	28
2.2.7	Code Obfuscation	28
2.2.8	Polymorphic and Metamorphic Malware	28
3	Proposed Tool	29
4	Results	30
5	Conclusions	31
	Bibliography	32

Chapter 1

Introduction

Introduction...

Chapter 2

Background

2.1 Malware

Malware, short for *malicious software*, is a general term for all types of programs designed to perform harmful or undesirable actions on a system. In fact in the context of IT security the term *malicious software* commonly means [1]:

Software which is used with the aim of attempting to breach a computer system's security policy with respect to Confidentiality, Integrity and/or Availability.

Malware consists of programming artefacts (code, scripts, active content, and other software) designed to disrupt or deny operation, gain unauthorized access to system resources, gather information that leads to loss of privacy or exploitation, and other abusive behaviour. Malware is not (and should not be confused with) defective software - software that has a legitimate purpose but contains harmful bugs (programming errors).

Different companies, organizations and people describe malware in various ways. For example **Microsoft** defines it in a generic way as:

Malware is a catch-all term to refer to any software designed to cause damage to a single computer, server, or computer network [2].

The **National Institute of Standards and Technology (NIST)**, on the other hand, presents multiple definitions for malware, describing it as "hardware, firmware, or software that is intentionally included or inserted in a system for a harmful purpose" [3].

In another more specific definition **NIST** affirms that Malware is:

A program that is inserted into a system, usually covertly, with the intent of compromising the confidentiality, integrity, or availability of the victim's data, applications, or operating system or of otherwise annoying or disrupting the victim [3].

Notice that, since the attacker can use a number of different means - such as executable code, interpreted code, scripts, macros etc. - to perpetrate its malicious intents, the term **software** should be understood in a broader sense in the above definitions.

Moreover, the computer system whose security policy is attempted to be breached is usually known as the **target** for the malware. Instead, the cybercriminal who originally launched the malware with the purpose of attacking one or more targets is generally referred to as the "*initiator* of the malware". Furthermore, depending on the malware type, the initiator may or may not exactly know what the set of targets is [1].

According to the above definitions software is defined as malicious in relation to an attempted breach of the target's **security policy**. In other words, software is often identified as malware based on its *intended use*, rather than the particular technique or technology used to build it.

2.1.1 Why is Malware used

Generally, cybercriminals use malware to access targets' sensitive data, extort ransoms, or simply cause as much damage as possible to the affected systems.

More generally malware serves a variety of purposes. For example, the most common cybercriminals' uses of malware are: [4]

- **To profit financially (either directly or through the sale of their products or services).** For example, attackers may use malware to infect targets' devices with the purpose of stealing their credit account information or cryptocurrency. Alternatively, they may sell their malware to other cybercriminals or as a service offering (*malware-as-a-service*).
- **As a means of revenge or to carry out a personal agenda.** For example, Brian Krebs of Krebs on Security was struck by a big DDoS attack in 2016 after having talked about a DDoS attacker on his blog.
- **To carry out a political or social agenda.** Nation-state actors (like state-run hacker groups in China and North Korea) and hacker groups such as Anonymous are a perfect example.
- **As a way to entertain themselves.** Some cybercriminals enjoy victimizing others.

Obviously there are also reasons for non-malicious actors to create and/or deploy some types of malware too - for example it can be used to test a system's security.

2.1.2 Malware types

There are numerous different ways of categorizing malware; one way is by *how* the malicious software spreads. Another one is by what it *does* once it has successfully infected its victim's computers (i.e. what is its payload, how it exploits or makes the system vulnerable).

By how they spread

Terms like *trojan*, *virus* and *worm* are commonly used interchangeably to indicate generic malware, but they actually describe three subtly different ways malware can infect target computers [5]:

- **Trojan horse.** Generally speaking, a *Trojan Horse*, commonly referred to as a "Trojan", is any program that disguises itself as legitimate and invites the user to download and run it, concealing a malicious payload. When executed, the payload - malicious routines - may immediately take effect and cause many undesirable effects, such as deleting the user files or installing additional malware or PUAs (Potentially Unwanted Apps).

Trojans known as *droppers* are often used to start a worm outbreak, by injecting the worm into users' local networks [8].

Trojans may hide in games, apps, or even software patches, or they may rely on social engineering and be embedded in attachments included in phishing emails.

Trojan horses cannot self-replicate. They rely on the system operators to activate. However, they can grant the attacker remote access permitting him to then perform any malicious activity that is in their interest. Trojan horse programs can affect the host in many different ways, depending on the payload attached to them [6].

- **Virus.** The term "computer virus" is used for describing a passive self-replicating malicious program. Usually spread via infected websites, file sharing, or email attachment downloads, it will lie dormant until the infected host file or program is activated. At that point it spreads to other executables (and/or boot sectors) by embedding copies of itself into those files. A virus, in fact, in order to spread from one computer to another, usually relies on the

infected files possibly ending up, by some means or another, in the target system. Viruses are therefore passive. The mean of transport (file, media file, network file, etc.) is often referred to as the virus *vector*. Depending on how complex the virus code is, it may be able to modify its copies upon replication. For the transport of the infected files to the target system(s), the virus may rely on an unsuspecting human user (who for example uses a USB drive containing the infected file) or initiate itself the transfer (for example, it may send the infected files as an e-mail attachment) [1].

Viruses may also perform other harmful actions other than just replicating, such as creating a backdoor for later use, damaging files, stealing information, creating botnets, render advertisements or even damaging equipment.

- **Worm.** On the other hand, a worm is a self-replicating, active malicious program that exploits various system vulnerabilities to spread over the network. Particularly, it relies on vulnerabilities present in the target's operating system or installed software. Worms usually consume a lot of bandwidth and processing resources due to continuous scanning and may render the host unstable, sometimes causing the system to crash. Computer worms may also contain "payloads" to damage the target systems. Payloads are pieces of code written to perform various nefarious actions on the affected computers among which stealing data, deleting files or creating bots - which can lead the infected systems to become part of a botnet [6].

These definitions lead to the observation that viruses require *user intervention* to spread, whereas a worm spreads itself automatically. A virus, however, cannot execute or reproduce unless the application it infected is running. This dependence on a host program makes viruses different from trojans, which require users to download them, and worms, which do not use applications to execute.

Furthermore, attackers can also install malware "manually" on a computer, either by gaining physical access to the target system or by using privilege escalation methods to obtain remote administrator access [7].

By what they do

There are a wide range of potential attack techniques used by malware, here are some of them:

- **Adware.** *Adware*, or "Advertising supported software", is any software package which automatically plays, displays, or downloads advertisements to a computer. Some adware may also re-direct the user's browser to dubious websites. These advertisements can be in the form of a pop-up ads or ad banners in websites, or advertisements displayed by software, that lure the user into making a purchase. The goal of Adware is to generate revenue for its author.

Often times software and application authors offer "free", or discounted, versions of their creations that come bundled with adware. Adware, in fact, is usually seen by the developers as a way to recover development costs. The income derived from ads may motivate the developer to continue developing, maintaining and upgrading his software product. On the other hand users may see advertisements as annoyances, interruptions, or as distractions from the task at hand [8].

Adware, by itself, is annoying but somewhat harmless, since it is solely designed to deliver ads; however, adware often comes bundled with spyware (such as keyloggers), and/or other privacy-invasive software that is capable of tracking user activity and steal information. Adware-spyware bundles are therefore much more dangerous than adware on its own [14].

- **Backdoor.** A *backdoor*, also called Remote Access Trojan (RAT), is a vulnerability deliberately buried into software's code that allows to bypass typical protection mechanisms, like credentials-based login authentication. In other words, it is a method of circumventing normal authentication procedures. Once a system has been compromised (by others types of malware or other methods), one or more backdoors may be installed. This is done with

the purpose of allowing the attacker easier access in the future without alerting the user or the system's security programs. Moreover, backdoors may also be installed before other malicious software, to allow attackers entry [8].

Many device or software manufacturers ship their products with intentionally hidden backdoors to allow company personnel or law enforcement to access the system when needed [9]. Alternatively, backdoors are sometimes hidden in programs also by intelligence services. For example, Cisco network routers, which process large volumes of global internet traffic, in the past were equipped with backdoors intended for US Secret Service use [10].

However, when used by malicious actors, backdoors grant access to attackers without the user knowledge, thus putting the system in real danger.

- **Browser Hijacker.** A *Browser Hijacker*, also called "hijackware", is a type of malicious program which considerably modifies the behaviour of the victim's web browser. For example it can force the browser to send the user to a new search page, slow down the loading, change the victim's home page, install unwanted toolbars, redirect the user to specific sites, and display unwanted ads without the user consent.

It can be used to make money off ads, to steal information from users, or to infect the systems with other malware by redirecting users to malicious websites [9].

- **Bots/Botnet.** In general, *bots* (short for 'robots') are software programs designed to automatically perform specific operations. Bots were originally developed to programmatically manage chat IRC channels - Internet Relay Chat: a text-based communication protocol appeared in 1989.

Some bots are still being used for legitimate and harmless purposes such as video programming, video gaming, internet auctions and online contest, among other functions. It is however becoming increasingly common to see bots being used maliciously. Malicious bots can be (and usually are) used to form botnets. A botnet is defined as a network of host computers (zombies/bots) that is controlled by an attacker - the *bot-master* [6]. Botnets are frequently used for DDoS (Distributed Denial of Service) attacks, but there are other ways that botnets can be useful to cybercriminals: [4]

- **Brute force & credential stuffing** - Bots can be used to carry out different types of brute force attacks on websites. For example they can use a pre-configured list of usernames and passwords combinations on website login pages with the hope of finding a winning combination, after enough tries.
 - **Data and content scraping** - Botnets can be used as web spiders to scour websites and databases to gather useful information - such as site content, pricing sheets, etc. - which can be used to obtain an unfair advantage against the competition.
 - **Botnet-as-a-service opportunities** - Botnets are sometimes rented out by their creators to all kinds of malicious users - including less tech-savvy ones. Doing so, even inexperienced attackers can carry out attacks, such as taking down a target's servers and networks with a DDoS, using these mercenary bots. This service model is sometimes called malware-as-a-service.
 - **Spambot** - A botnet can also be used to act as a spambot and render advertisements on websites.
 - **Malware distributor** - Finally Botnets can even be used for distributing malware disguised as popular search items on download sites.
- **Crypto-miner.** Crypto-miners are a relatively new family of malware. Cybercriminals employ this type of malicious tools to mine Bitcoin and/or other bitcoin-alike digital currencies on the target machine. The victim system's computing power is used for this, without the owner realising it. The mined coins end up in the attackers' digital crypto wallets.

Recently, a more modern method of crypto-mining that works within browsers (also called crypto-jacking), has become quite popular.

In some cases, the use of crypto-miners may be deemed legal. For example they could be used to monetize websites, granted that the site operator clearly informed visitors of the use of such tools [10].

Finally, according to ESET, most crypto-miners focus mostly on *Monero* as target cryptocurrency because it offers anonymous transactions and can be mined with regular CPUs and GPUs instead of expensive, specialized hardware [4].

- **File-less malware.** File-less malware is a type of memory-resident malware that uses legitimate code already existing within the target computer or device to carry out attacks. As the term suggests, it is malware that operates from a victim's computer memory, not from files on the hard drive, taking advantage of legitimate tools and software (known as "LOLBins" [4]) that already exist within the system. File-less malware attacks leave no malware files to scan and no malicious processes to detect. Since there are no files to scan, it is harder to detect and remove than traditional malware; this makes them up to ten times more successful than traditional malware attacks [11]. Furthermore, it also renders forensics more difficult because when the victim's computer is rebooted the malware disappears.
- **Keylogger.** Keystroke logging (often called *keylogging*) is the action of secretly tracking (or logging) keystrokes on a keyboard, without the person using the keyboard knowing that its actions are being monitored. The collected information is stored and then sent to the attacker who can then use the data to figure out passwords, usernames and payment details, for example. There are various methods used to perform keylogging, ranging from hardware and software-based approaches to the more sophisticated electromagnetic and acoustic analysis [8]. Key loggers can be inserted into a system through phishing, social engineering or malicious downloads.

There are various methods used to perform keylogging, ranging from hardware and software based approaches to electromagnetic and acoustic analysis.

To this extent keyloggers can be considered as a sub-category of spyware.

Keylogging also has legitimate uses, in fact it is often used by law enforcement, parents, and jealous or suspicious spouses. The most common use, however, is in the workplace, where employers monitor employee use of company computers.

- **RAM Scraper.** *RAM scraper* malware, also known as *Point-of-Sale (POS)* malware, targets POS systems like cash registers or vendor portals, harvesting data temporarily stored in RAM (Random Access Memory). Doing so the attacker can access unencrypted credit card numbers [9].
- **Ransomware.** *Ransomware*, also known as "encryption" or "crypto" Trojan, is a malicious program that, after having infected a host or network, holds the system captive and requests a ransom from the host/network users. In particular it encrypts data on the infected system (or anyway locks down the system so that the users have no access) and only unblocks it when the correct password - decryption key - is entered. The latter is not given to the victims until after they have paid the ransom to the attacker. Messages informing the system user of the attack and demanding a ransom are usually displayed. Without the correct decryption key, it's mathematically impossible for victims to decrypt and regain access to their files.

Digital currencies such as Bitcoin and Ether are the most common means of payment, making it difficult to track the cybercriminals. Moreover, paying the ransom does not guarantee the user to receive the necessary decryption key or that the one provided is correct and functions properly. Additionally, some forms of ransomware threaten victims to publicize sensitive information within the encrypted data.

Ransomware is one of the most profitable, and therefore one of the most popular, and dangerous kinds of malware programs of the past few years.

The "Five Uneasy E's" of ransomware, according to Tim Femister [12] - vice president of digital infrastructure at ConvergeOne - are:

- **Exfiltrate:** Capture and send data to a remote attacker server for later leverage.
- **Eliminate:** Identify and delete enterprise backups to improve odds of payment.
- **Encrypt:** Use leading encryption protocols to fully encrypt data.
- **Expose:** Provide proof of data and threaten public exposure and a data auction if payment is not made.

- **Extort:** Demand an exorbitant payment paid via cryptocurrency.

- **Rogue Security Software.** *Rogue Security Software* can be considered as a form of scareware. This type of malware program presents itself as a security tool to remove risks from the user's system. In reality, this fake security software installs more malware onto their system [9].
- **Rootkit.** A *rootkit* is generally thought as a type of malicious software, or a collection of software tools, designed to remotely access or control a computer without being detected by users or security programs. An attacker who has installed a rootkit on a system is able to remotely execute files, log user activities, access/steal information, modify system configurations, alter software (including security software), install hidden malware, mount attacks on other systems or control the computer as part of a botnet. Since a rootkit operates stealthily and continually hides its presence, its prevention, detection and removal can be difficult; in fact, typical security products are often not effective in detecting rootkits. Rootkit detection therefore often relies on manual methods such as monitoring the computer's behaviour for irregular activity, scanning system file signatures, and analysing storage dumps [14].

More recently, the term "rootkit" has also often been used to refer to concealment routines in a malicious program. These routines are highly advanced and complex and are written to hide malware within legitimate processes on the infected computer. In fact, once a malicious program has been installed on a system, it is essential that it remains hidden, to avoid detection and disinfection. The same is true when a human attacker directly breaks into a computer. Techniques known as rootkits allow for this concealment by modifying the host's operating system so that malware is hidden from the user. They can prevent a malicious process from being visible in the system's process list or prevent its files from being read [8].

Traditionally, rootkits can install themselves in kernel level (ring 0), although some sources state that they can install themselves all the way up to user level (ring 3). This means that they can get as much (or as little) access as necessary.

There are different types of rootkits, which are typically categorized by the reach of the system they affect: [9]

- **User-level/application level rootkits** - User-mode rootkits run in Ring 3, along with other applications as user. They can alter security settings, allowing the attacker to replace executables and system libraries.
 - **Kernel-level rootkits** - Kernel-mode rootkits run in ring 0, the highest operating system privileges (Ring 0). They manage to do so by modifying the core functionality of the operating system - the kernel. They usually add code or replace portions of the core operating system, including both the kernel and associated device drivers.
 - **Bootkit rootkits** - A Bootkit rootkit is a type of kernel-mode rootkit which infects startup code like the Master Boot Record (MBR), Volume Boot Record (VBR), or boot sector, subverting the kernel upon computer start up.
 - **Virtualization rootkits** - This type of rootkit, also called *Hypervisor rootkit*, runs in Ring -1 (before the kernel) and hosts the target operating system as a virtual machine. It manages to do so by exploiting hardware virtualization features. This in turn enables the rootkit to intercept hardware calls made by the original OS.
 - **Hardware/firmware rootkits** - A firmware rootkit uses device or platform firmware to create a persistent malware image in hardware. The rootkit hides in firmware, because the latter is not usually inspected for code integrity.
- **Scareware.** Scareware is a generic term for malware that uses social engineering to frighten and manipulate a user, inducing him into thinking their system is vulnerable or has been attacked. However, in reality no danger has actually been detected: it is a scam. The attack succeeds when the user purchases unwanted - and potentially dangerous - software in an attempt to eliminate the "threat". Generally, the suggested software is additional malware or allegedly protective software with no value whatsoever [10].

Both Rogue Security Software and Ransomware can be considered as scareware, together with other scam software.

Some versions of scareware act as a sort of shadow version of ransomware; they claim to have taken control of the victim's system and demand a ransom. However they are actually just using tricks - such as browser redirect loops - to fool the victim into thinking they have done more damage than they really have [7].

- **Spyware.**

Spyware, another name for *privacy-invasive software*, is a type of malicious software that uses functions in the infected host's operating system with the aim of spying on the user activity. Specifically it can collect various types of personal information about users, such as Internet browsing habits, credit card details and passwords, without their knowledge. The information gathered is then sent back to the responsible cybercriminal(s). The presence of spyware is typically hidden from the user, and can be difficult to detect.

However, the functions of spyware often go far beyond simple activity monitoring and information gathering. In fact, they may also interfere with the user's control of the computer in other ways, such as installing additional software and redirecting web browser activity. Spyware is known to change computer settings, often resulting in slow connection speeds, different home pages, and/or loss of Internet connection or functionality of other programs. They spread by attaching themselves to legitimate software, Trojan horses, or even by exploiting known software vulnerabilities [8].

Law enforcement, government agencies and information security organizations often use spyware to monitor communications in a sensitive environment or during an investigation. Spyware is however also available to private consumers, allowing them to spy on their employees, spouse and children [13].

Other cyber-threats

Other cyber threats which are not strictly malware are, for example:

- **Software Bug.** A software bug is an error, or flaw, in a computer program code or system that causes it to produce an incorrect or unexpected result, or to behave in unintended ways. Usually, most of these defects arise from human errors made in the program's source code. Some bugs may be caused by compilers or operating systems component used by the program.

Minor bugs only slightly affect the behaviour of a program. Therefore it can be a long time before they are discovered. On the other hand, more significant bugs can cause crashes or freezes. It is safe to say that almost all software has bugs and most bugs go unnoticed or have a slight impact on the user.

However, other bugs qualify as security bugs. These are the most serious type of bugs since they can allow attackers to bypass access controls such as user authentication, override access privileges, or steal data.

The frequency of bugs can be reduced through developer training, quality control, and code analysis tools [14].

- **Malvertising.** Malvertising is the use of legitimate ads or ad networks to covertly deliver malware to unsuspecting users' computers.

For example, a cybercriminal might pay to place an ad on a legitimate website. When a user clicks on the ad, code in the ad either redirects them to a malicious website or installs malware on their computer.

In some cases, the malware embedded in an ad might execute automatically without any action from the user, a technique referred to as a "drive-by-download".

- **Phishing.** Phishing is a type of social engineering attack commonly used to perform cyber attacks. Particularly in a phishing attack, the attacker attempts, through email messages, to trick users into divulging passwords (or anyway personal and financial information), downloading a malicious attachment or visiting a website that installs malware on their systems.

Some phishing emails are highly sophisticated and can deceive even experienced users, especially if the attacker has successfully compromised a known contact's email account and uses it to spread phishing attacks or malware such as worms. Others are less sophisticated and simply spam as many emails as possible with messages such as "Check your bank account details" [15].

There are different types of Phishing. Here are mentioned some of them: [9]

- *Deceptive Phishing* - The most common type. It uses an email headline with a sense of urgency from a known contact. This attack blends legitimate links with malicious code, modifies brand logos, and evades detection with minimal content.
- *Spear Phishing* - Spear phishing targets specific users or organizations by researching the victim to maximise trick potential. For example the attacker may explore social media, record out-of-office notifications, compromise API tokens etc. in order to better fool the target user.
- *Whaling* - Whaling is similar to spear phishing, but even more targeted. In fact, it targets chief officers of organizations using various social engineering tricks such as impersonating employees or co-workers and using phone calls - to name a few - to give a sense of legitimacy to malicious emails.
- *Vishing* - Vishing targets phone users. It uses the Voice over Internet Protocol (VoIP), technical jargon, and ID spoofing to trick a caller into revealing sensitive information.
- *Smishing* - Smishing also targets phone users. It uses, however malicious text messages (SMS).
- *Pharming* - Pharming leverages cache poisoning against the DNS with the objective of redirecting users to fake websites.

- **Spam.**

In cybersecurity, unsolicited emails are generally referred to as *spam*. Typically, spam includes emails carrying unsolicited advertisements, fraud attempts, links to malicious websites or malicious attachments. Most spam emails contain one or more of the following: [9]

- Poor spelling and grammar
- Unusual sender address
- Unrealistic claims
- Suspicious links

Spam might be one of the most universally understood forms of malicious attacks. As billions of users enable email for their everyday lives, it makes sense that malicious actors try to sneak into their inbox. Some of the most common types of spam emails include fake responses, PayPal, returned mail, and social media. All of which are disguised as legitimate but contain malware.

General considerations on malware types

Malware samples are usually categorised both by a means of infection and a behavioural category: for instance, WannaCry is a ransomware worm.

Moreover, a particular piece of malware may have various forms with different attack vectors: e.g., the banking malware called *Emotet* has been spotted in the wild as both a trojan and a worm [7].

Finally, many instances of malware fit into multiple categories: for example Stuxnet is both a worm, a virus and a rootkit.

Furthermore, in recent years, targeted attacks on mobile devices have also become increasingly popular. In fact, among the huge amount of available apps, an increasing quantity is not desirable; the problem is even worse when considering third-party app stores. Even when app store providers impose filters and manual checks to prevent the malicious apps from being available, some inevitably slip through. These mobile malware threats are as various as those targeting desktops and include Trojans, Ransomware, Advertising click fraud and more. They are mostly distributed through phishing and malicious downloads and are a particular problem for jail-broken phones, which tend to lack the default protections that were part of those devices' original operating systems.

2.1.3 Malware History

Malware history began in the 1960s. Then, hackers used to design computer viruses mainly for fun, as an exciting prank/experiment; their creations would generally display harmless messages and then spread to other computers [16]. There are numerous examples of malware created at that time within a laboratory setting: for example the *Darwing game* in 1962, *Creeper* in 1971, *Rabbit Virus* in 1974 and *Pervading Animal* in 1975.

In particular, the malware called *Creeper* was designed to infect mainframes on ARPANET. The program did not alter the machines' functions, nor it stole or deleted data. It only displayed the message "I'm the creeper: Catch me if you can" while illegitimately spreading from one mainframe to another. This malware was later upgraded with the ability to self-replicate and became the first known computer worm [17].

In the early 1980s, the concept of malware caught on in the technology industry, and numerous examples of viruses and worms appeared both on Apple and IBM personal computers. With the introduction of the World Wide Web and the commercial internet in the 1990s it eventually became widely popularized, so much that Yisreal Radai coined the term **malware** in 1990.

The previously mentioned 1960s and 1970s malware were all kept within a laboratory environment and never managed to escape to the wild. *Elk Cloner* (1981) was the first known virus to have been able to escape its creation environment. Then, following the success of that prank gone wild, the first Microsoft PC virus, called *Brain*, was created in 1986. Again, like *Elk Cloner*, Brain was mostly annoying rather than harmful, but it was also the first known virus capable of concealing its presence on the disk thus evading detection. In 1988 the first worm, called *Morris* worm, an experimental, self-propagating, self-replicating program was released on the internet [6]. In 1988 made its appearance also the first example of intentionally harmful virus, the *Vienna* virus, which encrypted data and destroyed files. This led to the creation of the first antivirus tool ever [16].

In the following decades malware has evolved both regarding its complexity and malware sample numbers.

The growth in malware complexity can be divided 5 different malware generations [6]:

- *First generation:* (DOS Viruses) malware mainly replicate with the assistance of human activity
- *Second generation:* malware self-replicate without help and share the functionality characteristics of the first generation. They propagate through files and media.
- *Third generation:* malware utilise the capabilities of the internet in their propagation vectors leading to big impact viruses.
- *Fourth generation:* malware are more organization-specific and use multiple vectors to attack mainly anti-virus software or systems due to the commercialisation of malware.
- *Fifth generation:* malware is used in cyberwarfare and the now popular malware as-a-service makes its appearance.

Each jump in generation is linked to an increase in malware complexity and more propagation vectors being available. Newer generations of malware always re-utilise older techniques while introducing newer ones. Finally newer generations are more and more evasive due to the commercial value in having access to exploited systems.

Here is an overview of the most famous malware or malware-related events in recent history:

- **Melissa** (1999) - Considered to be one of the first cases of social engineering in history, the Melissa mass-mailing macro virus infected thousands of computers worldwide by the end of 1999. The virus was spread via e-mail, using a malicious Word attachment named "list.doc" and as subject: "Important message from", followed by the victim's username. Once opened, the attachment would execute a macro that mass-mailed the virus to the first 50 people in the user's contact list and disabled multiple security features on Microsoft Word and Microsoft Outlook.
- **ILOVEYOU** (2000) - *ILOVEYOU*, sometimes referred to as *Love Bug* or *Love Letter for you*, was a computer worm that infected over ten million Windows personal computers in 2000. It spread as an email message with the subject line "ILOVEYOU" and the attachment "LOVE-LETTER-FOR-YOU.txt.vbs". Opening the attachment would activate the Visual Basic script which caused damages on the local machine, overwriting random files. Finally, the worm sent a copy of itself to all addresses in the Windows Address Book used by Microsoft Outlook.
- **SQL Slammer** (2003) - This malware exploited a buffer overflow bug in Microsoft's SQL Server and Desktop Engine database products, causing a denial of service on some Internet hosts and dramatically slowing general Internet traffic. It spread rapidly, infecting most of its 75,000 victims within ten minutes.
- **MyDoom** (2004) - Also known as *W32.MyDoom@mm*, *Novarg*, *Mimail.R* and *Shimgapi*, it was a computer worm which affected Microsoft Windows systems. It became known mostly because it tried hitting major technology companies, such as Google and Microsoft. It spread by email using attention-grabbing subjects, such as "?Error?", "?Test?" and "?Mail Delivery System?". It became the fastest-spreading e-mail worm ever, exceeding previous records set by the Sobig worm and ILOVEYOU, a record which as of 2021 has yet to be surpassed.
- **Conficker** (2008) - Also known as *Downup*, *Downadup* and *Kido*, this computer worm targeted the Microsoft Windows operating system. It used flaws in Windows OS software and dictionary attacks on administrator passwords to propagate while forming a botnet. The Conficker worm infected over 15 million Windows systems including government, business and home computers in more than 190 countries. This made it the largest known computer worm infection since the 2003 *Welchia*.
- **Zeus** (2007-2009) - Also known as *ZeusS*, or *Zbot*, *Zeus* was a Trojan horse malware that ran on Microsoft Windows. It was able to carry out many malicious and criminal tasks. However, it was most often used to steal banking information through the use of man-in-the-browser keystroke logging and form grabbing. It was also often used to install the *CryptoLocker* ransomware. Zeus spread mainly through drive-by downloads and phishing mails. First identified in mid 2007, it became more widespread in 2009.
- **Stuxnet Worm** (2010) - *Stuxnet* was an extremely sophisticated worm that infected computers worldwide. It was allegedly developed by US and Israeli intelligence to hinder the Iranian nuclear program. It was introduced into the target environment (Iran's nuclear power plant) via a flash drive. Stuxnet's escape from the target environment, which was air-gapped, was not expected. Once in the wild, Stuxnet spread aggressively but mostly harmed the target Iranian nuclear facility, where it damaged uranium-enrichment centrifuges, causing little damage outside of its intended target environment [11].
- **CryptoLocker** (2013) - It is considered to be one of the first widespread ransomware attacks. It targeted computers running Microsoft Windows and it propagated via infected

email attachments. When activated, it encrypted certain types of files stored on local and mounted network drives using RSA public-key cryptography. The private key was stored only on the malware's control servers. A message was then displayed offering to decrypt the data if a payment (through either bitcoin or other means) was made before a certain deadline. After such deadline the private key was deleted. However, paying the ransom did not always lead to the files being decrypted.

Its code now keeps getting repurposed in similar malware projects.

- **Mirai** (2016) - *Mirai* is an infamous malware which compromised vulnerable IoT (Internet of Things) devices - such as IP cameras and home routers - turning them into remotely controlled bots.

The Mirai botnet, one of the biggest (and worse) botnets in existence, was first found in August 2016 and has been used in some of the largest and most disruptive distributed denial of service (DDoS) attacks, including an attack on computer security journalist Brian Krebs' web site. The source code for Mirai has been published on Hack Forums as open-source, as a result its techniques have been adapted in many other malware projects.

- **Petya and NotPetya** (2016-2017) - These malware attacks spread globally, however their damages particularly targeted Ukraine, where the national bank was hit. The Petya ransomware family caused an estimated \$10 billion in damages worldwide [16]. *Petya* targeted Microsoft Windows-based systems, infecting the Master Boot Record (MBR) to execute a payload that encrypted the hard drive's file system table preventing Windows from booting. It subsequently demanded a ransom in Bitcoin to the user in order to regain access to the system.

Many variants of Petya were created in the subsequent months. In mid 2017, a new variant of Petya was used for a global cyber-attack, again primarily targeting Ukraine. The new variant spread using the EternalBlue exploit, the same one used by the WannaCry ransomware. This new version was called *NotPetya* to distinguish it from the 2016 variants.

- **WannaCry** (2017) - WannaCry is considered to be one of largest ransomware attacks in history. It targeted computers running Microsoft Windows by encrypting data and demanding ransom payments in Bitcoin. It mainly propagated through EternalBlue, an exploit developed by the U.S. National Security Agency (NSA) for older Windows systems. This exploit was stolen from NSA and leaked approximately one year before the attack. While Microsoft had released security patches against this exploit, some organizations had not applied them, or were using older Windows systems when the attack occurred.

The attack was stopped within a few days of its discovery thanks to emergency patches released by Microsoft and the discovery of a kill switch. Before being stopped, WannaCry was spread infecting systems at a terrifying rate of 10,000 PCs per hour [16]. In the end, the attack was estimated to have affected more than 200,000 computers across 150 countries, with total damages ranging from hundreds of millions to billions of dollars.

- **Emotet** (2018) - This malware, also known as *Heodo*, was first detected in 2014 as a banking trojan aimed at stealing banking credentials from infected hosts. Throughout 2016 and 2017, its creators updated and reconfigured it to work primarily as a "loader" - a type of malware that gains access to a system, and then allows its operators to download and execute additional payloads. These payloads can be any type of executable code, from Emotet's own modules to malware developed by other cybercriminals.

This malware usually makes its way on target systems via a macro virus attached to an email. The infected email appears to be a legitimate reply to an earlier message sent by the victim. When on the system, it is particularly difficult to combat because it evades signature-based detection, is persistent, and includes advanced spreader modules used to propagate effectively. Emotet authors have often used the malware to create a botnet of infected computers to which they sell access in Malware-as-a-Service to other cybercriminals, such as the Ryuk gang.

In 2020, Emotet spread again globally, infecting its victims with TrickBot and Qbot, which are used to steal banking credentials and spread inside networks. In January 2021, Europol

and Eurojust coordinated international actions allowed investigators to take control of and disrupt the Emotet infrastructure.

- **COVID-19 related attacks** (2020) - In 2020, many cybercriminals shamelessly took advantage of the people's fear of coronavirus during the COVID-19 pandemic through COVID-19 related phishing scams. Using fake communications, for example spoofing the World Health Organization, attackers deployed malware and got access to targets' sensitive information among other nefarious actions [16].

Another COVID-19 attack was that of a malicious Android app called *CovidLock*, which claimed to be a real-time coronavirus outbreak tracker but instead was a ransomware that attempted to trick the user into providing administrative access on their device and then locked it requesting a ransom.

2.2 Obfuscation and detection evasion

From the creation of the first malware in 1970 [18], there has been a strong competition between attackers and defenders. To defend from malware attacks, anti-malware groups have been developing increasingly complex (and clever) new techniques. On the other hand, malware developers have conceived and adopted new tactics/methods to avoid the malware detectors.

The first type of anti-malware tools were mostly based on the assumption that malware structures do not change appreciably during time. In fact, initially, the malware machine code was completely unprotected. This allowed analysts to exploit opcode sequences to recognise specific malware families. Recently, however, a big advancement led to the so-called "second generation" malware [19] which, to evade such opcode signatures, employs several obfuscation techniques and can create variants of itself. This posed a challenge to anti-malware developers.

Obfuscation is a technique that generally makes programs harder to understand [20], both for humans and automatic tools. To do so, it converts a software to a new, structurally different, version while retaining the same functionality as the original. Originally, this technology was conceived for legitimate purposes to protect the intellectual property of software developers; however it has been widely exploited by malware authors to evade detection [21]. Particularly, in order to elude anti-malware scanners, malware can, using obfuscation techniques, evolve their body into new generations [22], which eventually can be even harder to disassemble and analyse.

The first time a malware has been recognised to exhibit detection avoidance behaviour was in 1986 with the *Brain* virus [23]. In fact, such malware managed to conceal the infected disk section whenever the user attempted to read it, forcing the computer to display clean data instead of the infected part. From that moment on, the ever increasing popularity of detection evasion techniques among malware writers has shown that malware survival has become the number one priority: the longer the malware remains undetected, the more harm it can do and the more profitable it is to its writer [6].

2.2.1 Reverse-Engineering

2.2.2 Anti-reversing

2.2.3 Anti-disassembly

Anti-disassembly techniques use specially crafted code and/or data in a program to cause disassembly analysis tools to generate an incorrect program listing [24]. The attackers' usage of these techniques thus implies a time-consuming analysis for malware analysts, ultimately preventing the retrieval of the source code in a reasonable time.

Any executable code can be reverse engineered, but by armouring their code with anti-disassembly and anti-debugging techniques, attackers increase the skill level required by analysts. Furthermore, anti-disassembly techniques may also inhibit various automated analysis tools and

heuristic-based engines which take advantage of disassembly analysis to identify or classify malware.

These techniques exploit the inherent weaknesses present in disassembler algorithms. Moreover, disassemblers, in order to work properly, make certain assumptions on the code being analysed. However, when these assumptions are not met, there is an opportunity for malware authors to deceive the analyst.

For example, while disassembling a program, sequences of executable code can have multiple disassembly representations, some of which may be invalid and obscure the real purpose of the program. Thus, the malware authors, in order to add anti-disassembly functionality to their creations, can produce sequences of code that deceive the disassembler into outputting a list of instructions that differs from those that would be executed [24].

There are two types of disassembler algorithms: linear and flow-oriented (recursive). The linear one is easier to implement, but it is also more simplistic and error-prone.

Linear Disassemblers

The *linear* disassembly strategy is based upon the basic assumption that the program's instructions are organized one after the other, linearly. In fact, this type of disassemblers iterates over a block of code, disassembling one instruction at a time, sequentially, without deviating. More specifically, the tool uses the size of the currently disassembled instruction to figure out what bytes to disassemble next, without accounting for control-flow instructions [24].

Linear disassemblers are easy to implement and work reasonably well when working with small sections of code. They introduce, however, occasional errors even with non-malicious binaries. The main drawback of this technique is that it blindly disassembles code until the end of the section, assuming the data is nothing but instructions packed together, without being able to distinguish between code, data and pointers.

In a PE-formatted executable file, for example, the executable code is typically contained inside a single ".text" section. However, for almost all binaries, this code section contains also data, such as pointer values. These pointers will be blindly disassembled and interpreted by the linear disassembler as instructions.

Malware authors can exploit this weakness of linear-disassembly algorithms implanting data bytes that form the opcodes of multi-byte instructions in the code section.

Flow-Oriented Disassemblers

The *flow-oriented* (or *recursive*) disassembly strategy is more advanced than the previous one and is, in fact, the one used by most commercial disassemblers like *IDA Pro* [24].

Differently from the linear strategy, the flow oriented one examines each instruction, builds a list of locations to disassemble (the ones reached by code) and keeps track of the code flow.

This implies that, if disassembling a code section we find a JMP instruction, this type of disassembler will not blindly parse the bytes immediately following the JMP instruction's ones, but it will disassemble the bytes at the jump destination address.

This behaviour is more resilient and generally provides better results, but also implies a greater complexity.

In fact, while a linear disassembler has no choices to make about which instructions to disassemble at any given time, flow-oriented disassemblers have to make choices and assumptions, in particular when dealing with conditional branches and call instructions. Particularly, in the case of conditional branches, the disassembler needs to follow both the false branch (most flow-oriented disassemblers will process the false branch of any conditional jump first) and the true one. In typical compiler-generated code there would be no difference in output if the disassembler processes first one branch or the other. However, in handwritten assembly code and anti-disassembly code, taking first one branch or the other can often produce different disassembly for the same block of code, leading to problems in analysis.

Anti-Disassembly Techniques

Jump Instructions with the Same Target One of the most used anti-disassembly techniques consists of two consecutive conditional *jump* instructions both pointing to the same target [24].

Here is an example:

Listing 2.1. Jump Instructions with the Same Target

```
74 03      jz      loc
75 01      jnz     loc

loc:
```

In this case, the conditional jump '**jz loc**' is immediately followed by a jump to the same target but with opposite condition: '**jnz loc**'. This implies that the location **loc** will always be jumped to. Consequently, the combination of **jz** with **jnz** acts, in this case, like an unconditional **jmp** instruction. A disassembler, however, since it disassembles just one instruction at a time, won't recognize this combination as being an unconditional branch. During the disassembly process, in fact, if a **jnz** instruction is encountered, the disassembler will take the false branch of the instruction and will continue disassembling, even though this branch will never be executed in practice.

Jump Instructions with a Constant Condition Another common anti-disassembly technique is composed of a single conditional *jump* instruction with an always true (or false) condition [24].

Example:

Listing 2.2. Jump Instructions with a Constant Condition example

```
33 C0      xor      eax, eax
74 01      jz      loc

loc:
```

The first instruction in the example code, **xor eax, eax**, sets the **EAX** register to zero and, consequently, it sets the zero flag. The next instruction, **jz** (jump if zero flag is set), appears to be a conditional jump but in reality is not conditional at all. In fact the zero flag will always be set at this point in the program execution. The disassembler, however, will process the false branch first, even if in reality it would never trigger.

Impossible Disassembly The simple anti-disassembly techniques mentioned above are frequently coupled with the use of a, so called, *rogue byte*. A *rogue byte* is a data byte strategically placed after a conditional *jump* instruction in order to trick the disassembler. The byte inserted usually is the opcode for a multi-byte instruction, therefore disassembling it prevents the real following instruction from being properly disassembled. This byte is called *rogue byte* because it is not part of the program logic flow and it is inserted in the code with the only purpose of fooling the disassembler [24].

In all these cases, however, a reverse engineer is able to properly disassemble the code with the use of interactive disassemblers like IDA Pro, ignoring the *rogue bytes*.

However, there are some conditions in which no traditional assembly listing can accurately represent the instructions that are executed. Exploiting these conditions we obtain what are called *impossible disassembly* techniques. The code produced using these techniques can however be disassembled, but only using a vastly different representation of the code than what is provided by currently available disassemblers.

The core idea behind these techniques is to make the *rogue byte* part of a legitimate instruction that is executed at runtime. This way the *rogue byte* becomes not ignorable during disassembly.

In this scenario any given byte may be a part of multiple instructions that are executed. This is done using *jump* instructions. The processor, while running the code, will interpret and execute the bytes following the logical flow of the program, so there is no limitation on the number of instructions the same byte can be part of; a disassembler, however, has such limitations since it will usually represent a single byte as being part of a single instruction.

Example:

Listing 2.3. Impossible Disassembly example

```
EB
    JMP -1
FF
    INC EAX
C0
48    DEC EAX
```

In this simple example the first instruction is a 2-byte *jmp -1* instruction (**EB FF**). Its target is its own second byte. At run time this causes no errors because the **FF** byte is the first byte of the next instruction *inc eax* (**FF C0**).

However, when disassembling, if the disassembler interprets the **FF** byte as part of the *jmp* instruction, it won't be able to interpret it also as the beginning of the *inc eax* instruction. While the **FF** byte is in reality part of both instructions that actually execute, the disassembler is not able to recognise this.

The 4-byte example code increments the **EAX** register, and then decrements it, therefore it is essentially a complex **NOP** sequence. Being a simple, and small, sequence it could be inserted at any location in a program code in order to fool disassemblers. However this sequence it is also easily recognisable by reverse engineers and substituted with **NOP** instructions using IDA Pro or other instruments and/or scripts. Another alternative is to interpret this sequence as data bytes forcing the disassembler to skip it.

However this was only a simple example sequence. More complex and ingenious sequences can be made to fool disassemblers while being harder to detect.

Obscuring Flow Control

Control-flow analysis (CFA) is a static-code-analysis technique for determining the control flow of a program. Modern disassemblers like IDA Pro are able to correlate function calls and extract high-level information about the program knowing how functions are related to each other [24].

Control-flow analysis can however be easily defeated by malware authors.

The Function Pointer Problem Function pointers are a common programming idiom present in programming languages such as **C**, while being extensively used in the background in object oriented languages like **C++** and **Java** [24].

As opposed to referencing a data value, a function pointer points to executable code within memory. Dereferencing the function pointer yields the referenced function, which can be invoked and passed arguments to as in a normal function call. Since, doing so, the function is being invoked indirectly through a variable instead of directly through a fixed identifier or address, such invocation is also known as an "indirect" call. In assembly code this corresponds to a *call* instruction with a function pointer as argument.

Function pointers, however, greatly reduce the information that can be automatically extracted by the disassembler about the program control flow. Moreover, if function pointers are used in specially crafted, or non-standard code, the resulting code can be difficult to reverse-engineer without the use of dynamic analysis techniques.

As a result, function pointers, in combination with other anti-disassembly techniques, can greatly increase the complexity and difficulty of reverse-engineering.

Return Pointer Abuse Among the instructions capable of transferring control within a program we already mentioned the *call* and *jmp* instructions, however there are more [24].

The counterpart to the *call* instruction is *retn*. When a call instruction is reached during program execution, a return pointer is pushed on the stack, before jumping to the call instruction target. This return pointer in the stack will point to the address of the instruction immediately following the end of the *call* instruction itself. Therefore a *call* instruction can be seen as the combination of a *jmp* and *push*; a *retn* instruction, on the other hand, is the combination of *pop* and *jmp*.

The *retn* instruction pops the last value pushed to the stack and jumps to it; it is therefore typically used to return from a function call, but it could also be used for other purposes. When used for such other reasons the disassembler is generally fooled, because it still will interpret it as a return from a function call. Therefore it won't show any code cross-reference to the target being jumped to. As added benefit the disassembler will also prematurely terminate the function being analysed.

Misusing Structured Exception Handlers Another powerful anti-disassembly technique exploits the Structured Exception Handling (**SEH**) mechanism. Performing program flow control using this mechanism is able to fool both disassemblers and debuggers [24].

SEH provides programs a way to handle error conditions intelligently. *C++* and other programming languages heavily rely on exception handling (and therefore on **SEH**) when compiled for x86 systems.

Exceptions can be triggered for numerous reasons: for example when dividing by zero or accessing an invalid memory region. Moreover, software exception can also be raised by the code itself by calling the *RaiseException* function.

When an exception is raised it makes its way through the **SEH** chain, which is a list of functions specifically designed to handle exception, until it is caught by one exception handler in the chain. Each function in the list can either handle the exception (a.k.a. *catch* it) or pass it to the next handler in the list. *Unhandled exceptions* are the ones that make their way to the last handler. The last exception handler is the code responsible for triggering the 'unhandled exception' message to the user.

Exception handling is used in almost all programs and exceptions happen regularly in most processes (and are handled silently). A malicious actor could, however, exploit this mechanism to achieve covert flow control by adding his own specially crafted handler on top of the **SEH** chain.

This can be done at runtime simply pushing some specific values on the stack, effectively adding a new entry in the Exception handling chain. This procedure, however, is subject to the constraints imposed by the Software Data Execution Prevention (**Software DEP**), which is a security feature that prevents the addition of third-party exception handlers at runtime. However various workarounds to this protection can be used in the case of handwritten assembly code.

2.2.4 Anti-debugging

Another popular anti-analysis technique, besides anti-disassembly, is *anti-debugging*. Malware authors use anti-debugging techniques to recognise when their malicious program is under the control of a debugger or to interfere with the debugger behaviour. This is done in an attempt to slow down the malware analysts who use debuggers to understand how the malware operates. A malware using these techniques usually alter its normal control flow paths or causes crashes if it detects it is running in a debugger, thus interfering with analysis [24].

Windows Debugger Detection

In Windows OS various techniques can be used to detect if a process is being run in a debugger: from exploiting the Windows API itself, to manually checking memory structures for debugging artefacts [24].

Using the Windows API One of the most obvious, and simple, ways to know if a debugger is attached to a process is by using Windows API functions. Inside the Windows API there are, in fact, functions that were specifically designed to detect debuggers; moreover some functions that were originally created with other purposes can also be used for debugger detection [24].

Malware analysts can counter this technique by manually modifying the malware code during execution modifying the resulting flag after the call to make sure the desired path is taken, or by straight up removing/skipping the function call.

Here are some examples of common Windows API functions used for *anti-debugging*:

- **IsDebuggerPresent** This is the simplest API function that can be used for debugger detection. It determines whether the **current** process is being debugged by a user-mode debugger. It does so by getting the value of the field *IsDebugged* from the Process Environment Block (**PEB**) structure. In particular this functions returns zero if the process is not running within a debugger context and a non-zero value otherwise.
- **CheckRemoteDebuggerPresent** This API function is similar to the previously described one (*IsDebuggerPresent*). This function checks for a 'remote' debugger on the specified process. The term 'remote' in the name *CheckRemoteDebuggerPresent* does not imply that the debugger necessarily resides on a different machine; instead, it indicates that the debugger resides in a separate and parallel process. This function takes a process handle as argument, and will check if that process has a debugger attached. It can however be used also to check the current process by passing its handle.
- **NtQueryInformationProcess** This function can retrieve different kinds of information from a process. The first argument for this function is the process handle, the second one is the *ProcessInformationClass* parameter which specifies the information you want to get. When using the value *ProcessDebugPort* for this parameter, for example, the function will return a zero if the process is not currently being debugged; a non-zero value representing the debugger port number will instead be returned otherwise.
- **OutputDebugString** This function, originally designed to just send a string to a debugger for display, can be used to detect the presence of a debugger. In fact, in there is no debugger attached, the function will internally set the last-error code. In a few lines of code it is thus possible to know if a debugger is present or not:

Listing 2.4. *OutputDebugString* debugger detection

```
DWORD errorValue = 12345;
// set custom last error code
SetLastError(errorValue);

// try outputting string on debugger;
// if no debugger is present, it will set
// the last-error code to a new value
OutputDebugString("Test_for_Debugger");

if(GetLastError() == errorValue){
    // a debugger is present
    ExitProcess();
}
else{
    // no debugger was detected
    RunMaliciousPayload();
}
```

Manually Checking Structures Malware authors usually don't exploit the Windows API functions for detecting the presence of a debugger, but they prefer checking the PEB structure

(and others) by themselves. One of the reasons why they usually don't like using Windows API functions is that API calls can be easily hooked by a rootkit to return false information, thus thwarting this technique [24].

- **Checking the BeingDebugged Flag** The Windows PEB structure contains all user-mode parameters associated with a process, including the process's environment data such as environment variables, addresses in memory and debugger status, among other things.

Malware can 'manually' check the *BeingDebugged* flag within the PEB structure to understand if a debugger is attached its process. More precisely if this flag is zero it means that no debugger is attached.

Example of code listing performing 'manual' *BeingDebugged* check:

Listing 2.5. BeingDebugged manual check

```
mov eax, dword ptr fs:[30h]      ; get PEB address
mov ebx, byte ptr [eax+2]       ; get BeingDebugged flag value
test ebx, ebx                   ; test if the value is 0
jz NoDebuggerDetected          ; if 0, no debugger was detected
```

Malware analysts can counter this technique detecting this code sequence in the code and then wither manually changing the *BeingDebugged* flag to zero, or forcing the jump to be taken (or not) by manually modifying the zero flag before the jump instruction.

- **Checking the ProcessHeap Flag** The *ProcessHeap*, which is an undocumented location within a reserved array inside the PEB structure, contains the location of the first heap of a process allocated by the loader. This heap can be used for debugger detection since it contains some information telling if it was created within a debugger or not. In particular malware usually check the values of the fields called *ForceFlags* and *Flags*.

To overcome this technique, malware analysts can change the *ProcessHeap* flags manually or use a hide-debug plug-in for their debugger.

- **Checking NTGlobalFlag** Processes started within a debugger run slightly differently than others, therefore they create memory heaps differently. The information needed to determine how to create heap structures is stored at an undocumented location in the PEB. Practically, a value of *0x70* at this location indicates that the process is running within a debugger.

Again, in order to counter this technique, malware analysts can change the flags manually or use a hide-debug plug-in for their debugger.

Checking for System Residue Debugging tools typically leave traces of their presence on the system. Malicious programs can therefore be designed to search for these traces in the system in order to determine when it is being analysed. For example malware can search for references to debuggers in the registry keys [24].

Moreover, malware can also be designed to search the system for files and directories commonly related to debuggers, such as debugger program executables.

Furthermore, malware can also detect debugger residues in live memory, by viewing the current process listing or, more commonly, by performing a *FindWindow* in search for a debugger.

Identifying Debugger Behaviour

Debuggers are very useful to malware analysts because they can be used to set breakpoints in the code or even to single-step through a process running code to ease the reverse-engineering process. These operations, however, modify the process code and are therefore easily detectable [24].

INT Scanning A common anti-debugging technique used by malware authors consists in making the process scan its own code in search for an **INT 3** (opcode *0xCC*). **INT 3** is, in fact, a software interrupt used by debuggers: when setting a breakpoint the debugger replaces the target instruction in the running program with the opcode *0xCC* (INT 3) which causes the process to call the debug exception handler [24].

Malware analysts can counter this technique exploiting hardware breakpoints instead of software ones.

Performing Code Checksums Another anti-debugger technique consists in calculating the checksum of a section of the process' own code. This has the same net effects as scanning the code for software interrupts. However, instead of explicitly searching for a specific opcode (*0xCC*) in the process code, this check performs a cyclic redundancy check (CRC) or a MD5 checksum of the malware code [24].

Again this technique can be countered by using hardware breakpoints instead of software ones, or by modifying the program's control flow at runtime with a debugger.

Timing Checks One of the most widespread techniques for debugger detection is to perform *timing checks*. Processes, in fact, tend to run substantially slower when executed within a debugger context. Moreover analysts usually run programs in single steps in order to better understand the code behaviour, this in turn greatly increases execution time [24].

Using timing checks it is possible to detect a debugger in different ways:

1. Recording 2 timestamps before and after the execution of some operations and then comparing them. If the lag is greater than a specified threshold then a debugger is probably being used.
2. Recording 2 timestamps before and after raising an exception. If the current process is being debugged then the exception will be handled by the debugger itself more slowly than normal. Moreover, by default, debuggers ask for human intervention when an exception occurs, thus causing huge delays.

- **Using the *rdtsc* Instruction** The most common timing check method uses the *rdtsc* instruction. This instruction returns the number of ticks since the last system reboot. Malware authors thus use it as described above: *rdtsc* is called twice, once before and once after some other operations, and then the difference between the results is calculated. If too much time has elapsed between the two calls it means that a debugger is probably being used.

- **Using *QueryPerformanceCounter* and *GetTickCount*** These are two Windows API functions that can be used similarly to *rdtsc* for debugger detection.

More precisely *QueryPerformanceCounter* can be called to query a high-resolution counter available to processors which stores counts of activities performed by the processor.

The function *GetTickCount*, on the other hand, returns the number of milliseconds that have elapsed since the last reboot, very similarly to what the *rdtsc* instruction does.

Both of those functions, when used as described above, allow the malware to detect the presence of a debugger.

Anti-debugging though the use of timing checks can be discovered by malware analysts during debugging or static analysis by identifying specific sequences of instructions. Moreover, these checks usually detect debuggers only when the analyst is single-stepping through the code or setting a breakpoint between the two time related instruction calls. This implies that, to counter this technique, malware analysts could avoid setting breakpoints and single-stepping in those regions of code, or modify the result of the timestamps comparison as needed.

2.2.5 Anti-virtual machine

Malware analysts often use virtual machines (VMs) or other isolated environments like sandboxes, to analyse malware samples. With the purpose of evading analysis and bypassing security systems malware authors often design their code to detect isolated environments. The techniques used with such purpose are called *Anti-virtual machine* techniques (Anti-VM). Once a virtual machine is detected the evasion mechanism may alter the malware's behaviour, or it may even prevent the malicious code from running altogether [24].

VMware Artefacts

Virtual machines are designed to emulate real hardware functionality. To achieve that, however, some artefacts inevitably remain on the system, which can reveal that a virtual machine is indeed being used. These kind of artefacts can be specific files, processes, registry keys, services, network device adapters etc. [24].

Here are some examples of anti-virtual machine techniques applied to detect VMware virtualization software:

- **Checking for Processes Indicating a VM.** When a VMware virtual machine is running and VMware tools is installed, three VMware-related processes can be found in the system process listing: *VMwareService.exe*, *VMwareTray.exe* and *VMwareUser.exe*. A malicious software can therefore easily detect if VMware is being run searching through the process listing for the *VMware* string.
- **Checking for Existence of Files Indicating a VM.** The VMware default installation path usually also contains artefacts. Searching for the string *VMware* in such location may reveal the use of a virtualized environment.
- **Checking for Registry Keys.** VMware Tools may leave some artefacts also in the registry. More specifically the presence of specific registry entries may reveal the use of VMware.
- **Checking for Known Mac Addresses.** In order to connect a virtual machine to a network it needs to have its own virtual network interface card (NIC). This implies that VMware needs to create a MAC address for the virtual machine, to associate to its NIC. However, depending on VMware configuration, this may lead to the network adapter being able to identify VMware usage. VMware utilises, in fact, addresses with a specific starting sequence which depends on its current version. Therefore a malicious software just needs to check the system MAC address against common VMware values.

In order to counter anti-virtual machine techniques, malware analysts need to apply a two step process: identify the check for VMware artefacts and then 'manually' patch it. For example, depending on the anti-VM technique used, they may patch the malware code while debugging to artificially make all checks pass, or modify the name of VMware processes in order to make them undetectable by the malicious software.

Vulnerable Instructions

The virtual machine monitor program, which monitors the virtual machine execution, has some security weaknesses that may allow malware to detect its usage. In particular, in order to avoid performance issues deriving from fully emulating all instructions, VMware allows certain instructions to execute without being properly virtualized. This in turn means that certain instruction sequences may return different results when running within a VMware virtualized environment than they do on native hardware. This discrepancy can be used by malware authors to detect VMware usage [24].

However, those instructions previously mentioned are not typically used within a malicious program unless it is specifically performing VMware detection, because they are useless if executed in user mode. Therefore avoiding this type of anti-VM technique can be as easy as patching the malicious code to prevent it calling these instructions.

2.2.6 Packers and unpacking

2.2.7 Code Obfuscation

2.2.8 Polymorphic and Metamorphic Malware

Chapter 3

Proposed Tool

Description of the proposed tool..

Chapter 4

Results

Results analysis..

Chapter 5

Conclusions

Qui si inseriscono brevi conclusioni sul lavoro svolto, senza ripetere inutilmente il sommario.

Si possono evidenziare i punti di forza e quelli di debolezza, nonché i possibili sviluppi futuri o attività da svolgere per migliorare i risultati.

Bibliography

- [1] R. Sharp, “An introduction to malware.” <https://orbit.dtu.dk/en/publications/an-introduction-to-malware>, 2017
- [2] R. Moir, “Defining malware: Faq.” [https://docs.microsoft.com/en-us/previous-versions/tn-archive/dd632948\(v=technet.10\)?redirectedfrom=MSDN](https://docs.microsoft.com/en-us/previous-versions/tn-archive/dd632948(v=technet.10)?redirectedfrom=MSDN), 2009, Accessed: 2021-03-15
- [3] NIST, “malware.” <https://csrc.nist.gov/glossary/term/malware>, Accessed: 2021-03-15
- [4] C. Crane, “What is malware? 10 types of malware and how they work.” <https://www.thesslstore.com/blog/what-is-malware-types-of-malware-how-they-work/>, 2020, Accessed: 2021-03-15
- [5] Symantec, “Difference between viruses, worms, and trojans.” <https://knowledge.broadcom.com/external/article?legacyId=TECH98539>, 2019, Accessed: 2021-03-15
- [6] A. P. Namanya, A. Cullen, I. U. Awan, and J. P. Disso, “The world of malware: An overview”, 2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud), 2018, pp. 420–427, DOI 10.1109/FiCloud.2018.00067
- [7] J. Fruhlinger, “Malware explained: How to prevent, detect and recover from it.” <https://www.csoononline.com/article/3295877/what-is-malware-viruses-worms-trojans-and-beyond.html>, 2019, Accessed: 2021-03-15
- [8] P. Mullins, “Malware and its types.” http://www.idc-online.com/technical_references/pdfs/information_technology/Malware%20and%20its%20types.pdf, Accessed: 2021-03-15
- [9] S. Ingalls, “Types of malware and best malware protection practices.” <https://www.esecurityplanet.com/threats/malware-types/>, 2021, Accessed: 2021-03-15
- [10] MyraSecurity, “What is malware?.” <https://www.myrasecurity.com/en/what-is-malware/>, Accessed: 2021-03-15
- [11] K. Baker, “The 11 most common types of malware.” <https://www.crowdstrike.com/cybersecurity-101/malware/types-of-malware/>, 2021, Accessed: 2021-03-15
- [12] T. Femister, “Encryption happens last: The ransomware revolution.” <https://www.forbes.com/sites/forbestechcouncil/2020/08/18/encryption-happens-last-the-ransomware-revolution/?sh=4c267d34414b>, 2020, Accessed: 2021-03-15
- [13] McAfee, “What is malware?.” <https://www.mcafee.com/en-us/antivirus/malware.html>, Accessed: 2021-03-15
- [14] N. DuPaul, “Common malware types: Cybersecurity 101.” <https://www.veracode.com/blog/2012/10/common-malware-types-cybersecurity-101>, 2012, Accessed: 2021-03-15
- [15] Comtact, “What are the different types of malware?.” <https://comtact.co.uk/blog/what-are-the-different-types-of-malware>, 2019, Accessed: 2021-03-15
- [16] J. Regan, “What is malware? how malware works and how to prevent it.” <https://www.avg.com/en/signal/what-is-malware>, 2019, Accessed: 2021-03-15
- [17] B. Lutkevich, “malware.” <https://searchsecurity.techtarget.com/definition/malware>, Accessed: 2021-03-15
- [18] P. Szor, “The art of computer virus and defence”, Symantec press, 1st ed., 2005, ISBN: 978-0-321-30454-4

- [19] A. Sharma and S. K. Sahay, “Evolution and detection of polymorphic and metamorphic malwares: A survey”, *International Journal of Computer Applications*, vol. 90, Mar 2014, pp. 7–11, DOI [10.5120/15544-4098](https://doi.org/10.5120/15544-4098)
- [20] A. Balakrishnan and C. Schulze, “Code obfuscation literature survey.” <http://pages.cs.wisc.edu/~arinib/writeup.pdf>, 2005
- [21] E. Konstantinou, “Metamorphic virus: Analysis and detection.” <https://www.ma.rhul.ac.uk/static/techrep/2008/RHUL-MA-2008-02.pdf>, 2008, Technical Report of University of London
- [22] I. You and K. Yim, “Malware obfuscation techniques: A brief survey”, 2010 International Conference on Broadband, Wireless Computing, Communication and Applications, 2010, pp. 297–300, DOI [10.1109/BWCCA.2010.85](https://doi.org/10.1109/BWCCA.2010.85)
- [23] E. Skoudis and L. Zeltser, “Malware: Fighting malicious code”, Prentice Hall Professional, 2004, ISBN: 978-0-131-01405-3
- [24] M. Sikorski and A. Honig, “Practical malware analysis: The hands-on guide to dissecting malicious software”, No Starch Press, 1st ed., 2012, ISBN: 978-1-59327-290-6