portada

[**1. Abstract**](#_hqgoqgf28oqs) **2**

[**2. Preamble**](#_e1v6gdv2xx73) **2**

[**3. What is Malware**](#_yhkinkg2if6m) **2**

[3.1. Brief history. Analysis of most important malware through modern time](#_1qzv9o8447gu) 2

[3.2. Programming languages](#_lmo7ojungb5j) 8

[3.3. Propagation](#_7ez8ddrli62j) 11

[3.4. Infection requirements](#_36iy9z3o9a83) 13

[**4. Types of malware according to their purposes**](#_vq7sk4pne882) **13**

[4.1. Viruses](#_3klm1fwerhu6) 13

[4.2. Worms](#_wos494xgktyl) 14

[4.3. Trojans](#_uwagb32l93bf) 15

[4.4. Ransomware](#_y42ctzvc460h) 16

[4.5. Fileless Malware](#_rp2tynd4j9vd) 17

[4.6. Wiperware](#_yfu87argyjx7) 18

[4.7. Grayware](#_opq62vrddk10) 18

[4.7.1. Spyware](#_mf610kpzaaal) 19

[4.7.2. Adware](#_4u25ctto6l5h) 20

[4.8. Rootkits](#_f8te939vndw8) 20

[4.9. Cryptojacking](#_3bk0p8593skh) 20

[4.10. Rogue security software](#_edr578aqalao) 21

[4.11. RAM scraper](#_ago9wdushho4) 21

[4.12. Crimeware](#_npti8ylwm5wr) 21

[4.13. Bots](#_kfquxrpv098e) 22

[4.14. Less common malware: Unix, MacOS and Android](#_a47kxq7gisfc) 22

[**5. Prevention**](#_ds3yctcy5fij) **24**

[**6. Detection**](#_o6fkq4g4ls6h) **24**

[**7. Removal**](#_8pdrn77kkkjw) **24**

[**8. Experimental part:**](#_ne2yjq978sac) **24**

# Abstract

# Preamble

# What is Malware

As the very term suggests, malwares, short for malicious softwares, are intrusive software intended to harm computers, networks and other associated devices by stealing information, corrupting files and threatening users’ privacy. They spread mostly through networks and portable devices (less common at present), and transfer to devices without the knowledge of its owner.

Malware has always been a threat to computer security, and with the increase in the use of the Internet, its impact is becoming more severe because most of them are not easily detected.

Malware is often erroneously referred to as computer virus, so at this point of the paper it is important to clarify that these two concepts are not the same. The first one refers to any program or code that is created with the intent to do harm to a computer, network or server, while the second one is a type of malware.

## Brief history. Examples of important malware

The history of malware in chronological order, from early concepts to most destructive malware of modern years.

Between 1971 and early 2000, malware was mostly relegated to mischief and attempts by virus authors to see if something they had created would work. During the late 1980s, malwares were simple boot sectors and file infectors spread via floppy disk. In the 1990s, macro viruses, which spread via email attachment and exploited Microsoft Office products proliferated due to the increased use of email. Some examples of specific malware are shown below:

* **“The creeper”, first Proof of Concept (POC)[[1]](#footnote-0):** Developed by the engineer of BBN[[2]](#footnote-1) Bob Thomas in 1971, “The Creeper” is considered the first computer virus in history. It exhibited the behavior of a worm (see point 2.8), spreading via network protocol NCP through ARPANET[[3]](#footnote-2) computers and leaving a simple message.
* **“Elk cloner”, first Mac virus:** “Elk cloner” was developed in 1982 by a teenager and targeted Apple computers. This boot sector virus[[4]](#footnote-3) propagated when an infected disk was run, and on the fiftieth boot, it displayed a poem to the user.
* **“Brain”, the first PC virus:** In 1986 the brothers Amjad and Basit Farooq Alvi created a boot sector virus that showed a warning to individuals using a pirated copy of their medical software. The Internet was not of public access at the time, so the virus, “Brain”, spread via the copying of floppy disks until it became a global phenomenon.

It was not a destructive software, but it prevented the computer from booting and displayed a notification with the contact information of the brothers. They wanted affected individuals to call them to discuss how to obtain their software legally.

* **The Morris worm:** The Morris worm was created in 1988 by Robert Morris as a proof of concept, so it was not malicious. It exploited vulnerabilities in various programs and services and checked to see if an existing infection was present, behaviors that can still be observed on modern malware. Moreover, Morris programmed it for persistence[[5]](#footnote-4). However, as there was no way to stop the self-replication process, the worm caused high loads on devices, rendering them inoperable, and denials of service (DOS[[6]](#footnote-5)) on networks.
* **“AIDS”, world’s first ransomware:** In 1989, the “AIDS” Trojan became the world's first ransomware. At the time the human AIDS virus[[7]](#footnote-6) was a relevant topic worldwide, so the “AIDS” trojan was sent via mail (physical mail, not email) to AIDS researchers via 20 000 infected floppy disks.On the ninetieth reboot of the system, “AIDS” encrypted the disk and displayed a demand of $189 for a yearly lease.

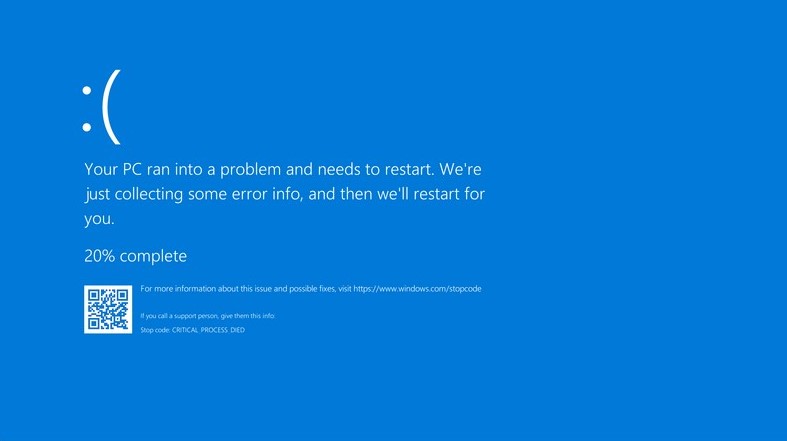
The creator, Dr. Joseph Popp, claimed that he created the ransomware to donate the funds he collected to AIDS research.

* **“Michelangelo”:** In 1992, “Michelangelo” was the next virus to make a significant impact, it received a lot of attention from the mainstream media. It was a 0-day virus[[8]](#footnote-7) that rewrote some sectors of the hard drive with void data, erasing their original content.
* **The first phishing attacks:** Between 1994 and 1995, internet access was quite expensive, so new programs appeared. One of the most famous, “AOHell”, contained a random account creator that used randomly created credit card accounts to open an account for free for a month on AOL chatrooms.

At the same time, fake automated AOL instant message bots sent indiscriminate IMs to targets asking them to verify their account credentials. Then the users of “AOHell” sold them or used them to access the internet freely.

* Other malwares that had an impact were Jerusalem, CIH and Melissa.

From the year 2000, the threat landscape has evolved from mischief to include profitable cybercrime and nation-state attacks. An increase in the use of exploit kits, programs used by cybercriminals to exploit system vulnerabilities, led to another increase in malware delivered online. Since 2007, when some ways to mass compromise websites eased distribution capabilities of malware, the number of attacks has grown exponentially.Socially engineered worms[[9]](#footnote-8) and spam proxies[[10]](#footnote-9) began to appear, as well as phishing[[11]](#footnote-10) and other credit card scams.

* **First botnet[[12]](#footnote-11):** By 2000, the first botnet “EarthLink Spam” appeared. Its function was to send massive amounts of spam, about 1.25 million messages in total.
* **I LOVE YOU:** In 2000, the “I LOVE YOU” worm spread around the world at record speed, using a novel approach: it was sent an an attachment in an email, and when the victim opened it, the worm looked for Microsoft Outlook address book and send out emails impersonating the victim and replicating itself as an attachment. This simple method is still used in modern malware.
* **Blaster**: In 2003, “Blaster” targeted a remote procedure call (RPC[[13]](#footnote-12)) vulnerability in 2003 operating systems to propagate worldwide. The worm’s goal was to prevent the computer from accessing updates, but luckily the author made a mis-take directing “Blaster'' to the wrong domain. Due to a buffer overflow[[14]](#footnote-13) bug, it also caused a denial of service (in the form of a BSOD[[15]](#footnote-14)) that could not be recovered through reboot. This was the first global Denial of Service attack.
* **Mytob/Zotob:** In 2005, “MyDoom”’s variants “Mytob” and “Zotob” were incredibly prolific and disruptive, taking down the operations of more than a hundred organizations. They combined the functionality of a worm, a botnet and accesses through backdoors[[16]](#footnote-15). Mytob was one of the first malwares to specifically work against antiviruses.
* **CoolWebSearch and BayRob:** “CoolWebSearch” was the first malware to hijack search results from Google (browser hijacker, see point 2.13) and overlay them with those from the perpetrators. Similarly, “BayRob” captured real results from eBay searches and injected fake ones for people to purchase.
* **Stuxnet:** In 2010, “Stuxnet” was the first malware in history to target industrial critical infrastructures. In this case, it caused nuclear centrifuges to overspin, resulting in a meltdown.
* **Flame:** “Flame” was considered the most advanced malware ever found at the time of discovery, that is, 2012. It could spread through LANs[[17]](#footnote-16) like a worm, record and capture screenshots and audio, record Skype conversations and send and receive certain files through bluetooth.
* **Reveton:** “Reveton” (2011/2012) was the archetype of modern ransomware, it settled the look and feel that still remains to this day. It had a professional appearance, it displayed different lock screen templates based on the location of the victim. The lock screens simulated local law enforcement organizations.
* **CryptoLocker:** In 2013, “CryptoLocker” was the first ransomware to demand payment via Bitcoin.
* **Lazarus team:** The Lazarus team of hackers in 2013 attacked in South Korea the broadcaster SBS and banking institutions in the DarkSeoul attack. Their malware, “Jokra”, overwrote devices’ Master Boot Record[[18]](#footnote-17), making the machines unable to start. The same group, in 2014, leaked confidential information about Sony Corporation.
* **Browser Locker and fake technical support scams:** Although technically not malware, these attacks mimic ransomware so the victim either pays the ransom or calls a fake support number created by the attackers. These attacks injected malicious JavaScript code on vulnerable legitimate websites. The script would then render the browser inoperable, frequently displaying warnings and demands in full-screen mode.
* **TeslaCrypt:** In 2015, “TeslaCrypt” became famous because, in the beginning, it infected game files, blocking maps and user profiles. However, evolved versions of TeslaCrypt were able to encrypt other files, such as PDF and Word.
* **The first IoT[[19]](#footnote-18) botnet:** “Mirai” was, in 2016, the first botnet to target IOT devices, primarily routers. It was mainly a DDOS[[20]](#footnote-19) botnet, and managed to take down a massive segment of the internet, causing disruptions all over the world.
* **WannaCry, Petya/NotPetya:** “WannaCry”, “Petya” and their versions were some of the most devastating ransomware attacks in history in terms of loss volume. “Petya”(2016) spread via emails with malicious attachments. It infected the Master Boot Record (see 15) of machines, which blocked the entire operating system. “Petya” and its most important versions like “NotPetya” caused more than USD 10 billion in financial losses. “WannaCry”(2017) spread via email scams, or phishing, successfully infecting more than 200 thousand users and organizations and causing an approximate loss of $4 billion.
* **SamSam:** In 2018 “SamSam” gained prominence after infecting the city of Atlanta, the Colorado Department of Transportation and the Port of San Diego, in the U.S., abruptly stopping services. At the same time, two iranian hackers were accused of using “SamSam” to attack more than 200 companies and public institutions, which caused an estimated loss of $30 million.
* **Botnets to mine crypto:** XMRig is a miner application written to mine for Monero cryptocurrency and is not malicious. However, in 2018 cybercriminals began installing XMRig on compromised machines and collecting the data for their own benefit. A variant of this attack targeted Android devices mainly via malicious APKs[[21]](#footnote-20).
* **GandCrab: Ransomware as a service:** With the organization GandCrab, a new, more voluminous and violent wave of attacks occurred in 2019: they rented a ransomware with the same name for cybercriminals to use. They enlarged and perfected the business model known as Ransomware-as-a-Service (RaaS), allowing the authors to work assiduously to update the malware so that it could evade antivirus and other security defenses while getting others to perform the actual breaches. Thanks to this, “GandCrab” far eclipsed the success of competing ransomware affiliate programs.
* **CovidLock:** In 2020, during the lockdown due to Covid-19, “CovidLock” ransomware infected targets through malicious files claiming to offer information about the disease. This threat affected only Android devices.

## Programming languages

Programming is an essential skill for people working in technological environments. A programming language consists of a set of instructions we can use to communicate with a computer and make it perform a specific task. It is mainly used to develop any kind of applications and websites. We can classify them according to their abstraction:

* **Low-level languages:** their instructions have direct control over the hardware and therefore their function depends on the physical structure of the computers that support them. They provide little or no abstraction from a computer's instruction set architecture, so the processor can run low-level programs directly without the need of a compiler or interpreter. There are three types of languages, in order from lower to higher complicity:
  + **Binary code** is the most basic and forms part of all information systems. It’s very easy to use, as it only needs two absolute values, 0 and 1.
  + **Machine language** is also formed of 0 and 1, and its function is to communicate the instructions directly to the computer.
  + **Assembly languages** are a bit more abstract, that is, human-understandable, so they need a tool such as an assembler to translate them into machine language.

Low-level languages have a simple and very adaptable code, with direct instructions that are easy to convert, and can run at high speed.

Programs written in low-level languages tend to be relatively non-portable, because they are specific for the hardware.

* **High-level languages:** they have strong abstraction from the details of the computer, so the programs are more or less independent of a particular type of computer. They may use natural language elements, making the process of developing a program simpler and more understandable. When code is written in a high-level language an interpreter or compiler has to translate it into low-level code so the computer can understand. Some examples of high- level languages are C, C# and C++, Python, Java and JavaScript, HTML, Fortran, Pascal, Go, Ruby, Swift, Perl, PowerShell, Golang and Cobol. We can also classify high-level languages according to the way they are executed:
  + **Compiled**: they are converted directly into machine code that the processor can execute. They tend to be faster and more efficient to execute, and give more control over hardware aspects like CPU usage or memory management. Compiled languages’ code needs to be built before execution, and rebuilt every time a change is made.
  + **Interpreted**: Interpreters, programs that translate the code to machine language, run through a program line by line and execute each command.

Nowadays, most programming languages have both compiled and interpreted implementations – the language itself is not necessarily compiled or interpreted.

When it comes to malware, most of it is written in either C or C++ or some other compiled language, although many times it depends on what platform the attacker is willing to target.

**Assembly** is the lowest-level language humans can read, and so, it can control more details on how the malware/program works.

**C,** being a very powerful and general-purpose programming language, is the main candidate when it comes to programming malware. Some reasons for this is that it has many windows-based libraries that efficiently control the computer's functionality, and its memory management is highly efficient. Moreover, it can be used to write all sorts of malware for different computing environments.

**Python** is also a popular language among malware programmers due to its easy syntax and the fact that it has several libraries that make designing security and offensive tools very easy, such as Nmap, regex, boto3, socket, and scapy.

**Java, JavaScript and Actionscript** are often used in web-based attacks, to create computer exploits[[22]](#footnote-21) (usually to perform SQL injection attacks[[23]](#footnote-22) and/or cross-site scripting[[24]](#footnote-23)).

When they target mobile devices, hackers also lean towards Java, because many Android mobile apps are written in it.

When it comes to more specific Windows malware, **Visual Basic for Applications (VBA)**, the language used to create Microsoft macros, can create malware that will run whenever someone downloads a Microsoft attachment and enables macros

According to a recent report published by BlackBerry’s Research & Intelligence division[[25]](#footnote-24), malware authors are turning to new programming languages - **Go, DLang, Nim and Rust** - as effective methods to hide their malicious code from security tools. To put it simple, they wrap a malware on a dropper[[26]](#footnote-25) written in one of these less common languages so the security software is unable to understand it. Then when the dropper is executed, the real malware infects the computer.

## Propagation

Just as in the past malware spread mainly via floppy disks, modern malware distribution is based on the internet. Still, it needs an attack vector, most of the times an interaction with the user, to establish its presence on an endpoint.

By far the most common method to spread malware is through **phishing emails.** These use **social engineering,** which is the oldest and still most effective method, to trick users into clicking on links or downloading files that contain malicious code.

**Malvertising** is also a common source of malware infections. This method consists of criminally controlled advertisements within Internet connected programs (usually web browsers) that look like legitimate online advertising, which are used to install malware and other threats with little to no user interaction required. There are two methods to infect a computer.

The first one presents a provocative bait that scares or tempts the user into clicking the ad, such as a warning or a prize alert.

The second one, known as **drive-by download**, is even more dangerous: the infected ad uses an invisible web page element so just loading the web page that hosts it redirects the user to an exploit page.

In both methods then a tiny piece of code is downloaded. It sends a sort of an acknowledgement message to a Command and Control[[27]](#footnote-26) server, which scans the computer and sends adequate malware to install according to the operating system and hardware specifications.

**Social network spam** is a relatively new approach for cybercriminals. This is, photos or videos shared on a social media site that, when clicked, take the user to a fake page that may mimic a real one such as YouTube or Instagram then requests the user to log in with their account or to download and install some kind of plugin (video player, image viewer…) that hides some malicious code.

**Remote Desktop Protocol** (RDP) is a connection protocol that enables a user to connect to another computer over a network connection. If not configured, it remains open or locked with a default password. A cybercriminal may use this backdoor to install malware.

While most malware nowadays has to be directly downloaded and run to infect a machine, there is still another class: the **network-based**. This kind of malware attempts to exploit any network-based vulnerabilities on a machine. If successful, it uses the machine’s network connection to probe for other machines on the network, so that it can spread and replicate itself freely inside the LAN.

And if that was not enough, if the propagation failed, it could also corrupt other devices’ disks thanks to mapped drives[[28]](#footnote-27), or even removable media like USB.

## Infection requirements

Each malware is made to run on certain versions of a specific operating system. For example, a Windows 10 malware will not be able to infect a Linux, macOS, or even Windows XP machine. Just as a Windows 10 machine will not run a Linux or macOS program.

This is because every operating system has its own API (Application Programming Interface), a huge set of functions to let programs perform common tasks, such as connect to a network or play a sound. Thanks to this standardization, the operating system is able to provide a simple way for all the programs to perform these actions, so they do not have to provide the specific instructions themselves, resulting in light-weight software.

In conclusion, for a malware to be able to run on a defined operating system, it has to be written according to that system’s API. Else, it will become innocuous. Also, if the OS’s antivirus software detects the malware, it will delete it immediately so the infection will not happen either.

# Types of malware according to their purposes

Although most of the real world malwares are hybrid, that is, a mix of two or more of the following types, this classification according to their purpose is useful to establish all the primary functions a malware can accomplish.

These terms are not exclusive, meaning that, for example, a computer virus can be a trojan if the file containing it is delivered by a trojan.

## Viruses

A computer virus is a fragment of code embedded in a legitimate program that can replicate and spread to other programs after a person first runs it on their system. It can wreak serious damage on a machine by modifying or destroying essential files, which may cause system crashes, program malfunctions and data loss. Virus cannot be controlled remotely. The major categories for viruses are:

* **Boot sector Virus:** Infects the computer storage sector where startup files are found. It executes malicious code during startup time, before many security layers are launched, rendering the machine unusable.
* **Web Scripting Virus:** Older browsers may have vulnerabilities which allow an attacker to run code on the local device, altering the functioning and settings of the browser and ultimately spreading to other applications.
* **Browser hijacker:** It can change the settings on a browser, just as book- marks and search preferences. It is used to redirect the user to malicious sites
* **Resident Virus:** Stays dormant until a specific payload is delivered.
* **Direct Action Virus:** Is attached to an executable file, and will only spread when the file is executed.
* **File Infector Virus:** Malicious code is injected into critical files that run the operating system, so when the system boots, the virus is activated.
* **Polymorphic Virus:** It changes its signature (pattern of bytes in its code that identify it) every time it is installed to avoid detection.
* **Encrypted Virus:** In order to avoid detection, the virus remains encrypted until it is executed.

Some examples of computer viruses are Elk cloner, Brain, BHP, MacMag, 1260, Shamoon, Michelangelo, Melissa and Jerusalem.

## Worms

A worm is a malicious program that can propagate or replicate itself from one computer to another without human interaction after having accessed a machine, usually, through a network or a LAN. Unlike viruses, worms do not modify any program files and can be controlled remotely. A worm can perform different actions, such as drop other, more harmful malware, delete certain files or steal data, open a backdoor, deplete hard drive space by writing null or random data or by creating too many copies of itself, or overload networks.

Many worms are designed only to spread, and do not modify the systems they pass through. However, even these can cause disruption by increasing network traffic and consuming bandwidth, and thus slowing down Internet connection for users.

Some famous computer worms are the Morris worm, Blaster, MyDoom, Code Red, Beagle, Conficker, ILOVEYOU, Ryuk, SQL Slammer and Stuxnet.

## Trojans

A trojan is any malware that seeks to mislead the user of its true intent. To put it simple, it disguises itself as a legitimate program, so the user downloads and executes it. It is generally spread by some form of social engineering, such as email attachments or fake advertising.

A trojan is a delivery strategy that cybercriminals use to distribute any kind of threats, from adware to ransomware, although many modern ones simply act as a backdoor for a cybercriminal to control the machine or capture sensitive information like usernames and passwords.

Some of the most common types of trojans are:

* **Downloader Trojan (dropper):** downloads and deploys other malicious code, such as rootkits, ransomware or keyloggers.
* **Backdoor Trojan:** creates remote access to the system, allowing cybercriminals to control the device, collect sensitive data or install other malware.
* **Spyware:** Spyware is sometimes considered a type of trojan because it is disguised as useful software. It observes user’s activities, collecting sensitive data like account credentials or banking details.
* **Rootkit Trojan:** aims to acquire root-level or administrative access to a machine.
* **Zombifying Trojan:** takes control of the computer to make it a slave in a network under a hacker’s control (botnet), and use that net for malicious purposes such as DDOS attacks.

ZeuS and Emotet are the most infamous trojans of the last few years.

Viruses, worms and trojans are frequently confused, so the table below may help clarify the differences and similarities among them:

|  | **Virus** | **Worm** | **Trojan** |
| --- | --- | --- | --- |
| **Consists of** | Piece of code embedded in a legitimate program. | Programs that propagate without human interaction. | Malware that disguises itself as legitimate software. |
| **Self-replication** | Yes (to other files on a system) | Yes (to other systems on a network) | No |
| **Can be remotely controlled** | No | Yes | Yes |
| **Spreading Rate** | Moderate | Fast | Rather slow |
| **Main purpose** | Alter the functioning of a system or the data on it. | Cause slowdown on network connections or system processes (eat system’s resources) | Deliver other malware, create backdoors or steal sensitive data. |
| **Execution** | When the executable file that contains it is run. | Via weaknesses on a system (when the infection is successful) | When it is run. |

## Ransomware

Ransomware is a type of malware that threatens to publish or block access to the victim’s personal data on a computer system unless a ransom is paid. There are two main types of ransomware, screen lockers and encryptors. The first ones are simpler, they block access to the system using a lock screen, claiming that the system has been encrypted when in fact, no file has been damaged. The second ones are more dangerous, because they do encrypt data on a machine, making it inaccessible without the description key.

The bulk of files is encrypted with symmetric encryption. This method uses a key which is a random sequence with a length of between 128 and 256 bits. Knowing that every bit may have two values, either 1 or 0, we can generate between and different keys, so it is nearly impossible to recover the encrypted files without the original key.

Then in more advanced ransomware, this symmetric key is asymmetrically encrypted with a public key, so even if the user finds it in the ransomware’s code, he will not be able to decrypt the files. The corresponding private key is owned by the attacker, so when the ransom is paid, he sends it to the victim to decrypt the symmetric key, and so the files.

Ransomware is the top variety of malicious software, found in 70% of cases where malware was identified[[29]](#footnote-28). It menaces all the statement of the CIA triad model[[30]](#footnote-29), Confidentiality, Integrity and Availability, which also makes it one of the most dangerous types.

One of the largest and most serious ransomware attacks took place in 2017 and was called **WannaCry.** Between 230,000 and 300,000 victims from roughly 150 countries were asked to pay a ransom in Bitcoin, resulting in a financial loss of more than $4 billion. Other ransomwares that had a significant impact were Locky, BadRabbit, Ryuk, CryptoLocker, Petya and NotPetya, GoldenEye, GandCrab and B0r0nt0k.

## Fileless Malware

Unlike traditional malware, which relies on a file being written to a disk, fileless malware is designed to work in volatile system areas such as the system registry, in-memory processes and service areas, so its existence on the system lasts only until the process is terminated or the system is rebooted. It can also hide its code inside existing benign files, trusted applications like PowerShell or Windows Script or invisible registry keys, which thanks to the trust model used by security applications to not monitor whitelisted programs, allows it to go unnoticed. It can also exist on a hard disk other than the one from the affected computer.

Fileless attacks are much more effective than regular attacks because they are often undetectable by most security solutions and forensic analysis.

In essence, fileless malware can do anything that “regular” malware can do, but for practical reasons, there is often a limitation in the amount of fileless malicious code that can remain in volatile memory. Due to that, for more complex programs like ransomware, the fileless malware might act as a dropper. However, some of its most common functions are the harvest of credentials with the purpose of privilege escalation, creation of backdoors, data (such as network credentials/configuration) exfiltration or to act as a dropper.

## Wiperware

Wiperware is maybe the most destructive form of malware. Its sole purpose is not to is not to steal money or sell information to cybercriminals but rather to cause damage by erasing the hard drive of the computer it infects, that is, deleting all data and programs.

The reason behind these attacks, as opposed to the ones that seek for money or attention, may be either political, as a protest, or simply cover the tracks of a separate data theft.

Some examples of this practice are Shamoon, Meteor, NotPetya, ZeroCleare, WhisperGate,

## Grayware

Grayware is not exactly a type of malware: it is not necessarily harmful but is often unpleasant or irritating, so if its effects are detrimental enough it is classified as one. It is also classified as a Potentially Unwanted Program[[31]](#footnote-30) that manage to get into a system through the download of other programs. Grayware is an intrusive application which causes disruptions to the machine, but it can even be entirely legal in certain scenarios.

The most common types of grayware are:

### Spyware

Spyware is one of the most common forms of malware or malicious behavior thanks to its apparently harmless and discrete functioning. Its purpose is to monitor and capture personal and sensitive data from the device and user, and send or sell it to third parties, such as advertisers or cybercriminals, without their consent. This information (internet usage, credit card, and bank account details, or user credentials) can then be used for advertising purposes (to show more relevant ads), which would be the least intrusive form, or sold to criminal groups as potential targets for zombifying trojans and other types of malware, or identity thefts. Websites may also engage in spyware behaviors like web tracking.

* **Adware**: it can be a type of spyware when it not only displays an awful amount of ads and malicious pop-ups, but also monitors user’s activity to show more relevant ones.
* **Infostealer**: collects specific information from devices, such as usernames and passwords, email addresses, browser history and webpages’ login information.
* **Keylogger**: generally a trojan, collects all of the information that the user types into their device’s keyboard and stores it in an encrypted file.
* **Tracking cookies:** dropped onto a device by a website and then used to follow the user’s online activity.

Spyware is not considered a malware when it is related to marketing and advertisements, despite acting without the user's consent and knowledge. In any case, it violates user privacy and may endanger device's security and integrity.

Pegasus, GhostRat, CoolWebSearch and Gator are some real life examples.

### Adware

Adware is unwanted software designed to display advertisements in the form of pop-ups, banners, and videos within an application, operating system or most commonly a web browser.

Adware has the purpose of generating revenue for its developer. It usually does not cause any damage on the system, it only represents a nuisance for the users.

However, as mentioned in the previous section, when adware also has the function of collecting data such as browser history or location and displays ads based on those, it can be considered a type of spyware.

Fireball, Filetour and Deskad are the most influential adware of the last few years.

## Rootkits

Rootkits are collections of malware, usually trojans, designed to enable access to a target device and control an area of the software that is not accessible for a normal user. In other words, they give to an attacker the highest privileges in a system (root access). They are able to conceal their presence while remaining active.

## Cryptojacking

Cryptojacking is a form of malware that hides on a computer and uses the machine’s computing power to generate cryptocurrencies, forms of digital money that exist only in the online world, primarily Monero or Zcash. It is designed to remain completely hidden from the user.

The primary impact of cryptojacking is performance-related, it does not threaten user’s data integrity nor expects the user to pay a ransom, instead it can lead to slowdowns and crashing due to the overexertion of computer resources. Nevertheless, it can increase costs for the individuals affected because as it uses high levels of computer power, it requires a lot of electricity.

## Rogue security software

Rogueware is a malware, usually a kind of trojan, that relies on social engineering to deceive users into believing their computer is afflicted with a virus. Sometimes its objective is only to scare the victim (In that case, it is also called scareware). However, most of the times, it attempts to induce the user to pay to download a phony anti-malware software or a malware eradication service, which installs other, more dangerous malware.

An early example that gained infamy was SpySheriff and its clones

## RAM scraper

A RAM scraper attack is the intrusion into a retail sales terminal’s RAM[[32]](#footnote-31) to collect consumers’ credit card sensitive information. RAM scraper malware scans the memory of digital devices to collect that information (credit card numbers and personal identification numbers, PIN) and sends it to the attacker for the purpose of exploitation.

## Crimeware

Crimeware is any computer program or malware designed for the express purpose of conducting criminal activities online. Crimeware programs are meant to automate the theft of information, allowing the thief to gain access to a person’s financial accounts online. This can be accomplished by redirecting the web browser to malicious sites that mimic legitimate ones, controlled by the thief, where the user is asked to enter their login credentials (phishing), or sites infected with network-based malware. This kind of malware can also enable remote access of applications, allowing criminals to break into their private networks, or steal passwords stored on system’s cache[[33]](#footnote-32). The most sophisticated ones will even install other malware such as keyloggers to gather even more data.

A crimeware attack that is targeted to a specific enterprise or organization is called a targeted threat.

## Bots

Bots are automated pieces of software that perform some kind of defined, repetitive action. A malicious bot, however, is the one that when installed in a computer, can be automated or commanded remotely by an attacker to gain full control over the device. As the purpose of the bot is to perform its function for as long as possible while avoiding detection, most of the times the user does not even realize that their computer is being manipulated. Because of this, a computer infected by a bot is referred to as a zombie.

A large group of zombie computers connected to each other is called a botnet. Cybercriminals use botnets to perform other attacks, such as Distributed Denials of Service (DDoS), an Account Takeover attacks[[34]](#footnote-33), spam phishing emails, steal data and more.

To summarize, a malicious bot converts a computer into a zombie device and connects it to a botnet that a cybercriminal can control at will and use to perform other types of attacks.

## Less common malware: Unix, MacOS and Android

As mentioned on point 1.4, the machines a certain malware can infect are defined by the specific API they are dedicated to. The large majority of examples listed so far are intended to affect different versions of Microsoft Windows OS, not because it is the only one vulnerable to these threats, but because it is by far the most common, hence the most profitable.

Some people do think that Linux/Unix OS distributions are immune to malware. The truth is that, although they are certainly very well protected against it, these systems are still vulnerable. But then why is there so few Linux/Unix malware? In the past, it has been suggested that Linux had so few malware because as its market is rather low, it is a less profitable target. But that is not exactly correct, because there are a lot of non-desktop devices, such as web servers or workstations, which are Unix-derived. Another strong reason is that, as Linux/Unix OS implements a multi-user environment where users are granted different grades of privileges, to gain control over a system or to cause any serious damage, the malware would have to gain root access, which is not as easy as it is in Windows. In the last years, Mirai, PNScan, Gafgyt and GonnaCry have been the most famous Linux threats.

MacOS is said to rarely suffer malware or virus attacks. This is mainly because system software updates to resolve vulnerabilities, just as utilities to find and remove malware, are released very frequently. Some examples of recent MacOS malware are XLoader, XcodeSpy, Silver Sparrow, GoSearch22,GenericSuspicious and FakeFileOpener.

Android malware is really no different than the different types of malware discussed previously which affect desktop or laptop computers. Android devices are considered the most susceptible to malware infection in part because the Google Play Store ecosystem, which should be legitimate and completely safe, has fewer security measures in place when developers are introducing their apps to the public, unconsciously allowing malware, mainly trojans, to infect off-guard users. And if that was not enough, users can download content directly from the internet onto their device, or via their computer, completely bypassing the PlayStore. Of course, traditional forms of malware propagation are applicable too. Android does not come with a preinstalled security software, so it is highly recommendable to install some.

# Prevention

# Detection and removal

# Experimental part:

As I stated in the introduction of this paper, the main objective is to demonstrate that we are closer to malware than we think. To do so, I will attempt to create one completely from scratch, with no more prior knowledge than Windows system administration and architecture, and basic Python programming.

Python is far from being the best choice when it comes to malware programming, but the fact that it is one of the easiest programming languages existing will support the point that programming malware is not difficult in essence.

To avoid repeating the word malware, I will first of all name this project as ‘mocaMW’.

## First design and features

The first step I made in this project was to establish what will be the purpose of ‘mocaMW’. From the types previously listed, the one that I found more threatening from the point of view of the user is ransomware, because it extorts them directly, it is one of the few that face the victim, it does not hide nor steal data or money silently.

And so, the first design of mocaMW was inspired in the ransomware model of file encryption, as shown in the diagram. The functions this version was meant to incorporate were one for encryption and one for decryption.

The main feature of ransomware is the encryption in itself, so I started to search information about it and how to implement it on Python.

Encryption is a form of cryptography that scrambles plain text into unintelligible cipher text using a sort of password called key. In general, in an encrypted communication, a host encrypts a message with a key and then sends it to a recipient, who decrypts it, either with the same key or the corresponding private one, so it can be read again. There are two types of encryption:

* **Symmetric encryption:** Only one key is used to both encrypt and decrypt data. To encrypt, an algorithm scrambles the data in a way that it cannot be understood, and to decrypt, the same algorithm reverses the scramble. The key can be a specific password or a random sequence generated by a RNG[[35]](#footnote-34). Its length may be 128, 192 or 256 bits.

Symmetric encryption is considered less secure compared to asymmetric encryption, although it is the older, faster and most efficient method for encrypting large chunks of data.

It can be easily carried through in Python thanks to the module ‘Fernet’ from the ‘cryptography’ library.

* **Asymmetric encryption:** It uses two separate but mathematically related keys; the public one, which encrypts data, and the private one, which decrypts it. Usually, the public key generated based on the private key. The public key does not need to be secret, because in any case, data can not be decrypted without the private key.

Asymmetric encryption is designed for securing sensible data and symmetric key exchanges in public channels. It uses large keys, usually more than 1024 bits long, being the most common length 2048 bits, which makes it more secure and strong, but also converts it in a high resource-consuming process and makes it take longer time, so it is only suitable to encrypt small chunks of data.

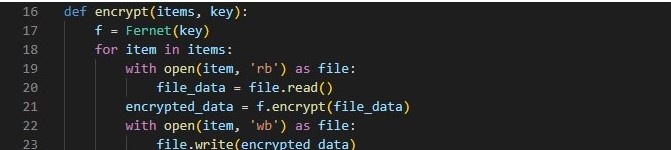
Python’s ‘cryptography’ library has also some modules and functions which can be used in asymmetric encryption, but none as specific as ‘Fernet’ for symmetric.

Now, taking into account the fact that the malware should be able to encrypt any kind of data, independently of its size or type, along with the ease to incorporate each method in Python, I opted for symmetric encryption using the ‘Fernet’ module.

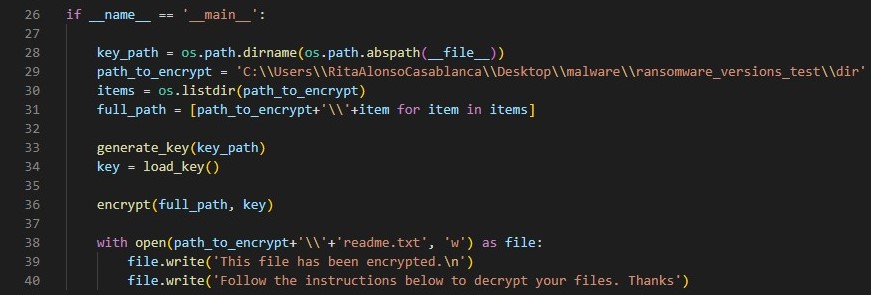
## Version 1 [Stable 1.2]

The functioning of the ‘Fernet’ module is very simple. It includes a classmethod generate\_key(), which generates the key, an encrypt function and a decrypt function. To maintain code modularity, I created two separated files, one to generate the key and encrypt, and the other to decrypt.

In this code, I imported the necessary libraries, ‘os-sys’ and ‘cryptography’ and then defined a function to generate a key. In line 5, the variable ‘key’ is defined as the result of the Fernet method generate\_key(). Then the content of this variable is written on a file so the decrypt script can also access it. Note that the function takes as a parameter the key\_path variable, which is used in the open method to indicate the full path where the is stored. (I will provide the key\_path later, when this function is called). I also created a function to read the key from the keyfile and store it into a variable so we can use it (not shown).



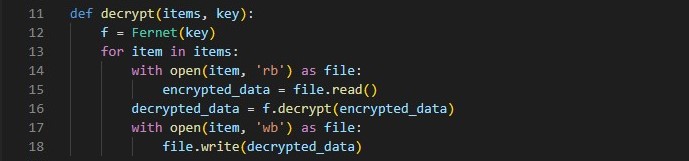
This is the function which encrypts the files. It takes two arguments (I will set them when it is called), ‘items’ represents a list with all the names of the files in the directory, that is, all the files that have to be encrypted, and the second one, ‘key’, represents the key that will be loaded into a variable. Basically, this means that for every element in the list ‘items’, it will read it and store its binary into a variable. Then, it will encrypt the variable and write the content back into the file also in binary form.

Finally, we need to execute all these functions, and I will do so with an if statement:

The condition in line 26 means that the code will be run when the file is run directly as main, that is, when it is not imported and run through another file. Then, the key\_path variable states the absolute path to the current folder, that is, the folder from which the program is running and where the key has to be generated. The path\_to\_encrypt variable stores the path of the folder we want to encrypt, the items variable, as I said before, returns a list with all the files in the folder, and finally the full\_path variable creates a list with the absolute paths to all the files in the directory to encrypt joining the path\_to\_encrypt with each of the names in items list.

Then, we generate and load the key and finally encrypt the files.

I added the code in lines 38-40 to create a text file in the encrypted folder and display a message.



For the decryption file, I imported the same libraries and wrote the same function to load the key. The decrypt function is practically the same as the encrypt one but backwards: for every path in the ‘items’ list, it opens and reads the file on that path as binary and stores it to a variable. Then it decrypts the variable, opens the file anew and writes the decrypted data.

I did a lot of function tests before achieving something close to what I wanted. The first full program I wrote seemed to work fine, but on the decryption file I came up against a bug that took me quite some time to solve[[36]](#footnote-35). It was a cryptography.fernet.InvalidToken error, which appears when the data to decrypt is not valid. The problem was that, as shown in image , a text file was generated on the encrypted folder, but contrary to image , the statement in line 25 was missing, so the readme.txt file would not be removed. The error raised because a file that has not been encrypted cannot be decrypted.

1. A POC is a demonstration of the viability or functioning of a certain idea or concept. [↑](#footnote-ref-0)
2. Rayteon BBN is an American research and development company which works on technological matters. [↑](#footnote-ref-1)
3. Before the internet existed, there was the Advanced Research Projects Agency Network (ARPANET), a project which in 1969 managed to connect remote computers. It also developed NCP, the first network transport layer to enable data to flow from one computer to another. [↑](#footnote-ref-2)
4. Boot sector virus is a malware that infects the computer storage sector where startup files are found. They execute malicious code during startup time,before many security layers are executed. [↑](#footnote-ref-3)
5. Persistence occurs when an attacker maintains long-term access to systems despite disruptions such as restarts or reboots. [↑](#footnote-ref-4)
6. In a Denial Of Service attack, the resources of a machine or network are made unavailable by flooding the targeted resource with superfluous requests in an attempt to overload systems [↑](#footnote-ref-5)
7. Known as VIH/SIDA [↑](#footnote-ref-6)
8. The code on a 0-day malware is executed only on a certain day, such as March 6th, in this case. [↑](#footnote-ref-7)
9. Worms that are spread disguised as a tantalizing video or image file, or as desirable software in order to trick the user. [↑](#footnote-ref-8)
10. A proxy server is an intermediary server separating end users from the websites they browse. [↑](#footnote-ref-9)
11. Phishing is a type of social engineering where an attacker sends a fraudulent message designed to trick a person into revealing sensitive information, like usernames or passwords. [↑](#footnote-ref-10)
12. A botnet is a group of compromised computers under the command and control of an operator. [↑](#footnote-ref-11)
13. Software communication protocol that one program can use to request a service from a program located in another computer on a network, without knowing that it is remote. [↑](#footnote-ref-12)
14. Occurs when the volume of data exceeds the storage capacity of the memory buffer (memory storage regions that temporarily hold data while it is being transferred from one location to another). As a result, the program overwrites adjacent memory locations. [↑](#footnote-ref-13)
15. Windows displays a Blue Screen of Death when the system runs into an error which may be solved after a system reboot, or may not be solvable [↑](#footnote-ref-14)
16. A way to access a computer system or encrypted data that bypasses the system's customary security mechanisms [↑](#footnote-ref-15)
17. Local Area Network is the network within a router, created by a switch. [↑](#footnote-ref-16)
18. The MBR is a boot sector category that provides information about the hard disk partitions and the OS so it can be loaded for the system boot. Without the MBR, the system is unable to start. [↑](#footnote-ref-17)
19. The Internet of Things is the network of physical objects that are embedded with sensors, software, and other technologies able to exchange data with other devices and systems over the internet. [↑](#footnote-ref-18)
20. In a Distributed Denial Of Service attack, unlike in a DOS (see 6), the resources of a machine or network are made unavailable by flooding the targeted resource with superfluous requests sent from a large number of hosts, commonly “zombie” hosts of a botnet. [↑](#footnote-ref-19)
21. Android Package Kit (.apk) is the file format that Android uses to distribute and install apps. [↑](#footnote-ref-20)
22. An exploit is a piece of software that takes advantage of a vulnerability in an operating system, application or any other software code to cause unintended behavior, which might include giving control of the computer to the attacker or allowing privilege escalation.. [↑](#footnote-ref-21)
23. Web security vulnerability that allows an attacker to interfere with the queries that an application makes to its database, that is, allowing him to view or modify data that he should not be able to retrieve. [↑](#footnote-ref-22)
24. XXS is a web security vulnerability that allows an attacker to compromise the interactions that users have with a vulnerable application. [↑](#footnote-ref-23)
25. The BlackBerry Research & Intelligence Team. (2021, July 26). *Old dogs new tricks: Attackers adopt exotic programming languages*. BlackBerry Blog. Retrieved December 7, 2022, from <https://blogs.blackberry.com/en/2021/07/old-dogs-new-tricks-attackers-adopt-exotic-programming-languages> [↑](#footnote-ref-24)
26. A dropper is a kind of trojan that has been designed to install other malware to a computer. The malware contained within the dropper can avoid detection by virus scanners. [↑](#footnote-ref-25)
27. A C2 server is the one that cybercriminals use to issue commands to control malware and bots of a botnet. [↑](#footnote-ref-26)
28. A mapped drive is a shortcut to a drive that is physically located on a different computer. They can be used to reach resources on different computers on a local network, as well as files on a website or server. [↑](#footnote-ref-27)
29. Verizon. (2022). *DBIR report 2022 - Results and analysis*. Verizon Business. Retrieved December 8, 2022, from <https://www.verizon.com/business/resources/reports/dbir/2022/results-and-analysis-intro/> [↑](#footnote-ref-28)
30. The CIA triad model is designed to guide policies for information security within an organization. [↑](#footnote-ref-29)
31. PUPs are programs that are regarded as undesirable despite being frequently downloaded. [↑](#footnote-ref-30)
32. Random Access Memory is a computer's short-term memory, where the data that the processor is currently using is stored. [↑](#footnote-ref-31)
33. Cache is a storage location used to temporarily store data used by servers, apps, and browsers to speed load times. Cache would allow a browser to load certain resources without downloading them from the server every time it accesses the page, so it would load faster. [↑](#footnote-ref-32)
34. Account takeover is a form of online identity theft in which a cybercriminal illegally gains unauthorized access to an account belonging to someone else. [↑](#footnote-ref-33)
35. Random Number Generator [↑](#footnote-ref-34)
36. For a detailed report refer to https://github.com/Moca15-ar/ransomware\_tdr/blob/main/v1.0.0/v1.1.0/ [↑](#footnote-ref-35)