

Basic precautions for malware analysis:

Before analyzing malware, one must understand that malware is often destructive. This means that when malware is being analyzed, there is a high chance of damaging the environment in which it is being analyzed. This damage can be permanent, and it might take more effort to get rid of this damage than the effort made to analyze the malware. Therefore, creating a lab setup that can withstand the destructive nature of malware is necessary.

Virtual Machines:



A lab setup for malware analysis requires the ability to save the state of a machine (snapshot) and revert to that state whenever required. The machine is thus prepared with all the required tools installed, and its state is saved. After analyzing the malware in that machine, it is restored to its clean state with all the tools installed. This activity ensures that each malware is analyzed in an otherwise clean environment, and after analysis, the machine can be reverted without any sustained damage.

Virtual Machines provide an ideal medium for malware analysis. Some famous software used for creating and using Virtual Machines includes [Oracle VirtualBox](#) and [VMWare Workstation](#). These applications can create snapshots and revert to them whenever required, making them

well-suited for our malware analysis pursuit. In short, the following steps portray the usage of Virtual Machines for malware analysis.

1. Created a fresh Virtual Machine with a new OS install
2. Set up the machine by installing all the required analysis tools in it
3. Take a snapshot of the machine
4. Copy/Download malware samples inside the VM and analyze it
5. Revert the machine to the snapshot after the analysis completes

Following these steps ensures that your VM is not contaminated with remnants of previous malware samples when analyzing new malware. It also ensures that you don't have to install your tools again and again for each analysis. Selecting the tools to install in your malware analysis VM can also be hectic. One can use one of the freely available malware analysis VMs with pre-installed tools to ease this task. Let's review some common malware analysis VMs most popular among security researchers.



FLARE VM:

The FLARE VM is a Windows-based VM well-suited for malware analysis created by Mandiant (Previously FireEye). It contains some of the community's favorite malware analysis tools. Furthermore, it is also customizable, i.e., you can install any of your own tools to the VM. FLARE VM is compatible with Windows 7 and Windows 10. For a list of tools already installed in the VM and installation steps, you can visit the [GitHub page](#) or the [Mandiant blog](#) for the VM. Since it is a Windows-based VM, it can perform dynamic analysis of Windows-based malware.

An instance of FLARE VM is attached to this room for performing practical tasks. Please click the Start Machine button on the top-right corner of this task to start the machine before proceeding to the next task. The attached VM has a directory named mal on the Desktop, which contains malware samples that we would be analyzing for this room.



REMnux:

REMnux stands for Reverse Engineering Malware Linux. It is a Linux-based malware analysis distribution created by Lenny Zeltser in 2010. Later on, more people joined the team to improve upon the distribution. Like the FLARE VM, it includes some of the most popular reverse engineering and malware analysis tools pre-installed. It helps the analysts save time that would otherwise be spent in identifying, searching for, and installing the required tools. Details like installation and documentation can be found on [GitHub](#) or the [website](#) for distribution. Being a Linux-based distribution, it cannot be used to perform dynamic analysis of Windows-based malware. REMnux was previously used in the Intro to Malware Analysis room and will also be used in the upcoming rooms.

Please click the Start Machine button on the top-right corner of this task to start the machine before proceeding to the next task. The machine will start in a split-screen view. If the VM is not visible, use the blue Show Split View button at the top-right of the page. The attached VM has a directory named mal on the Desktop, which contains malware samples that we would be analyzing for this room. Alternatively, you might use the following information to log into the machine:

Machine IP: MACHINE_IP

Username: Administrator

Password: letmein123!

In the [Intro to Malware Analysis](#) room, we identified that searching for strings is one of the first steps in malware analysis. A string search provides useful information to a malware analyst by identifying important pieces of strings present in a suspected malware sample. To learn a little more about strings, we can look at [this room](#) dedicated to strings.

How a string search works:

A string search looks at the binary data in a malware sample regardless of its file type and identifies sequences of ASCII or Unicode characters followed by a null character. Wherever it finds such a sequence, it reports that as a string. This might raise the question that not all sequences of binary data that looks like ASCII or Unicode characters will be actual strings, which is right. Many sequences of bytes can fulfill the criteria mentioned above but are not strings of useful value; rather, they might include memory addresses, assembly instructions, etc. Therefore, a string search leads to many False Positives (FPs). These FPs show up as garbage in the output of our string search and should be ignored. It is up to the analyst to identify the useful strings and ignore the rest.



What to look for?

Since an analyst has to identify actual strings of interest and differentiate them from the garbage, it is good to know what to look for when performing a string search. Although a lot of useful information can be unearthed in a string search, the following artifacts can be used as Indicators of Compromise (IOCs) and prove more useful.

- Windows Functions and APIs, like SetWindowsHook, CreateProcess, InternetOpen, etc. They provide information about the possible functionality of the malware
- IP Addresses, URLs, or Domains can provide information about possible C2 communication. The Wannacry malware's killswitch domain was found using a string search

- Miscellaneous strings such as Bitcoin addresses, text used for Message Boxes, etc. This information helps set the context for further malware analysis

Basic String Search:

In the [Intro to Malware Analysis](#) room, we learned about the `strings` utility, which is pre-installed in Linux machines and can be used for a basic string search. Similarly, the FLARE VM comes with a Windows utility, `strings.exe`, that performs the same task. This Windows `strings` utility is part of the Sysinternals suite, a set of tools published by Microsoft to analyze different aspects of a Windows machine. Details about the `strings` utility can be found in [Microsoft Documentation](#). The `strings` utility comes pre-installed in the FLARE VM attached to this room. The good thing about the command line `strings` utility is that it can dump strings to a file for further analysis. In the attached VM, executing the following command will perform a basic string search in a binary.

```
C:\Users\Administrator\Desktop>strings <path to binary>
```

Several other tools included in the FLARE VM can be used for string search. For example, CyberChef (Desktop>FLARE>Utilities>Cyberchef) has a recipe for basic string search as well. PEstudio (Desktop>FLARE>Utilities>pestudio) also provides a string search utility. PEstudio also provides some additional information about the strings, like, the encoding, size of the string, offset in the binary where the string was found, and a hint to guess what the string is related to. It also has a column for a blacklist, which matches the strings against some signatures.

pestudio 9.22 - Malware Initial Assessment - www.winitor.com

file settings about

	encoding (2)	size (bytes)	file-offset	blacklist (25)	hint (9)	value (8556)
indicators (37)	ascii	25	0x000901A2	x	-	QueryPerformanceFrequency
virustotal (error)	ascii	14	0x0009025A	x	-	VirtualProtect
dos-header (offset)	ascii	19	0x000902A2	x	-	GetCurrentProcessId
dos-stub (456 bytes)	ascii	15	0x000902F8	x	-	DeviceIoControl
rich-header (n/a)	ascii	16	0x000903A8	x	-	TerminateProcess
file-header (Apr.2016)	ascii	16	0x000903BC	x	-	GetCurrentThread
optional-header (GUI)	ascii	11	0x000903D0	x	-	OpenProcess
directories (6)	ascii	10	0x00090410	x	-	DeleteFile
sections (file)	ascii	13	0x0009042A	x	-	FindFirstFile
libraries (4) *	ascii	12	0x0009043C	x	-	FindNextFile
functions (133)	ascii	9	0x0009047E	x	-	WriteFile
exports (n/a)	ascii	13	0x000904EC	x	-	CreateProcess
tls-callbacks (n/a)	ascii	24	0x000905D2	x	-	CreateToolhelp32Snapshot
.NET (n/a)	ascii	14	0x000905EE	x	-	Process32First
resources (executable) *	ascii	13	0x00090600	x	-	Process32Next
strings (8556)	ascii	16	0x0009064C	x	-	SetPriorityClass
debug (n/a)	ascii	25	0x00090670	x	-	QueryFullProcessImageName
manifest (administrator)	ascii	19	0x000906BA	x	-	GetForegroundWindow
version (n/a)	ascii	13	0x000907C0	x	-	OpenClipboard
certificate (n/a)	ascii	14	0x000907D0	x	-	CloseClipboard
overlay (n/a)	ascii	16	0x000907E2	x	-	SetClipboardData
	ascii	14	0x000907F6	x	-	EmptyClipboard
	ascii	9	0x00090808	x	-	SendInput
	ascii	13	0x00090814	x	-	MapVirtualKey
	ascii	16	0x0009086A	x	-	GetClipboardData
	ascii	5	0x00038399	-	format-string	D8%G
	ascii	4	0x000586E3	-	file	ds.H
	ascii	9	0x0009015E	-	file	ntdll.dll
	ascii	12	0x000906AA	-	file	KERNEL32.dll

sha256: 2606E7850597F8C4D22E7C28CCAE1FEDD98DC91B307EA56AD9C69B5EED78C363 cpu: 64-bit file-type: executable subsystem: GUI entry-point: 0x0005A48C signature: n/a

The above screenshot from PEstudio shows strings found by PEstudio in a malware sample. This can be done by selecting strings in the left pane after loading the PE file in PEstudio. The blacklist here shows a bunch of Windows API calls, which PEstudio flags as potentially used in malicious processes. You can learn about these APIs using resources like [MalAPI](#) or [MSDN](#).

Obfuscated strings:

Searching for strings often proves one of the most effective first steps in malware analysis. As seen in [the case of Wannacry](#), effective use of string search can often disrupt malware propagation and infection. The malware authors know this and don't want a simple string search to thwart their malicious activities. Therefore, they deploy techniques to obfuscate strings in their malware. Malware authors use several techniques to obfuscate the key parts of their code. These techniques often render a string search ineffective, i.e., we won't find much information when we search for strings.

Mandiant (then FireEye) launched FLOSS to solve this problem, short for **FireEye Labs Obfuscated String Solver**. FLOSS uses several techniques to deobfuscate and extract strings that would not be otherwise found using a string search. The type of strings that FLOSS can extract and how it works can be found in [Mandiant's blog post](#).

To execute FLOSS, open a command prompt and navigate to the Desktop directory. From there, use the following command.

```
C:\Users\Administrator\Desktop>floss -h
```

This command will open the help page for FLOSS. We can use the following command to use FLOSS to search for obfuscated strings in a binary.

```
C:\Users\Administrator\Desktop>floss --no-static-strings <path to binary>
```

Please remember that the command might take some time to execute, and you might see what appear to be some error messages before the results are generated.

Answer the questions below

On the Desktop in the attached VM, there is a directory named 'mal' with malware samples 1 to 6. Use floss to identify obfuscated strings found in the samples named 2, 5, and 6. Which of these samples contains the string 'DbgView.exe'?

✓ Correct Answer

🔍 Hint

When analyzing malware, it is often required to identify unique malware and differentiate them from each other. File names can't be used for this purpose as they can be duplicated easily and might be confusing. Also, a file name can be changed easily as well.



Hence, a hash function is used to identify a malware sample uniquely. A hash function takes a file/data of arbitrary length as input and creates a fixed-length unique output based on file contents. This process is irreversible, as you can't recreate the file's contents using the hash. Hash functions have a very low probability (practically zero) of two files having different content but the same hash. A hash remains the same as long as the file's content remains the same. However, even a slight change in content will result in a different hash. It might be noted that the file name is not a part of the content; therefore, changing the file name does not affect the hash of a file.

Besides identifying files, hashes are also used to store passwords to authenticate users. In malware analysis, hash files can be used to identify unique malware, search for this malware in different malware repositories and databases, and as an Indicator of Compromise (IOC).

Commonly used methods of calculating File hashes:

For identification of files, a hash of the complete file is taken. There are various methods to take the hash. The most commonly used methods are:

- Md5sum
- Sha1sum
- Sha256sum

The first two types of hashes are now considered insecure or prone to collision attacks (when two or more inputs result in the same hash). Although a collision attack for these hash functions is not very probable, it is still possible. Therefore, sha256sum is currently considered the most secure method of calculating a file hash. In the attached VM, we can see that multiple utilities calculate file hashes for us.

Finding Similar files using hashes:

Another scenario in which hash functions help a malware analyst is identifying similar files using hashes. We already established that even a slight change in the contents of a file would result in a different hash. However, some types of hashes can help identify the similarity among different files. Let's learn about some of these.

Imphash:

The imphash stands for "import hash". Imports are functions that an executable file imports from other files or Dynamically Linked Libraries (DLLs). The imphash is a hash of the function calls/libraries that a malware sample imports and the order in which these libraries are present in the sample. This helps identify samples from the same threat groups or performing similar activities. More details on the Imphash can be found on Mandiant's blog [here](#).

We can use PEstudio to calculate the Imphash of a sample.

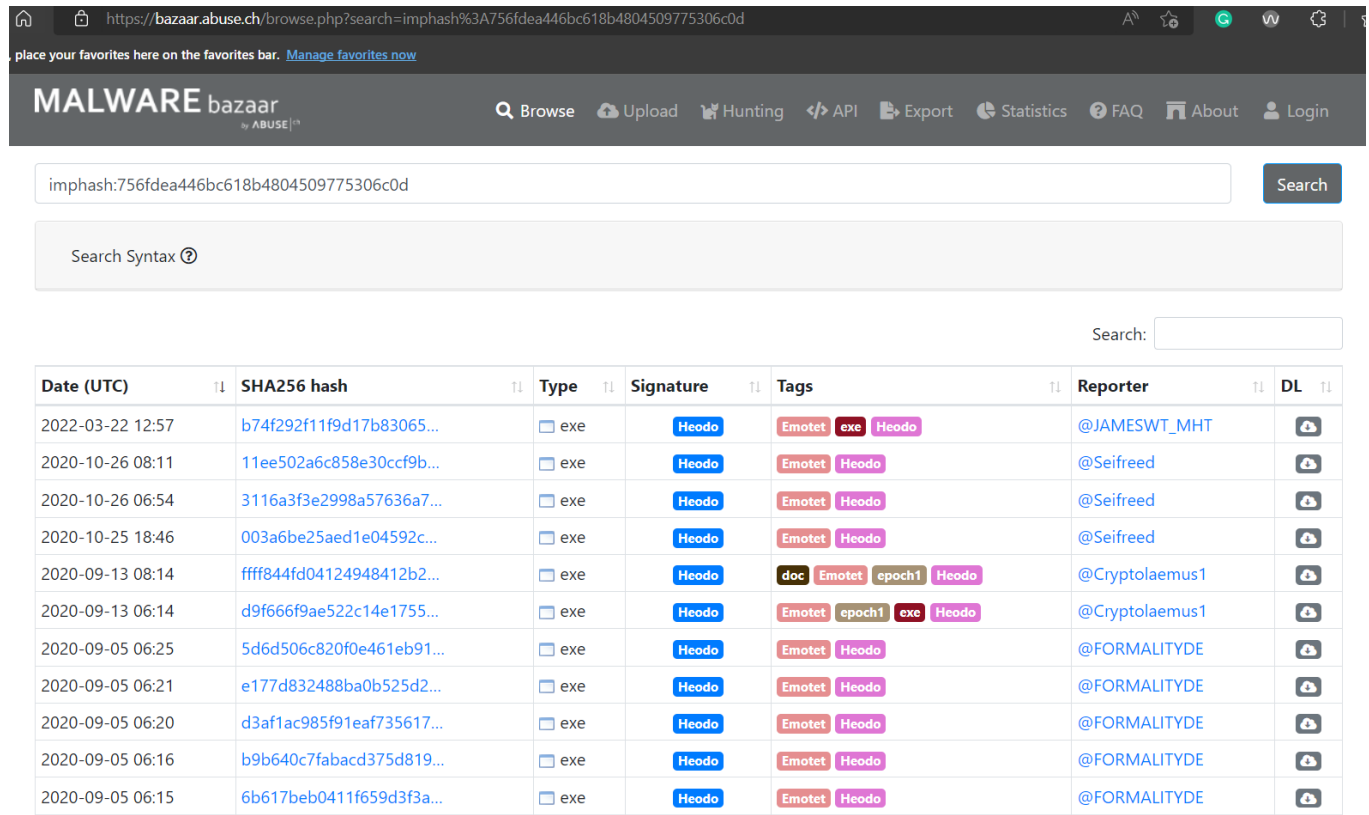
pestudio 9.22 - Malware Initial Assessment - www.winitor.com

file settings about

c:\users\administrator\desktop\mal\6

property	value
md5	7DEBB7345BCFFDF1C5328F51524EE16B
sha1	86875DEFA552651DAC3C7360607E66971595A8E2
sha256	2606E7850597F8C4D22E7C28CCAE1FEDD98DC91B307EA56AD9C69B5EED7BC363
first-bytes-hex	4D 5A 90 00 03 00 00 04 00 00 00 FF FF 00 00 B8 00 00 00 00 00 00 40 00 00 00 00 00 00 00
first-bytes-text	M Z @
file-size	870912 (bytes)
entropy	6.507
imphash	7997426B781E3208CC031754360DF27E
signature	n/a
entry-point	48 8B C4 48 89 58 08 48 89 68 10 48 89 70 18 48 89 78 20 41 56 48 83 EC 20 4C 8B F1 49 8B F0 33 C9
file-version	n/a
description	n/a
file-type	executable
cpu	64-bit
subsystem	GUI
compiler-stamp	0x5721FD88 (Thu Apr 28 05:09:44 2016)
debugger-stamp	n/a
resources-stamp	0x00000000 (empty)
import-stamp	0x00000000 (empty)
exports-stamp	n/a
version-stamp	n/a
certificate-stamp	n/a

Any malware samples with the same imports in the same order will have the same imphash. This helps in identifying similar samples.



The screenshot shows the Malware Bazaar website interface. At the top, there's a navigation bar with links like Browse, Upload, Hunting, API, Export, Statistics, FAQ, About, and Login. Below this is a search bar where the imphash '756fdea446bc618b4804509775306c0d' has been entered. A search button is to the right. Below the search bar, there's a 'Search Syntax' link. The main content area displays a table of search results. The table has columns for Date (UTC), SHA256 hash, Type, Signature, Tags, Reporter, and DL. The results show 12 entries, all of which are 'exe' files, have a 'Heodo' signature, and are tagged with 'Emotet' and 'Heodo'. The reporters for these samples are @JAMESWT_MHT, @Seifreed, @Cryptolaemus1, and @FORMALITYDE.

Date (UTC)	SHA256 hash	Type	Signature	Tags	Reporter	DL
2022-03-22 12:57	b74f292f11f9d17b83065...	exe	Heodo	Emotet, exe, Heodo	@JAMESWT_MHT	DL
2020-10-26 08:11	11ee502a6c858e30ccf9b...	exe	Heodo	Emotet, Heodo	@Seifreed	DL
2020-10-26 06:54	3116a3f3e2998a57636a7...	exe	Heodo	Emotet, Heodo	@Seifreed	DL
2020-10-25 18:46	003a6be25aed1e04592c...	exe	Heodo	Emotet, Heodo	@Seifreed	DL
2020-09-13 08:14	ffff844fd04124948412b2...	exe	Heodo	doc, Emotet, epoch1, Heodo	@Cryptolaemus1	DL
2020-09-13 06:14	d9f666f9ae522c14e1755...	exe	Heodo	Emotet, epoch1, exe, Heodo	@Cryptolaemus1	DL
2020-09-05 06:25	5d6d506c820f0e461eb91...	exe	Heodo	Emotet, Heodo	@FORMALITYDE	DL
2020-09-05 06:21	e177d832488ba0b525d2...	exe	Heodo	Emotet, Heodo	@FORMALITYDE	DL
2020-09-05 06:20	d3af1ac985f91eaf735617...	exe	Heodo	Emotet, Heodo	@FORMALITYDE	DL
2020-09-05 06:16	b9b640c7fabacd375d819...	exe	Heodo	Emotet, Heodo	@FORMALITYDE	DL
2020-09-05 06:15	6b617beb0411f659d3f3a...	exe	Heodo	Emotet, Heodo	@FORMALITYDE	DL

In the above screenshot from [Malware Bazaar](#), all these samples have the same imphash. We can see that all of these samples are classified as the same malware family. We can see that their sha256 hash is vastly different and doesn't provide any information as to their similarity. However, the same imphash helps us identify that they might belong to the same family.

Fuzzy hashes/SSDEEP:

Another way to identify similar malware is through fuzzy hashes. A fuzzy hash is a Context Triggered Piecewise Hash (CTPH). This hash is calculated by dividing a file into pieces and calculating the hashes of the different pieces. This method creates multiple inputs with similar sequences of bytes, even though the whole file might be different. More information on SSDEEP can be found on this [link](#).

Multiple utilities can be used in the attached VM to calculate ssdeep, like CyberChef. However, the ssdeep utility has been placed on the Desktop to make it easier. The following command shows the help menu of the utility.

Finding similar files using ssdeep

```
C:\Users\Administrator\Desktop>ssdeep-2.14.1\ssdeep.exe -h
ssdeep version 2.14.1 by Jesse Kornblum and the ssdeep Project
For copyright information, see man page or README.TXT.
```

```
Usage: ssdeep [-m file] [-k file] [-dpgvrsblcxa] [-t val] [-h|-V] [FILES]
-m - Match FILES against known hashes in file
-k - Match signatures in FILES against signatures in file
-d - Directory mode, compare all files in a directory
-p - Pretty matching mode. Similar to -d but includes all matches
-g - Cluster matches together
-v - Verbose mode. Displays filename as its being processed
-r - Recursive mode
-s - Silent mode; all errors are suppressed
-b - Uses only the bare name of files; all path information omitted
-l - Uses relative paths for filenames
-c - Prints output in CSV format
-x - Compare FILES as signature files
-a - Display all matches, regardless of score
-t - Only displays matches above the given threshold
-h - Display this help message
-V - Display version number and exit
```

```
FLARE Sun 09/18/2022 17:28:35.11
C:\Users\Administrator\Desktop>
```

Let's calculate the hashes of all the samples in the mal directory in the attached VM.

Calculating ssdeep

```
C:\Users\Administrator\Desktop>ssdeep-2.14.1\ssdeep.exe mal\*
ssdeep,1.1--blocksize:hash:hash,filename
3072:C3twbyJdvGwRCf/swDQhe0AmN4hMRl37G:8Eac0AmN6C,"C:\Users\Administrator\Desktop\mal\1"
768:fMjB/JpMfHDWqpuXDvod3UmQmv4acY2GS2C9xjwhU:UFQlpSDvoJrbvUfGS2q,"C:\Users\Administrator\Desktop\mal\2"
1536:C3tvICAqw8IKVn2wJk0c8PoYJvGwRCwAL6pILgl7vBIQtCnDkbZ3e0AmV2u4hnnM:C3twbyJdvGwRCf/swDQhe0AmN4hM,"C:\Users\Administrator\Desktop\mal\3"
24576:u7DtLSdAlZvEFZhbs7buPPcedehP5XLnk03hGL8Siw9zVZprY8fWg5r110:80KVizL3cvtk03hmVVZBYu5r1M,"C:\Users\Administrator\Desktop\mal\4"
12288:z+IIs67xrXWxgxMdpLNGvIcGZwwDVHJXuDzKYzIE5P/XiS6lYSz8uahDtbNL6WTW:z+PsGlsFGgcQDJJQ,"C:\Users\Administrator\Desktop\mal\5"
24576:UCsTPcqE9S7td0DN+6ybJVCy2pvZHG0zPBjRj4AFsb:UC07tsp+6ybJVChpRjvs,"C:\Users\Administrator\Desktop\mal\6"
```

```
FLARE Sun 09/18/2022 17:41:12.38
C:\Users\Administrator\Desktop>
```

We can try the other options shown in the help file per the requirement. When we have the ssdeep hashes, we can match these hashes together to identify similar files. This helps us identify similar files if we have a bulk of data. The documentation link provided above has very good examples of usage. The following terminal window shows one of the examples relevant to our use case to match files. For this, we can use the `-d` operator. The `-r` operator runs the ssdeep utility recursively, and the `-l` operator outputs relative paths of the files.

Finding matching files using ssdeep

```
C:\Users\Administrator\Desktop>ssdeep-2.14.1\ssdeep -l -r -d Incoming Outgoing
Trash
Outgoing/Corporate Espionage/Our Budget.doc matches Incoming/Budget 2007.doc
(99)
Outgoing/Personnel Mayhem/Your Buddy Makes More Than You.doc matches
Incoming/Salaries.doc (45)
Trash/DO NOT DISTRIBUTE.doc matches Outgoing/Plan for Hostile Takeover.doc
(88)
```

```
FLARE Sun 09/18/2022 17:41:12.38
C:\Users\Administrator\Desktop>
```

The results show files that match each other. The number in the bracket at the end is the percentage of matches among the files.

Answer the questions below

In the samples located at Desktop\mal\ directory in the attached VM, which of the samples has the same imphash as file 3?

✓ Correct Answer

Using the ssdeep utility, what is the percentage match of the above-mentioned files?

✓ Correct Answer

In the previous task, we learned how hashes could identify identical files. We also found out that hashes can be changed by changing even a single byte of data in a file and how specific hashes like imphash and ssdeep help us identify file similarities. While using imphash or ssdeep provides a way to identify if some files are similar, sometimes we just need to identify if a file contains the information of interest. Hashes are not the ideal tool to perform this task.

Signatures:



Signatures are a way to identify if a particular file has a particular type of content. We can consider a signature as a pattern that might be found inside a file. This pattern is often a sequence of bytes in a file, with or without any context regarding where it is found. Security researchers often use signatures to identify patterns in a file, identify if a file is malicious, and identify suspected behavior and malware family.

Yara rules:

Yara rules are a type of signature-based rule. It is famously called a pattern-matching swiss army knife for malware researchers. Yara can identify information based on binary and textual patterns, such as hexadecimal and strings contained within a file. TryHackMe has a [dedicated room](#) for Yara rules if it interests you.

The security community publishes a [repository](#) of open-source Yara rules that we can use as per our needs. When analyzing malware, we can use this repository to dig into the community's collective wisdom. However, while using these rules, please keep in mind that some might depend on context. Some others might just be used for the identification of patterns that can be non-malicious as well. Hence, just because a rule hits doesn't mean the file is malicious. For a better understanding, please read the documentation for the particular rule to identify the use case where it will be applicable in the best possible manner.

Proprietary Signatures - AntiVirus Scans:

Besides the open-source signatures, Antivirus companies spend lots of resources to create proprietary signatures. The advantage of these proprietary signatures is that since they have to be sold commercially, there are lesser chances of False Positives (FPs, when a signature hits a non-malicious file). However, this might lead to a few False Negatives (FNs, when a malicious file does not hit any signature).

Antivirus scanning helps identify if a file is malicious with high confidence. Antivirus software will often mention the signature that the file has hit, which might hint at the file's functionality. However, we must note that despite their best efforts, every AV product in the market has some FPs and some FNs. Therefore, when analyzing malware, it is prudent to get a verdict from multiple products. The [VirusTotal](#) website makes this task easier for us, where we can find the verdict about a file from 60+ AV vendors, apart from some very useful information. We also touched upon this topic in our Intro to Malware Analysis room. Please remember, if you are analyzing a sensitive file, it is best practice to search for its hash on Virustotal or other scanning websites instead of uploading the file itself. This is done to avoid leaking sensitive information on the internet and letting a sophisticated attacker know that you are analyzing their malware.

Since we have covered Yara rules in detail in the Yara room and Virustotal scanning in the Intro to malware analysis room, we will not cover them again here. However, the FLARE VM has another very cool tool that can be used for signature scanning.

Capa:

Capa is another open-source tool created by Mandiant. This tool helps identify the capabilities found in a PE file. Capa reads the files and tries to identify the behavior of the file based on signatures such as imports, strings, mutexes, and other artifacts present in the file. For further detail into the background of Capa, we can visit its [Github page](#) or Mandiant's [blog post](#) introducing Capa.

Using Capa is simple. On the command prompt, we just point capa to the file we want to run it against.

```
C:\Users\Administrator\Desktop>capa mal.exe
```

The `-h` operator shows detailed options.

Capa

```
C:\Users\Administrator\Desktop>capa -h
usage: capa.exe [-h] [--version] [-v] [-vv] [-d] [-q] [--color]
```

```
{auto,always,never}] [-f {auto,pe,sc32,sc64,freeze}] [-b {vivisect,smda}] [-r
RULES] [-t TAG] [-j] sample
```

The FLARE team's open-source tool to identify capabilities in executable files.

positional arguments:

sample path to sample to analyze

optional arguments:

-h, --help show this help message and exit
--version show program's version number and exit
-v, --verbose enable verbose result document (no effect with --json)
-vv, --vverbose enable very verbose result document (no effect with --json)
-d, --debug enable debugging output on STDERR
-q, --quiet disable all output but errors
--color {auto,always,never} enable ANSI color codes in results, default: only

during interactive session

-f {auto,pe,sc32,sc64,freeze}, --format {auto,pe,sc32,sc64,freeze} select sample format, auto: (default) detect file type automatically, pe: Windows PE file, sc32: 32-bit shellcode, sc64: 64-bit shellcode, freeze: features previously frozen by capa
-b {vivisect,smda}, --backend {vivisect,smda} select the backend to use
-r RULES, --rules RULES path to rule file or directory, use embedded rules by default
-t TAG, --tag TAG filter on rule meta field values
-j, --json emit JSON instead of text

By default, capa uses a default set of embedded rules.

You can see the rule set here:

<https://github.com/fireeye/capa-rules>

To provide your own rule set, use the '-r' flag:

```
capa --rules /path/to/rules suspicious.exe
capa -r /path/to/rules suspicious.exe
```

examples:

```
identify capabilities in a binary
capa suspicious.exe
```

```
identify capabilities in 32-bit shellcode, see '-f' for all supported
```

```
capa -f sc32 shellcode.bin
```

```
capa -v suspicious.exe
```

```
capa -vv suspicious.exe
```

```
capa -t "create TCP socket" suspicious.exe
```

C:\Users\Administrator\Desktop>

Capa example

[illegible][illegible]

-----+

1

1

6ec74cc0a9b5697efd3f4cc4d3a21d9ffe6e0187b770990df8743fbf4f3b2518

1

1

-----+

-----+

-----+

1

-----+

DEFENSE EVASION	Obfuscated Files or Information::Indicator Removal
from Tools [T1027.005]	
	Obfuscated Files or Information [T1027]
DISCOVERY	Application Window Discovery [T1010]
	System Information Discovery [T1082]
EXECUTION	Command and Scripting Interpreter [T1059]
	Shared Modules [T1129]
-----+	-----+
-----+	
-----+	
MBC Objective	MBC Behavior
-----+	-----+
-----+	
ANTI-STATIC ANALYSIS	Disassembler Evasion::Argument Obfuscation
[B0012.001]	
CRYPTOGRAPHY	Encrypt Data::RC4 [C0027.009]
	Generate Pseudo-random Sequence::RC4 PRGA
[C0021.004]	
FILE SYSTEM	Delete File [C0047]
	Read File [C0051]
	Write File [C0052]
OPERATING SYSTEM	Console [C0033]
PROCESS	Allocate Thread Local Storage [C0040]
	Set Thread Local Storage Value [C0041]
	Terminate Process [C0018]
-----+	-----+
-----+	
-----+	
-----+	

CAPABILITY	NAMESPACE
-----	-----+
contain obfuscated stackstrings analysis/obfuscation/string/stackstring	anti-
encrypt data using RC4 PRGA manipulation/encryption/rc4	data-
contains PDB path	executable/pe/pdb
contain a resource (.rsrc) section executable/pe/section/rsrc	
accept command line arguments	host-interaction/cli
manipulate console interaction/console	host-
query environment variable interaction/environment-variable	host-
delete file interaction/file-system/delete	host-
read file interaction/file-system/read	host-
write file (2 matches) interaction/file-system/write	host-
enumerate gui resources	host-interaction/gui
get disk information interaction/hardware/storage	host-
get hostname interaction/os/hostname	host-
get thread local storage value (3 matches) interaction/process	host-
set thread local storage value (2 matches) interaction/process	host-
terminate process (5 matches) interaction/process/terminate	host-
link function at runtime (8 matches) linking	linking/runtime-
link many functions at runtime linking	linking/runtime-
parse PE exports (2 matches)	load-code/pe
parse PE header (4 matches)	load-code/pe
-----+	-----+
-----+	

```
FLARE Sun 09/18/2022 18:34:13.15  
C:\Users\Administrator\Desktop>
```

We can see that Capa has mapped the identified capabilities according to the MITRE ATT&CK framework and [Malware Behavior Catalog \(MBC\)](#). In the last table, we see the capabilities against the matched signatures and the number of signatures that have found a hit against these capabilities. As we might see, it also tells us if there are obfuscated stackstrings in the sample, allowing us to identify if running FLOSS against the sample might be helpful. To find out more information about the sample, we can use the `-v` or the `-vv` operator, which will show us the results in verbose or very verbose mode, identifying addresses where we might find the said capability.

Answer the questions below

How many matches for anti-VM execution techniques were identified in the sample?

86

✓ Correct Answer

Does the sample have to capability to suspend or resume a thread? Answer with Y for yes and N for no.

Y

✓ Correct Answer

What MBC behavior is observed against the MBC Objective 'Anti-Static Analysis'?

Disassembler Evasion::Argument Obfuscation [B0012.001]

✓ Correct Answer

At what address is the function that has the capability 'Check HTTP Status Code'?

0x486921

✓ Correct Answer

🔍 Hint

So far in this room, we have covered techniques that work regardless of the file type of the malware. However, those techniques are a little hit-and-miss, as they don't always provide us with deterministic information about the malware. The PE headers provide a little more deterministic characteristics of the sample, which tells us much more about the sample.



The PE header:

The programs that we run are generally stored in the executable file format. These files are portable because they can be taken to any system with the same Operating System and dependencies, and they will perform the same task on that system. Therefore, these files are called Portable Executables (PE files). The PE files consist of a sequence of bits stored on the disk. This sequence is in a specific format. The initial bits define the characteristics of the PE file and explain how to read the rest of the data. This initial part of the PE file is called a PE header.

Several tools in the FLARE VM can help us analyze PE headers. PESTudio is one of them. We already familiarized ourselves with PESTudio in a previous task, so we will just use that in this task as well.

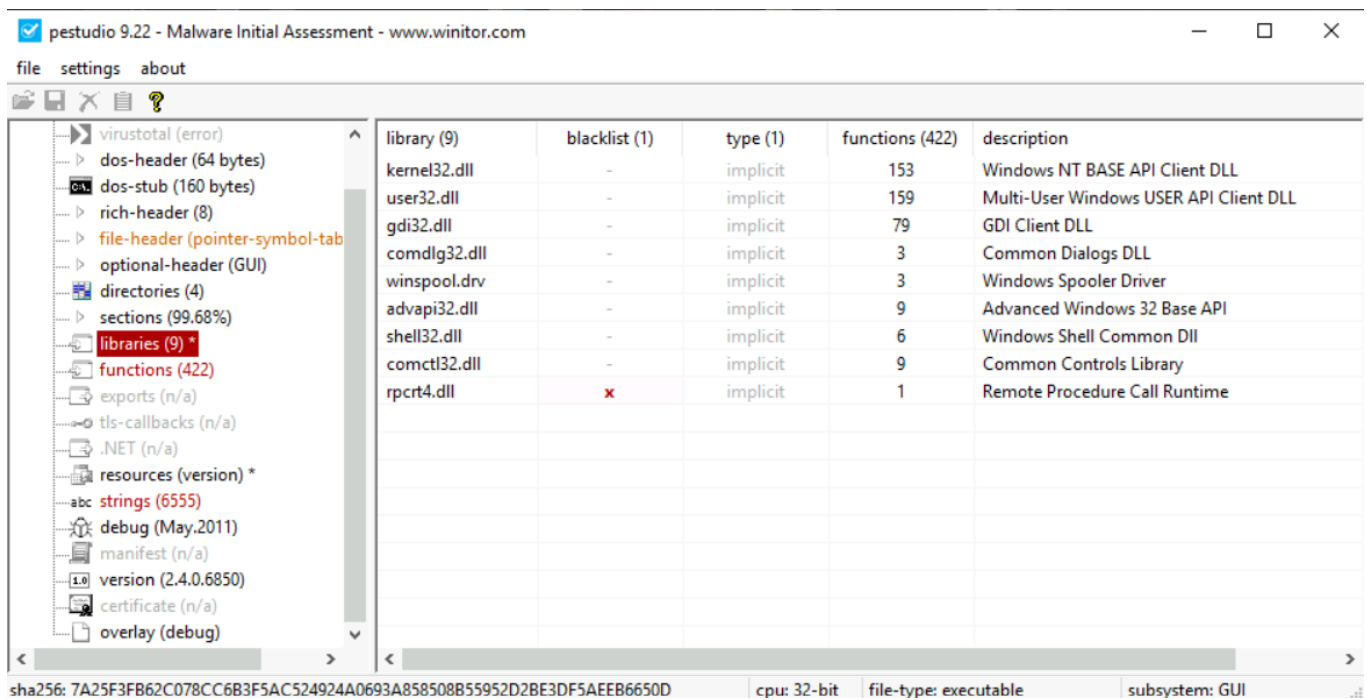
The PE header contains rich information useful for malware analysis. We will learn about this data in detail in the [Dissecting PE headers](#) room. However, we can get the following information from a PE header as an overview.

Linked Libraries, imports, and functions:

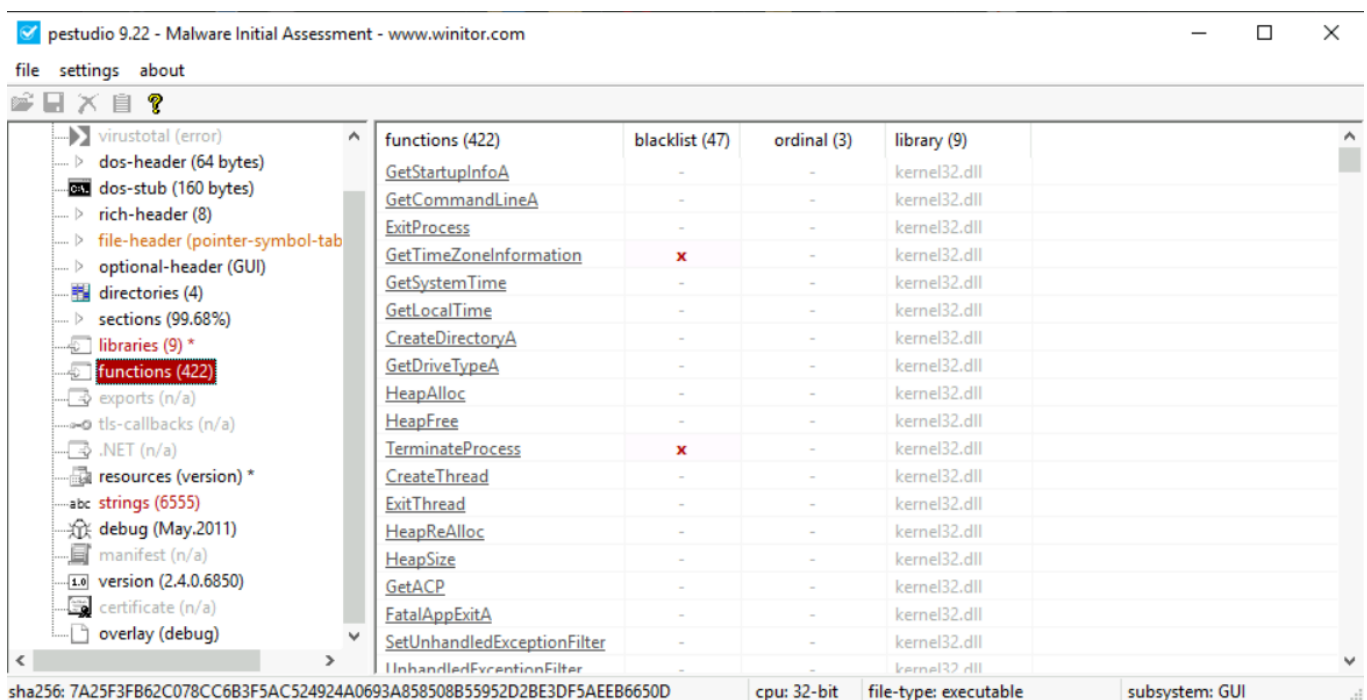
A PE file does not contain all of its code to perform all the tasks. It often reuses code from libraries, often provided by Microsoft as part of the Windows Operating System. Often, certain functions from these libraries are imported by the PE file. The PE header contains information about the libraries that a PE file uses and the functions it imports from those libraries. This information is very useful. A malware analyst can look at the libraries and functions that a PE file imports and get a rough idea of the functionality of a malware sample. For example, if a malware sample imports the CreateProcessA function, we can assume that this sample will create a new process.

Similarly, other functions can provide further information about the sample. However, it must be noticed that we don't know the context in which these functions are called by just looking at the PE headers. We need to dig deeper into that, which we will cover in the upcoming rooms.

PEStudio has a libraries option in the right pane, which, when selected, shows us the libraries that a PE file will use.



The functions option just below shows the functions imported from these libraries.



Identifying Packed Executables:

As we have seen so far, static analysis bares a lot of information that can be used against the malware. Malware authors understand this as a problem. Therefore, they often go to great lengths to thwart analysis. One of the ways they can do this is by packing the original sample inside a shell-type code that obfuscates the properties of the actual malware sample. This technique is called packing, and the resultant PE file is called a packed PE file. Packing greatly

reduces the effectiveness of some of the malware analysis techniques we have learned about so far. For example, if we try to search for strings in a packed executable, we might not find anything useful because of packing. Similarly, searching for similar samples using ssdeep might return more samples packed with the same packer instead of samples behaviorally similar to the sample of interest. Some signatures might also be evaded due to a malware sample being packed.

As we will see in the upcoming [Dissecting PE headers](#) room, we can identify packed executables by analyzing the PE header of a malware sample. The PE header contains important information such as the number of sections, permissions of different sections, sizes of the sections, etc. This information can give us pointers to help identify if the malware sample is packed and, if so, which type of packer has packed the executable.

Answer the questions below

Open the sample Desktop\mal\4 in PEstudio. Which library is blacklisted?

rpcrt4.dll

✓ Correct Answer

🔍 Hint

What does this dll do?

Remote Procedure Call Runtime

✓ Correct Answer

🔍 Hint