

Malware Analysis (CS6038)

Week 03.1 Static Analysis

Scott Nusbaum

nusbausa@ucmail.uc.edu

January 29, 2019

Overview

- Revisit Last weeks failed Demo
- Static Analysis
 - Definition
 - File Structures
 - Tools overview
 - Examples

Demo Retry

- Show Metasploit connection and send across info

Static Analysis

Static Analysis is the process of documenting your observations about what identifying characteristics a malware sample exhibits. The goal of this process is that, after analysis, you have extracted some identifying characteristics from a malware sample that can be used to help you search further for more samples of that malware (and, hopefully, others that are similar to it).

We distinguish **static analysis** to focus on how a sample “looks”, for the purpose of identifying any samples of it that may be dormant and inactive within your attack surface. This is different from dynamic analysis, where we are trying to define the actions it takes or may take when executed on a system.

Files & Structures

- EXE
 - MZ/PE Headers
- ELF
- PDF
- PNG
- JPG
- Doc vs Docx
- Undocumented File Structures

Exe Headers

PE File format

offset	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0x00000000	0x5A4D (MZ)		lastsize	PagesInFile		relocations		headerSizeInParagraph	MinExtraParagraphNeeded	MaxExtraParagraphNeeded	Initial (relative) SS					
0x00000010	Initial (relative) SP		checksum	Initial IP		Initial (relative) CS		FileAddOfRelocTable	OverlayNumber	reserved		reserved				
0x00000020	reserved		reserved	OEMIdentifier		OEMInformation		reserved	reserved	reserved		reserved				
0x00000030	reserved		reserved	reserved		reserved		reserved	reserved	0x80 (offset to PE signature)						
0x00000040	This block contains instructions to display the message "This program cannot be run in DOS mode" when run in MS-DOS															
0x00000050																
0x00000060																
0x00000070																
0x00000080																
0x00000090	NumberOfSymbols (0 for image)			SizeOfOptionalHeaders		Characteristics		0x10B (exe)		InMajVer	InMnrVer	SizeOfCode				
0x000000A0	SizeOfInitializedData			SizeOfUninitializedData			AddressOfEntryPoint					BaseOfCode				
0x000000B0	BaseOfData			ImageBase			SectionAlignment					FileAlignment				
0x000000C0	MajorOSVersion		MinorOSVersion		MajorImageVersion		MinorImageVersion		MajorSubsystemVersion		MinorSubsystemVersion		Win32VersionValue			
0x000000D0	SizeOfImage			SizeOfHeaders			Checksum			Checksum		DllCharacteristics				
0x000000E0	SizeOfStackReserve			SizeOfStackCommit			SizeOfHeapReserve			SizeOfHeapCommit						
0x000000F0	LoaderFlags			NumberOF RVA and Sizes			.edata offset			.edata size						
0x0000100	.idata offset			.idata size			.rsrc offset			.rsrc size						
0x0000110	.pdata offset			.pdata size			attribute certificate offset (image)			attribute certificate size (image)						
0x0000120	.reloc offset (image)			.reloc size (image)			.debug offset			.debug size						
0x0000130	Architecture (reserved - 0x0)			Architecture (reserved - 0x0)			Global Ptr offset			must be 0x0						
0x0000140	.tls offset			.tls size			Load config table offset (image)			Load Config table size (image)						
0x0000150	Bound import table offset			Bound import table size			IAT (Import address table) offset			IAT (Import address table) size						
0x0000160	Delay import descriptor offset (image)			Delay import descriptor size (image)			CLR runtime header offset (object)			CLR runtime header size (object)						
0x0000170	Reserved (must be 0x0)			Reserved (must be 0x0)			Section header - Name									
0x0000180	VirtualSize			VirtualAddress			SizeOfRawData			PointerToRawData						
0x0000190	PointerToRelocations			PointerToLineNumbers			NumberOfRelocations		NumberOfLineNumbers		Characteristics					
0x00001A0	Section header - Name										VirtualSize			VirtualAddress		
0x00001B0	SizeOfRawData			PointerToRawData			PointerToRelocations			PointerToLineNumbers						
0x00001C0	NumberOfRelocations		NumberOfLineNumbers		Characteristics		Section header - Name..									

		Size in bytes
MS-DOS header	File header	64
PE Signature		4
COFF header		20
Standard fields		28
Windows-Specific fields		68
Data directories		variable
Section table (each section header is 40 bytes)		variable

```

1 00000000: 4d5a 9000 0300 0400 0000 ffff 0000 MZ
2 00000010: b800 0000 0000 0000 4000 0000 0000 .....@.....
3 00000020: 0000 0000 0000 0000 0000 0000 0000 .....
4 00000030: 0000 0000 0000 0000 0000 0000 d800 0000 .....
5 00000040: 0e1f ba0e 00b4 09cd 21b8 014c cd21 5468 .....!..L.!Th
6 00000050: 6973 2070 726f 6772 616d 2063 616e 6e6f is program canno
7 00000060: 7420 6265 2072 756e 2069 6e20 444f 5320 t be run in DOS
8 00000070: 6d6f 6465 2e0d 0d0a 2400 0000 0000 0000 mode...$.
9 00000080: e57d 4aa8 a11c 24fb a11c 24fb a11c 24fb .}J...$...$...$.
10 00000090: 2f14 7bfb a31c 24fb a11c 25fb 3d1c 24fb /.{...$...%...$.
11 000000a0: 2214 79fb b01c 24fb f53f 14fb a81c 24fb ".y...$..?....$.
12 000000b0: 661a 22fb a01c 24fb 5269 6368 a11c 24fb f."...$.Rich..$.
13 000000c0: 0000 0000 0000 0000 0000 0000 0000 0000 .....
14 000000d0: 0000 0000 0000 0000 5045 0000 4c01 0500 .....PE..L...
15 000000e0: 1a5d 0f43 0000 0000 0000 0000 e000 0f01 .].C.....
16 000000f0: 0b01 0600 005a 0000 00da 0100 0004 0000 .....Z.....
17 00000100: 3533 0000 0010 0000 0070 0000 0000 4000 53.....p....@.
18 00000110: 0010 0000 0002 0000 0400 0000 0000 0000 .....
19 00000120: 0400 0000 0000 0000 0040 0300 0004 0000 .....@.....
20 00000130: 0000 0000 0200 0000 0000 1000 0010 0000 .....
21 00000140: 0000 1000 0010 0000 0000 0000 1000 0000 .....
22 00000150: 0000 0000 0000 0000 c072 0000 b400 0000 .....r.....
23 00000160: 00d0 0200 0070 0000 0000 0000 0000 0000 .....p.....
24 00000170: 0000 0000 0000 0000 0000 0000 0000 0000 .....
25 00000180: 0000 0000 0000 0000 0000 0000 0000 0000 .....
26 00000190: 0000 0000 0000 0000 0000 0000 0000 0000 .....
27 000001a0: 0000 0000 0000 0000 0000 0000 0000 0000 .....
28 000001b0: 0070 0000 9002 0000 0000 0000 0000 0000 .p.....
29 000001c0: 0000 0000 0000 0000 0000 0000 0000 0000 .....
30 000001d0: 2e74 6578 7400 0000 cc59 0000 0010 0000 .text...Y.....
31 000001e0: 005a 0000 0004 0000 0000 0000 0000 0000 .Z.....
32 000001f0: 0000 0000 2000 0060 2e72 6461 7461 0000 .... \.rdata..
33 00000200: cc10 0000 0070 0000 0012 0000 005e 0000 .....p.....^..
34 00000210: 0000 0000 0000 0000 0000 0000 4000 0040 .....@..@
35 00000220: 2e64 6174 6100 0000 14b4 0100 0090 0000 .data.....
36 00000230: 0004 0000 0070 0000 0000 0000 0000 0000 .....p.....
37 00000240: 0000 0000 4000 00c0 2e6e 6461 7461 0000 ....@....ndata..

```

Elf Header

```
>hexdump -C ./compile_me.elf | head -n 10
00000000  7f 45 4c 46 02 01 01 00 00 00 00 00 00 00 00 00 |.ELF.....|
00000010  02 00 3e 00 01 00 00 00 30 04 40 00 00 00 00 00 |..>.....0.@...|
00000020  40 00 00 00 00 00 00 00 00 1a 00 00 00 00 00 00 |@.....|
00000030  00 00 00 00 40 00 38 00 09 00 40 00 1f 00 1c 00 |...@.8...@...|
00000040  06 00 00 00 05 00 00 00 40 00 00 00 00 00 00 00 |.....@.....|
00000050  40 00 40 00 00 00 00 00 40 00 40 00 00 00 00 00 |@ @ @ @ @ @ @ @|
00000060  f8 01 00 00 00 00 00 00 f8 01 00 00 00 00 00 00 |.....|
00000070  08 00 00 00 00 00 00 00 03 00 00 00 04 00 00 00 |.....|
00000080  38 02 00 00 00 00 00 00 38 02 40 00 00 00 00 00 |8.....8.@...|
00000090  38 02 40 00 00 00 00 00 1c 00 00 00 00 00 00 00 |8@.....|
kh3m@kh3m-machine:~/Research/ELF/tests/baseline/compile_options$

>readelf -l ./compile_me.elf | head -n 20

Elf file type is EXEC (Executable file)
Entry point 0x400430
There are 9 program headers, starting at offset 64

Program Headers:
  Type           Offset             VirtAddr           PhysAddr
  FileSiz        MemSiz            Flags             Align
PHDR            0x0000000000000040 0x0000000000040040 0x0000000000004000
INTERP          0x00000000000001f8 0x00000000000001f8 R E               8
LOAD            0x0000000000000238 0x0000000000040238 0x0000000000040238
LOAD            0x000000000000001c 0x000000000000001c R                 1
[Requesting program interpreter: /lib64/ld-linux-x86-64.so.2]
LOAD            0x0000000000000000 0x0000000000040000 0x0000000000040000
LOAD            0x0000000000000077c 0x0000000000000077c R E             200000
LOAD            0x00000000000000e10 0x00000000000600e10 0x00000000000600e10
LOAD            0x00000000000000228 0x00000000000000230 RW             200000
```


PDF Header

PDF File Specification

```
1 00000000: 2550 4446 2d31 2e33 0a25 c4e5 f2e5 eba7 %PDF-1.3%.
2 00000010: f3a0 d0c4 c60a 3420 3020 6f62 6a0a 3c3c .....4 0 obj.<<
3 00000020: 202f 4c65 6e67 7468 2035 2030 2052 202f /Length 5 0 R /
4 00000030: 4669 6c74 6572 202f 466c 6174 6544 6563 Filter /FlateDec
5 00000040: 6f64 6520 3e3e 0a73 7472 6561 6d0a 7801 ode >>.stream.x.
6 00000050: 8d54 db8e d330 107d f757 0c6c 9a26 2d71 .T...0.}.W.l.&-q
7 00000060: 6dc7 b951 aebb 2004 4fac 1489 0796 0754 m..Q.. .0.....T
8 00000070: b52c 520b 6c0b ffcf 193b 4eb2 55b8 3452 .,R.l....;N.U.4R
9 00000080: 3d39 1ecf 9c39 33ce 1d5d d31d 29b2 aaa0 =9...93...)..
10 00000090: 0355 c65b 7b67 29a9 0ded b1d9 19b7 f481 .U. [{g).....
11 000000a0: bef1 6bef 77dc 3a68 7575 d262 7322 ed9e ..k.w.:huu.bs"..
12 000000b0: d3e6 bfe3 eda6 b28b 2ebb 4f8d ff3f e77d .....0..?.}
```

PNG Header

File Specification

Offset (h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	
00000000	89	50	4E	47	0D	0A	1A	0A	00	00	00	0D	49	48	44	52	%PNG....IHDR
00000010	00	00	00	6F	00	00	00	73	08	02	00	00	00	19	B3	CB	...o...s.....'E
00000020	D7	00	00	00	01	73	52	47	42	00	AE	CE	1C	E9	00	00	*....sRGB.®Î.é..
00000030	00	04	67	41	4D	41	00	00	B1	8F	0B	FC	61	05	00	00	..gAMA..±..üa...
00000040	00	09	70	48	59	73	00	00	0E	C3	00	00	0E	C3	01	C7	..pHYs...Ã...Ã.Ç
00000050	6F	A8	64	00	00	00	3C	49	44	41	54	78	5E	ED	C1	01	o``d...<IDATx^iÁ.
00000060	0D	00	00	00	C2	A0	F7	4F	6D	0F	07	04	00	00	00	00Ã ÷Om.....
00000070	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00000080	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00000090	70	AE	06	96	0A	00	01	1E	C4	F7	41	00	00	00	00	49	p@.-....Ã÷A....I
000000A0	45	4E	44	AE	42	60	82										END®B` ,

JPEG Header

File Specification

```
0: FF D8 FF E0 00 10 4A 46 49 46 00 01 02 01 01 2C y0ya..JFIF.....
10: 01 2C 00 00 FF E1 06 94 45 78 69 66 00 00 49 49 ...ya..Exif..II
20: 2A 00 08 00 00 00 06 00 1A 01 05 00 01 00 00 00 *.....
30: 56 00 00 00 1B 01 05 00 01 00 00 00 5E 00 00 00 V.....^...
40: 28 01 03 00 01 00 00 00 02 00 8E 26 9B 9C 01 00 (. ....|&|...
50: 16 00 00 00 66 00 00 00 9D 9C 01 00 2C 00 00 00 ....f...|...
60: 7C 00 00 00 9F 9C 01 00 32 00 00 00 A8 00 00 00 |...|...2...
70: DA 00 00 00 00 00 2C 01 00 00 01 00 00 00 2C 01 U.....
80: 00 00 01 00 42 00 6C 00 75 00 65 00 20 00 48 00 ....Blue..H.
90: 69 00 6C 00 6C 00 73 00 00 00 4D 00 69 00 63 00 i.l.l.s..M.i.c.
A0: 72 00 6F 00 73 00 6F 00 66 00 74 00 2C 00 20 00 r.o.s.o.f.t...
B0: 49 00 20 00 74 00 68 00 69 00 6E 00 6B 00 2E 00 I..t.h.i.n.k...
C0: 2E 00 2E 00 00 00 53 00 6F 00 6D 00 65 00 20 00 ....S.o.m.e.
D0: 6C 00 61 00 6E 00 64 00 73 00 63 00 61 00 70 00 l.a.n.d.s.c.a.p.
E0: 65 00 20 00 73 00 6F 00 6D 00 65 00 77 00 68 00 e..s.o.m.e.w.h.
F0: 65 00 72 00 65 00 00 00 03 00 03 01 03 00 01 00 e.r.e.....
100: 00 00 06 00 FF FF 01 02 04 00 01 00 00 00 04 01 ...y...
110: 00 00 02 02 04 00 01 00 00 00 88 05 00 00 00 00 .....|...
120: 00 00 FF D8 FF E0 00 10 4A 46 49 46 00 01 02 01 ..y0ya..JFIF...
130: 00 48 00 48 00 00 FF EE 00 0E 41 64 6F 62 65 00 .H.H..yi..Adobe.
140: 64 80 00 00 00 01 FF DB 00 84 00 0C 08 08 08 09 d|...y0..
150: 08 0C 09 09 0C 11 0B 0A 0B 11 15 0F 0C 0C 0F 15 .....
160: 18 13 13 15 13 13 18 11 0C 0C 0C 0C 0C 11 0C
```

Windows OLE Documents

[MS-CFB Files](#), also known as OLE are a container format common to many Microsoft applications and systems. Many people associate these with the older MSOffice files, DOC, XLS, PPT, etc. These are even more complex structures, mimicking a filesystem within a file, complete with hierarchy and block-based storage allocation

Unstructured

Unstructured Data is content for which you do not have any assigned meaning or context associated with its positioning. Upon initial review, much of the content within malware that you have yet to analyze fits this unstructured definition.

The task of Reverse Engineering includes attempting to derive what the meaning of unstructured data is within a malicious artifact.

Tools Overview

- Python
- Awk | sed | grep
- Readelf
- Strings
- [Exiftool](#)
- File
- Objdump
 - ”-M intel” for Intel Syntax
- [Jadx](#)
- [Rhino](#) | js | [nodejs](#) (apt install nodejs)
- Js-beautify (apt install node-js-beautify)
- [Powershell Beautify](#)
- [Binwalk](#)
- [Windows SysInternals](#)
- [File Analyzer](#), [PEView](#), [PE Studio](#)
- [Wireshark](#), [Tshark](#), [TcpDump](#)
- [Zeek \(Bro\)](#)
- [Hexer](#) | [xxd](#)
- [Didier Stevens](#)

Demo Disassemblers

- [IDA Pro](#)
- [Hopper](#)
- [Radare2](#)
- [Objdump](#)

Examples

- Malware Deobfuscation
 - Kovter (Fileless Malware)