In-Memory Malware Analysis

PV204 Laboratory of security and applied cryptography II

Course handouts and notes. // Václav Lorenc

In-Memory Malware Analysis

PV204 Laboratory of security and applied cryptography II

Before we start…

A short introduction, how a common attack (let's assume we are talking about targeted attacks) is usually performed:

1. Reconnaissance
2. Weaponization
3. Delivery
4. **Exploitation**
5. **C2**
6. **Exfiltration**

Malware phases/stages:

1. System Infection / Exploit
2. Dropper / Downloader (multi-stated)
3. Callbacks
4. Configuration / Updates
5. Removal

How to use in-memory analysis?

1. Run malware / acquire memory.
2. Analyze memory, have fun.
3. ???
4. Profit!

How to setup your lab system?

- Oracle VirtualBox (or another virtualization technology capable of making snapshots and memory dumps).
- Windows XP/7/8 license or trial
- Memory dumping tool
- Adobe Reader and any other vulnerable software you consider as important.

- *(if possible, disable swap)*

- *Debuggers*
- *Disassembles*
- *Text editors*

Reverse Engineering for Beginners (x86)

Registers

8bit	16bit	32bit	64bit
AL	AX	EAX	RAX
BL	BX	EBX	RBX
CL	CX	ECX	RCX
DL	DX	EDX	RDX
	SI	ESI	RSI
	DI	EDI	RDI
	BP	EBP	RBP
	SP	ESP	RSP

And of course, instructions pointer (`IP/EIP/RIP`) and flags (`flags/rflags`), segment registers (`CS, SS, DS, ES, GS, FS`), FPU registers, SSE, SSE2, …

BTW: `BX` register can be used for loops (like many other registers), but `LOOP` instruction works with `(E)CX`.

Instructions "families"

Evergreen	The past	The future
push	aaa	crc32
mov	xlat	aesenc
call	verr	pcmpistrm
retn	smsw	vfmsubadd132ps
jmp	lsl	movbe

Function Entry/Exit

```
push ebp
mov ebp, esp
sub esp, X
```

```
mov esp, ebp
pop ebp
ret X ; sizeof(x) + sizeof(y) + sizeof(z)
```

More Information

Very nice PDF published by Dennis Yurichev with introduction into Reverse Engineering, assembly and some more advanced topics: http://yurichev.com/non-wiki-files/RE_for_beginners-en.pdf

Other interesting/necessary topics

Memory addressing

- Segmentation/Paging/Virtualization
- Process vs. Kernel space and addresses; **DLL/process injection**

Executable formats

- Legacy DOS formats (DOS, EXE)
- Portable Executable (PE), symbol imports, DLL loading

Non-documented instructions/behavior

- Intel vs. AMD; Virtual Machines

Calling conventions

- cdecl
- stdcall (most common on Windows)
- fastcall

Anti-debugging tricks

- Exceptions, interrupts, time checking, debuggers detection, PE header manipulation, PEB manipulation…

Anti-VM tricks

- looking for uncommon sequences and/or behavior (CPUID instruction, e.g.); BIOS analysis
- Windows registry keys presence

Code obfuscation/packing

- Virtualization (own instruction interpreters)
- UPX/ASPack – the most famous packers
- Usually combined with the previous techniques

Getting current IP (instructions pointer)

- near/far calls or jumps, useful for shellcodes

Win32 API Calls

- API hash strings, trampolines, … (to prevent detection)

Crypto Algorithms

- xor, aes, rc4, rsa

Data Execution Prevention & others

- DEP (Data Execution Prevention)
- ASLR (Address Space Layout Randomization)
- The previous techniques can (of course) be bypassed: ROP (Return Oriented Programming)

Images

All the images were taken from Corkami Project webpage (https://code.google.com/p/corkami/).

Index of images:

COM 101 a DOS executable walk-through

VERSION 1.0
26 MARCH 2013
CREATIVE COMMONS 3.0 BY

ANGE ALBERTINI
CORKAMI.COM

DISSECTED FILE

SHA-1: 396E4D1778B1DDE6F5C78C14FE2852DC53CB80A3
DOWNLOAD a PE:101.CORKAMI.COM

```
C:\>SIMPLE.COM
Hello World!
```

```
0E 1F BA 0E 01 B4 09 CD 21 B8 01 4C CD 21 48 65 ........!..L.!He
6C 6C 6F 20 57 6F 72 6C 64 21 0D 0D 0A 24       llo.World!...$
```

SIMPLE.COM

TRIVIA

COM FILES ARE USED BY MICROSOFT SINCE 1981
AND STILL RUN ON ANY 32-BIT WINDOWS TODAY

CODE

WHAT'S EXECUTED

```
Offset:    0000
Address: cs:0100
0E 1F BA 0E 01 B4 09 CD 21 B8 01 4C CD 21 .. ..  ........!..L.!..
```

X86 16BITS ASSEMBLY*

```
push CS
pop  DS

1 mov  DX, msg
  mov  AH, 9
  int  0x21

2 mov  AX, 0x4C01
  int  0x21
```

EQUIVALENT C CODE

```
// copy CODE segment to DATA segment

print("Hello World!\r\n");

return 1;
```

*THE DOS CODE OF A STANDARD PORTABLE EXECUTABLE FILE IS IDENTICAL

DATA

INFORMATION USED BY THE CODE

```
Offset:    000E
Address: cs:010E
.. .. .. .. .. .. .. .. .. .. .. .. .. .. 48 65  ..............He
6C 6C 6F 20 57 6F 72 6C 64 21 0D 0D 0A 24        llo.World!...$
```

STRING

```
"Hello World!\r\n" $-terminated
```

EXPLANATIONS ARE SIMPLIFIED, FOR CONCISENESS

KEY CONCEPTS

FILE SPECIFICATIONS

NO HEADER
LIMITED TO ≈64KiB
MERGED CODE AND DATA
USUALLY WITH NO BOUNDARIES

MEMORY SEGMENTATION

DOS MEMORY IS CUT INTO BLOCKS
THEY ARE CALLED SEGMENTS
ADDRESSES ARE DEFINED BY:
A SEGMENT
AN OFFSET WITHIN THAT SEGMENT
THEY ARE WRITTEN `segment:offset`

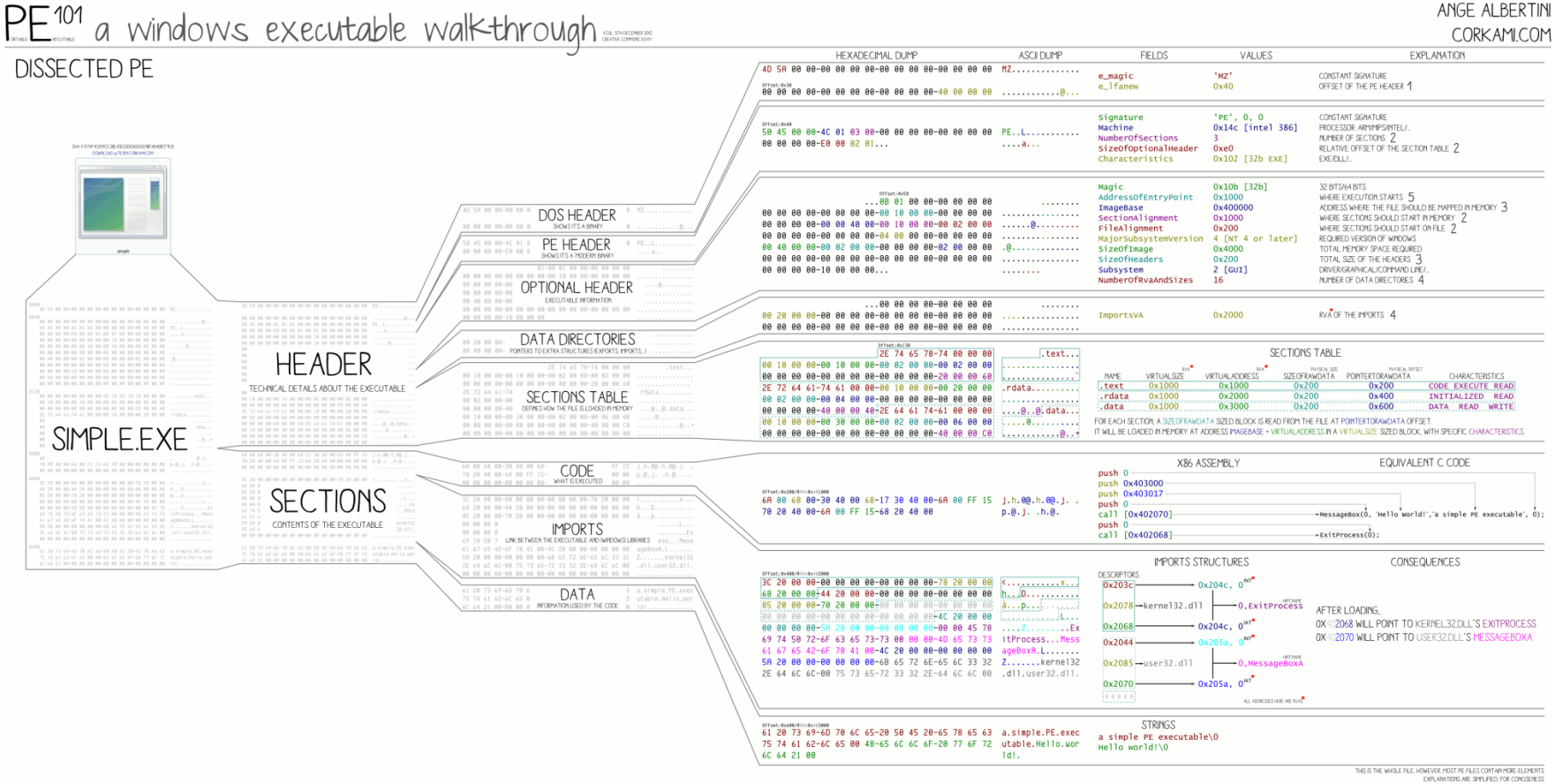
LOADING PROCESS

THE FILE IS ENTIRELY LOADED
OFFSET `0` IS MAPPED AT `cs:0x100`
EXECUTION STARTS THERE

OPERATING SYSTEM'S FUNCTIONS

CALLED VIA INTERRUPTS + FUNCTION CODE IN `AH`
EXAMPLES:

1. INTERRUPT `0x21` FUNCTION `0x09`
   PRINTS A $-TERMINATED STRING AT `DS:DX`
2. INTERRUPT `0x21` FUNCTION `0x4C`
   TERMINATES THE PROCESS
   RETURN VALUE GIVEN IN `AL`


PE 101 a windows executable walk-through
ANGE ALBERTINI
CORKAMI.COM
DISSECTED PE
HEXADECIMAL DUMP
ASCII DUMP
FIELDS
VALUES
EXPLANATION
DOS HEADER
PE HEADER
OPTIONAL HEADER
DATA DIRECTORIES
SECTIONS TABLE
HEADER
SIMPLE.EXE
SECTIONS
CODE
IMPORTS
DATA
SECTIONS TABLE
X86 ASSEMBLY
EQUIVALENT C CODE
IMPORTS STRUCTURES
CONSEQUENCES
STRINGS
LOADING PROCESS
1 HEADERS
2 SECTIONS TABLE
3 MAPPING
4 IMPORTS
5 EXECUTION
NOTES

In-Memory Analysis

Even though doing an advanced reverse engineering could be life changing experience, analyzing malware in this depth usually needs some good clarification. In many security teams, one of the first steps in incident classification is a triage.

For that reason, in-memory analysis of a running malware might be beneficial.

However, there are many other good reasons why a security engineer should do in-memory analysis first. You should use it when…

- Doing a rapid threat assessment – very efficient method.
- Infected host is online and available for the analysis, not restarted yet.
- There is a chance that the original binary is gone (transient infections).
- There was no original binary stored on the infected host.
- You cannot read JavaScript/Java or any other exploited application's code and vulnerability triggered.
- Messing with packers and obfuscated code can be annoying. Indeed.
- Even though you can pass many defenses put in place by attackers, after many hours you can find one that cannot be broken because of missing DLL/configuration file.
- Some data are being exfiltrated memory can contain important evidence – attackers steps, passwords, tools used, …
- …

Recommended Tools

There are some publicly available tools that can be used for memory analysis. They differ in capabilities, supported systems and licensing.

1. Volatility Framework
2. HBGary Responder
3. Mandiant Redline
4. F-Response
5. …
6. grep/strings :)

My preferred tools are Volatility and HBGary Responder (Pro Edition). The first one is great for its broad scope of system that can be analyzed, the second for graphical debugger.

Memory Acquisition Tools and Techniques (Windows OS)

1. Virtual machine memory dump
    - Not applicable for many hosts (laptops, servers).
    - Super useful for malware analysis when the malware doesn't do any anti-VM tricks.
        - Don't forget to configure as little memory as possible for the running system; it will significantly speed up your analysis.
    - VirtualBox, VMWare can do this, VMWare more convenient.
2. FastDump (Pro)
    - HBGary solution, small footprint, one of the best tools available.
    - Cannot be obtained easily, official pages don't work well; Pro version is expensive.
    - **Can acquire memory that is currently swapped-out!**
        - This can be critical for non-VM systems!
3. Memoryze
    - Free tool by Mandiant.
    - Quite big footprint, XML files, installer.
4. Win32dd.exe
    - Not available anymore, replaced by MoonSols Windows Memory Toolkit.
5. MoonSols Windows Memory Toolkit
    - Community Edition available.
    - Single binary, small footprint.
6. Forensic tools (EnCase, Mandiant, Access Data, …)
    - Remote acquisitions (over the network), compressed images, using already-installed drivers, thus no tampering with the system memory.
    - Corporate tools -- inhuman expensive ;)

Memory Analysis Tools

Mandiant Redline

- http://www.mandiant.com/resources/download/redline
- For free, available for Windows XP, Vista and 7 (32-bit and 64-bit).
- Malware risk scoring index.
    - Helps to assess system quickly.
    - List with system and well-known good apps, suspicious mutexes, etc.

HBGary Responder (CE/Pro)

- http://www.hbgary.com/hbgary-releases-responder-ce
- Community Edition is available against registration.
- Available for Windows XP, Vista and 7.
- Nothing really awesome unless you use Pro Edition.
    - Simple disassembler with graphing features, priceless.
    - Digital DNA – very good process/memory scoring system.
- Can be extended by C# plug-ins.

Volatility Framework

- https://code.google.com/p/volatility/
- Open source!
- Capable of memory analysis of Windows, Linux (Android) and MacOS systems.
- Extensible, written in Python.
- No GUI yet.

What Can Be Found in Memory?

Almost everything!

- **Malware :) (rootkits included)**
- System information
  - Hardware and software
- Processes and threads
  - Loaded DLLs
- Network sockets, URLs, IP addresses
- Open files (and pipes)
- Mutexes/Handles
- User-generated content
  - Passwords, clipboards, caches
- Encryption keys!
  - E.g. for TrueCrypt (can be automated)
- Registry hives
- Event logs
- *(Screen preview!)*

Also, you can search for system inconsistencies – hidden processes, hidden drivers, non-system handlers (interrupts handled by non-system processes).

What to Search For?

- IRP (**I/O Request Packets)**
  - Mostly focused at NTFS, DISK, FAT, TCPIP, NDIS and KBDCLASS drivers.
  - Look for a single hook IRP_MJ_DEVICE_CONTROL.
  - Use your brain and Google.
- SSDT (**System Service Dispatch Table)**
  - Used by legit systems (HIPS, malware protection) and malware.
  - Differentiating is not easy; use your brain and Google.
    - Hooking to unresolved drivers is suspicious.
    - Unsigned drivers are suspicious.
    - Only a few hooks in place can be suspicious.
- IDT (**Interrupt Descriptor Table**)
  - It allows attackers to subvert memory manager, keyboard events, system calls.
  - Not much used, any IDT hook is suspicious :)
- Hidden processes, DLLs, drivers, …
- Process injections
- Process path/user inconsistencies
  - E.g. `svchost.exe` executed from `c:\windows\sytem32\dllcache\`
  - Running as a non-standard user.
- Open sockets, network connections
- Mutexes/Handles
- URLs (URL-like strings)
  - URL-like strings can also be interesting! (e.g. `http://%d%S/config.html`)
- Anything suspicious!
  - **Malware can be digitally signed!**

Memory Injection

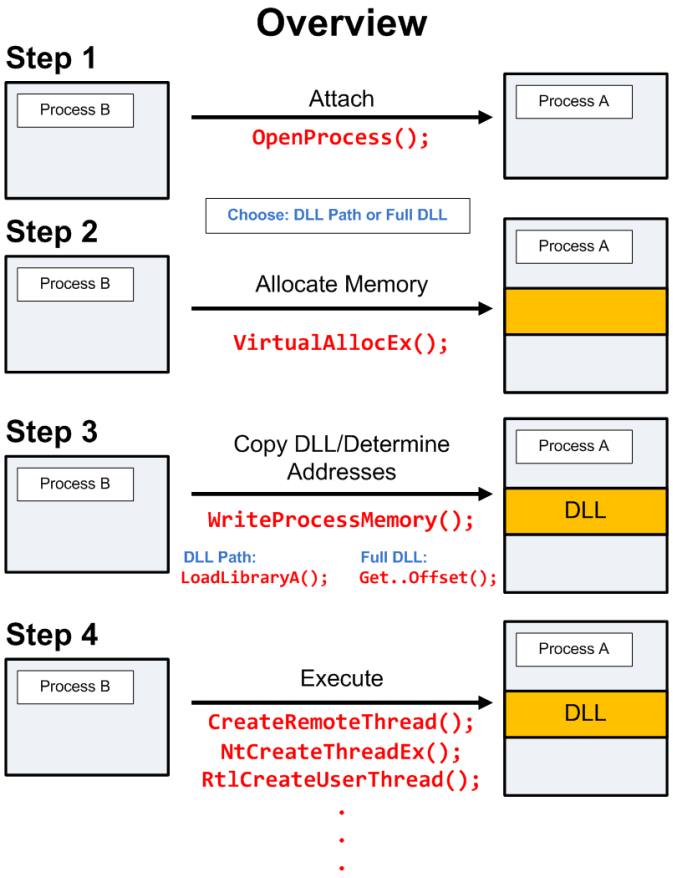
What is it and why is it so important?

A very simple definition: a mechanism of inserting dynamic library / malicious code in the process of confidence.

Why? After successful injection the malware can use all the benefits of the original process. Thus, if the malware injected the process of Internet Explorer, it can now bypass Windows Firewall and run its code in any port. With some care, malware can also spy on the original process; can re-define some of the original functions and/or event handlers.

Technical details and possible ways how the code can be injected into a running process can be found at:

1. http://resources.infosecinstitute.com/code-injection-techniques/
2. http://www.codeproject.com/Articles/4610/Three-Ways-to-Inject-Your-Code-into-Another-Proces
3. http://blog.opensecurityresearch.com/2013/01/windows-dll-injection-basics.html (source of the following image)



Well-known Suspicious Mutexes

Virus/Tool Name	Mutex
Conficker	`.*-7` and `.*-99`
Sality.AA	`Op1mutx9`
Flystud.??	`Hacker.com.cn_MUTEX`
NetSky	`'D'r'o'p'p'e'd'S'k'y'N'e't'` `YY99knPY`
Sality.W	`u_joker_v3.06`
Poison Ivy	`)!VoqA.I4`
koobface	`35fsdfsdfgfd5339`

Expected Paths

Process	Expected Paths
lsass.exe	`\windows\system32`
services.exe	`\windows\system32`
csrss.exe	`\windows\system32`
explorer.exe	`\windows`
spoolsv.exe	`\windows\system32`
smss.exe	`\windows\system32`
svchost.exe	`\windows\system32`
iexplore.exe	`\program files` `\program files (x86)`
winlogon.exe	`\windows\system32`

Suspicious Imports

Scenario #1

- `GetProcAddress`
- `LoadLibrary`
  - What the binary can do with these calls?
  - What if the binary doesn't contain any other import?

Scenario #2

- `CreateToolhelp32Snapshot`
- `Process32Next`
- `Process32First`
  - What the binary/process can do with these calls?

Scenario #3

- `Ws2_32.dll / msock32.dll`
- `Wininet.dll`
- `Netapi32.dll`
  - Network-related imports; can they be present in `calc.exe`? And in `svchost.exe`?

Volatility Cheat Sheet

Source: https://blogs.sans.org/computer-forensics/files/2012/04/Memory-Forensics-Cheat-Sheet-v1_2.pdf

- `vol.py -h / vol.py plugin -h / vol.py plugin --info` (help)
- `vol.py -f image.file imageinfo` (info about the image, useful for further steps)
- `vol.py -f image.file --profile=profile plugin` (sample command line)
  - `export VOLATILITY_LOCATION=image.file`
  - `export VOLATILITY_PROFILE=WinXPSP3x86`
    - can save you some typing
- `vol.py -f image.file --profile=profile <plugin>`
  - `psxview` (look for hidden processes)
  - `apihooks`
  - `driverscan`
  - `ssdt / driverirp / idt`
  - `connections / connscan` (WinXP, list of open TCP connections / all TCP connections)
  - `netscan` (Win7, scan for connections and sockets)
  - `pslist / psscan` (high-level process list vs. scan for EPROCESS blocks)
  - `malfind / ldrmodules` (find injected code, dump sections / detect unlinked DLLs)
  - `hivelist` (find and list available registry hives) / `hashdump`
  - `handles / dlllist / filescan` (list of open handles / DLL files / FILE_OBJECT handles)
  - `cmdscan / consoles` (find the history of cmd.exe / console buffer)
  - `shimcache` (application compatibility info)
  - `memdump / procmemdump / procexedump`

Homework

The task is simple. Analyze the given memory image, find all irregularities there, prepare a formal report and document your findings. The more details and information you provide, the more points you get.

Please note that you have to present me not only your results, but also the tools you used in the whole process described in your report. If you use Redline for one part of the analysis and Volatility for another please document it. I have to be able to follow your investigation in order to validate the steps you took.

File: homework.zip

Hints:

- should be similar to one of the lecturing samples,
- watch for suspicious connections,
- do not forget to check execution times,
- find suspicious URLs, if there are any,
- can you find the source of the infection? And any technical details? (Google is your friend)

References and further reading

1. https://code.google.com/p/volatility/wiki/PublicMemoryImages
2. https://www.mandiant.com/blog/precalculated-string-hashes-reverse-engineering-shellcode/
3. https://github.com/iagox86/nbtool/blob/master/samples/shellcode-win32/hash.py
4. http://blog.spiderlabs.com/2013/04/basic-packers-easy-as-pie.html (simple unpacking, not UPX)
5. http://zeltser.com/remnux/ -- REMnux. All you need to do a reverse engineering.
6. http://www.deer-run.com/~hal/Detect_Malware_w_Memory_Forensics.pdf
7. http://downloads.ninjacon.net/downloads/proceedings/2011/Michael_J_Graven-Finding_Evil_in_Live_Memory.pdf
8. http://www.skullsecurity.org/blog/2013/ropasaurusrex-a-primer-on-return-oriented-programming (Practical introduction into Return Oriented Programming)