

恶意代码分析与防治技术 第2章 基本静态分析技术

王志

zwang@nankai.edu.cn

updated on 2022-09-25

南开大学 网络空间安全学院 2022-2023学年



#### **NSA TAO**

• 美国国家安全局(NSA)"特定入侵行动办公室"(TAO)









# 网络攻击事件

- 2022年6月22日,西北工业大学发布《公开声明》称,该校遭受境外网络攻击。
- 国家计算机病毒应急处理中心和360公司从多个信息系统和上网终端中提取到了多款木马样本。
- 全面还原了相关攻击事件的总体概貌、技术特征、攻击武器、攻击路径和攻击源头。









#### 新型冠状病毒 与网络安全病毒**特点分析**

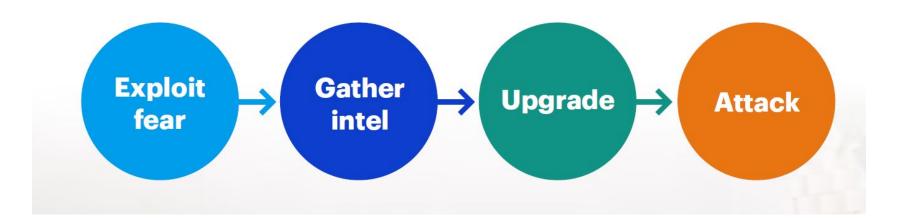
a Stelles

新型冠状 病毒	病毒名称	永恒之蓝(WannaCry) 勒索病毒
野生动物-中华菊头蝠	传染源	影子经纪人一利用美军NSA 网络攻击武器库工具
人	感染对象	服务器终端
经呼吸道飞沫传播 亦可通过接触传播 存在粪-口传播可能性	传播途径	利用MS17-010漏洞攻击微软 SMB服务, 通过139和445端口感染
人传人、人通过路网 交通,各地爆发	传染路线	机器之间互相传染 通过网络复制传播
发热、乏力、干咳, 逐渐呼吸困难	症状	文件全部加密,变更文件名, 桌面背景包含勒索语言和支付方式
传染强度高于非典	传染强度	传染强度大,一个机器中招, 几个小时同网全部机器被加密
严重, 致死率目前在2~3%左右	严重性	严重, 最核心数据全部被加密
人群普遍易感	易感群体	系统版本较老,未升级MS17-010 补丁的,开放139、445端口的, 后期主要感染非互联网系统
尚无	特效药物	尚无很好的手段 被加密后破解难度很大
3~7天,最长14天	潜伏期	永恒之蓝演变成潜伏挖矿病毒





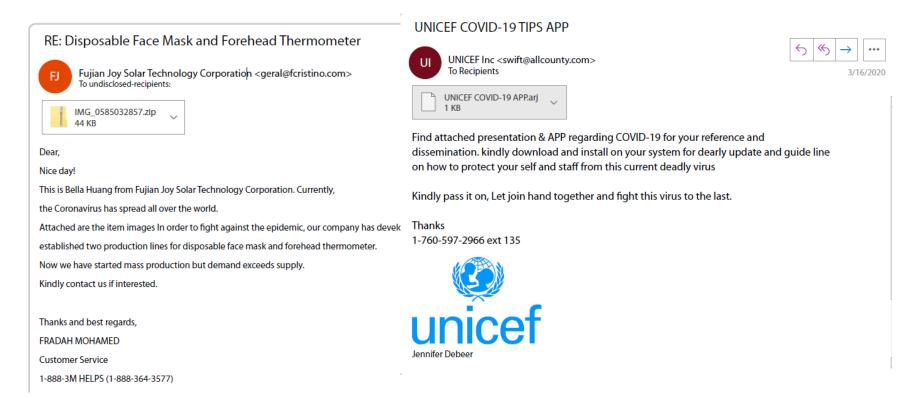
## COVID-19 changed the threat landscape







# Exploit fear



In April, Google reported it was blocking 18 million spam emails related to COVID-19 per day!





#### 允公允铭日新月异

## Malware as a business

- Malware DistributionServices
  - On November 23, GootKit
     pushing the Revil ransomware
     to machines only in Germany.







# 本章知识点

- 1. Antivirus Scanning
- 2. Hashes
- 3. Finding Strings
- 4. Packed and Obfuscated Malware
- 5. Portable Executable File Format
- 6. Linked Libraries and Functions
- 7. Dependency Walker
- 8. The PE File Headers and Sections





Basic Static Analysis



# Static Analysis

- Reverse engineering the code or structure of a binary executable to understand its functionality.
- Static analysis:
  - The program is not run at this time.







## Basic Static Analysis

- No disassembly
- Provides good pointers to guide dynamic and advanced analysis
- Lots of **tools** involved!







## Techniques

- Antivirus scanning
- Hashes
- A file's strings, functions, and headers





Antivirus Scanning



# 杀毒软件











































































## 允公允帐日新月异

# Antivirus Scanning

- Known malware
  - File signatures
  - Heuristics
- Unknown malware
  - Obfuscation
  - Polymorphic: syntax obfuscation
  - Metamorphic: semantic obfuscation





## Collection of Antivirus Tools

VirusTotal is a free service that analyzes sus the quick detection of viruses, worms, trojans	•
No file selected	Choose File
Maximum file siz	ze: 64MB





# Zvirustotal

VirusTotal is a free service that **analyzes suspicious files and URLs** and facilitates the quick detection of viruses, worms, trojans, and all kinds of malware.

File	<b>W</b> URL	Q Search	
7			Choose File

By clicking 'Scan it!', you consent to our Terms of Service and allow VirusTotal to share this file with the security community. See our Privacy Policy for details.

Scan it!





## virustotal

SHA256: 7a371cc054ee14f0614b90cd6797001b5fd18c70c45c463f9f1a161ba08498ec

File name: 1eaee1e2-fad8-4ef0-b20f-1863e02100e1.doc

Detection ratio: 0 / 55

Analysis date: 2017-02-15 04:46:13 UTC ( 0 minutes ago )

0 0 0

Analysis

♠ File detail

Additional information

Comments

√ Votes

Antivirus	Result	Update
ALYac	•	20170215
AVG	•	20170215
AVware	•	20170215
Ad-Aware	•	20170215





Hashing



# Hashing

- Method to uniquely identify malware
- MD5
  - Message-Digest Algorithm
  - 128-bit hash
- SHA1
  - Secure Hash Algorithm
  - 160-bit hash





## Hashes

- Input:
  - A file or string with arbitrary length
- Output:
  - fixed-length hash
- Uniquely identifies a file well in practice
  - MD5 collisions but they are not common
  - Collision: two different files with the same hash





## 允公允帐日新月异

## HashCalc

H HashCalc		X
Data Format: File ▼	Data: C:\Users\student\Desktop\p3.pcap	- III
☐ HMAC	Key Format: Key:	
<b>✓</b> MD5	52583b5e2c99d19c046915181fd7b29b	
☐ MD4		
✓ SHA1	991d4e880832dd6aaebadb8040798a6b9f163194	
☐ SHA256		

南阁大學

**Nankai University** 





## Hash Uses

- Label a malware file
- Share the hash with other analysts to identify malware
- Search the hash online to see if someone else has already identified the file





Finding Strings



# Strings

- String: a sequence of printable characters.
- Computer can only understand 0 and 1
- Use 0 and 1 to represent characters
  - ASCII
  - UNICODE





# Strings

#### • ASCII

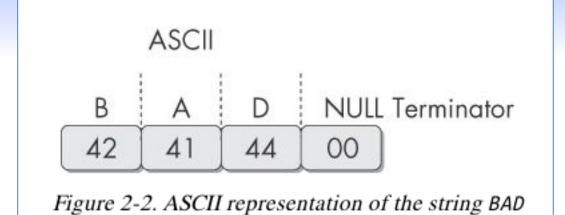
- American Standard Code for Information Interchange
- 8 bits long

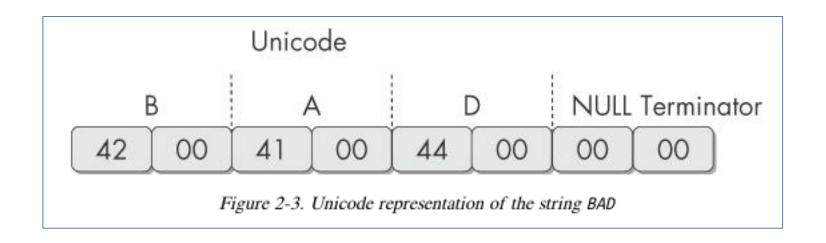
#### • UNICODE

- Universal Coded Character Set
- 135 modern or historic scripts
- 16 bits long















# The strings Command

- Search binary executable for ASCII and Unicode strings
- Three or greater sequence of characters
- Followed by a terminator





## The strings Command

```
C:>strings bp6.ex_
VP3
VW3
t$@
D$4
99.124.22.1 4
e-@
GetLayout 1
GDI32.DLL 3
SetLayout 2
M}C
Mail system DLL is invalid.!Send Mail failed to
send message. 5
```





Packed and Obfuscated Malware



#### Packed and Obfuscated Malware

- Goals: Make malware more difficult to reverse engineering and detect
- Obfuscation: conceal execution information
- Packer: compress the size of binary file
  - a subset of obfuscation







## Packer and Obfuscation

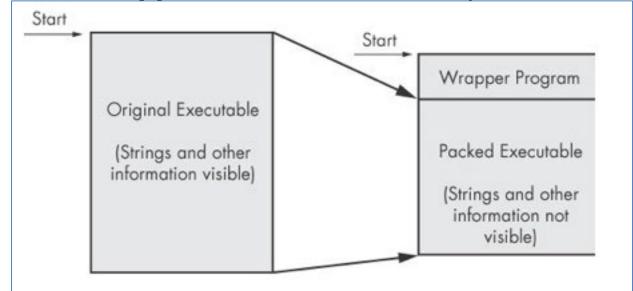
- Legitimate Program
  - Many strings
- Packed or Obfuscated Malware
  - few strings





## Packing Files

- The code is compressed, like a Zip or RAR file
- This makes the strings and instructions unreadable
- All you'll see is the wrapper small code that unpacks the file when it is running







## Demo: UPX

```
root@kali: ~/126
File Edit View Search Terminal Help
root@kali:~/126# cat chatty.c
#include <stdio.h>
main()
char name[10];
printf("This program contains readable strings\n");
printf("Enter your name: ");
scanf("%s", name);
printf("Hello %s\n", name);
root@kali:~/126# gcc -static chatty.c -o chatty
root@kali:~/126# upx -o chatty-packed chatty
                       Ultimate Packer for eXecutables
                          Copyright (C) 1996 - 2011
                Markus Oberhumer, Laszlo Molnar & John Reiser
UPX 3.08
                                                                Dec 12th 2011
        File size
                          Ratio
                                     Format
                                                 Name
    592800 ->
                 272588
                          45.98% linux/elf386
                                                 chatty-packed
Packed 1 file.
root@kali:~/126# ls -l
total 852
 -rwxr-xr-x 1 root root 592800 Aug 16 20:34 chatty
 -rw-r--r-- 1 root root     174 Aug 16 20:27 chatty.c
 -rwxr-xr-x 1 root root 272588 Aug 16 20:34 chatty-packed
root@kali:~/126#
```





## 允公允帐日新月异

# Detecting Packers with PEiD

File: C:\ma	alware\orig_af2.ex_			[AAA]
Entrypoint:	0000EEA0	EP Section:	UPX1	>
File Offset:	000050A0	First Bytes:	60,BE,15,A0	>
Linker Info:	6.0	Subsystem:	Win32 console	>
UPX 0.89.6 Multi Scan ▼ Stay on t		arkus & Laszlo  Options Abo	ut E <u>×</u>	it

Figure 2-5. The PEiD program





# Packing Obfuscates Strings

```
root@kali:~/126# strings chatty | wc
   1962   4498   33817
root@kali:~/126# strings chatty-packed | wc
   3950   4290   23623
root@kali:~/126#
```





#### NOTE

Many PEiD plug-ins will run the malware executable without warning! (See Chapter 3 to learn how to set up a safe environment for running malware.) Also, like all programs, especially those used for malware analysis, PEiD can be subject to vulnerabilities. For example, PEiD version 0.92 contained a buffer overflow that allowed an attacker to execute arbitrary code. This would have allowed a clever malware writer to write a program to exploit the malware analyst's machine. Be sure to use the latest version of PEiD.





## Portable Executable File Format





## PE Files

- Portable Executable File Format
- Used by Windows executable files, and DLLs
- Contains the information necessary for Windows to load the binary executable
- Almost every file executed on Windows is in PE format







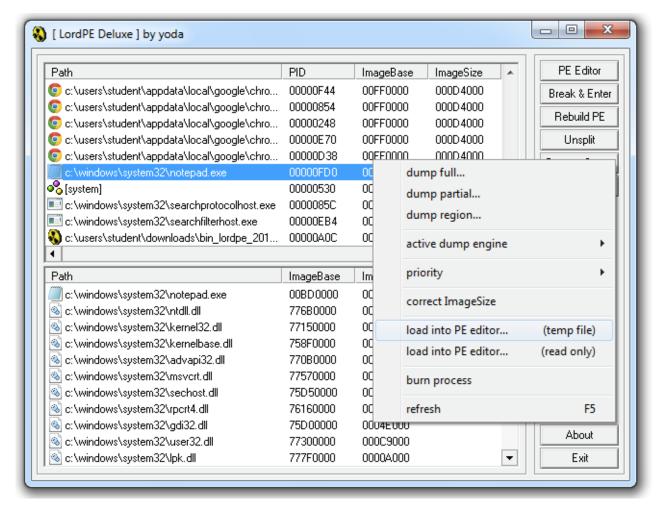
## PE Header

- Information about the code
- Type of application
- Required library functions
- Space requirements





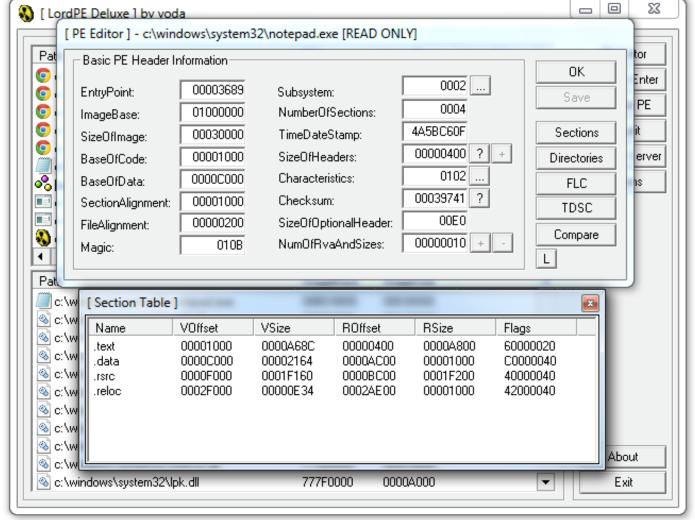
## LordPE Demo







## Main Sections









允公允继日新月异

There are a lot more sections

But the main ones are enough for now

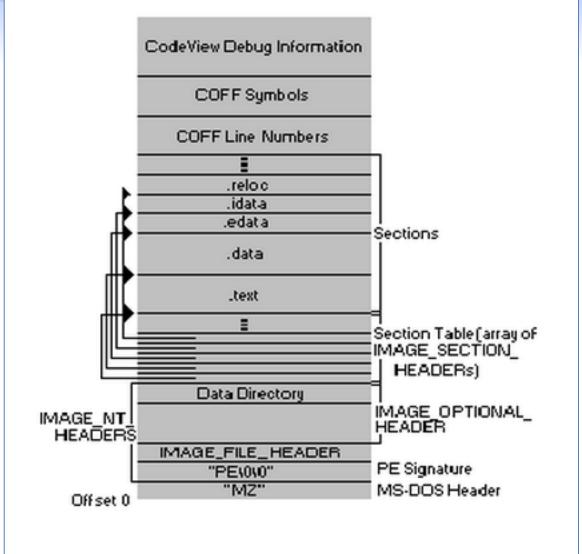


Figure 1. The PE file format







Linked Libraries and Functions





# **Imports**

- Functions used by a program that are stored in a different program, such as library
- Connected to the main EXE by Linking
- Can be linked three ways
  - Statically
  - At Runtime
  - Dynamically





# Static Linking

- Rarely used for Windows executables
- Common in Unix and Linux
- All code from the library is copied into the executable
- Bigger file size
- More memory space

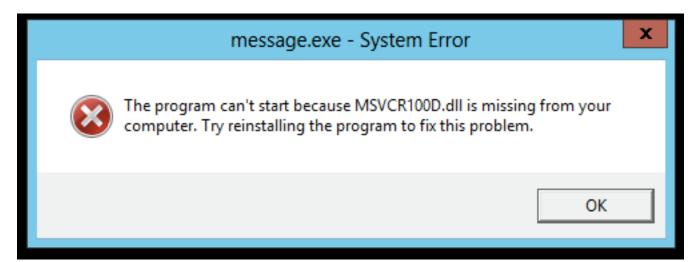






# Dynamic Linking

- Most common method
- Host OS searches for necessary libraries when the program is loaded







# Runtime Linking

- Unpopular in friendly programs
- Common in malware, especially packed or obfuscated malware
- Connect to libraries only when needed, not when the program starts
- Most commonly done with the **LoadLibrary** and **GetProcAddress** functions







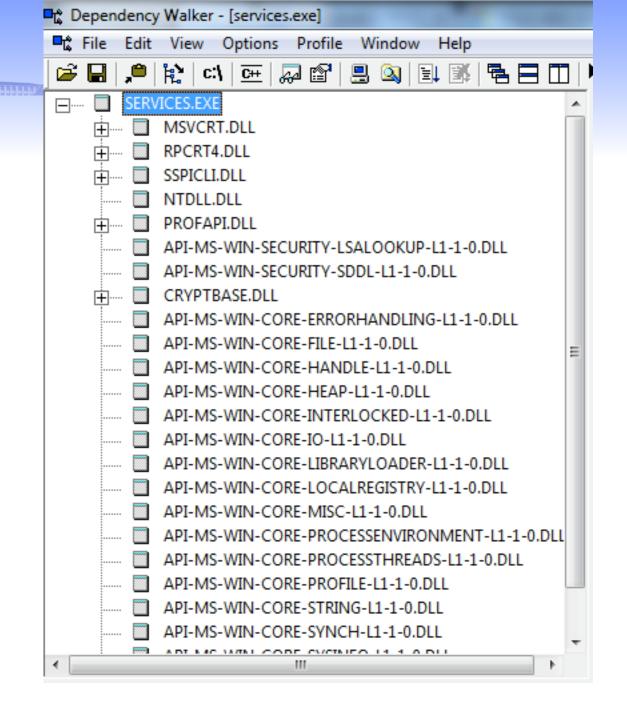
## Clues in Libraries

- The PE header lists every library and function that will be loaded
- Their names can reveal what the program does
- URLDownloadToFile indicates that the program downloads something

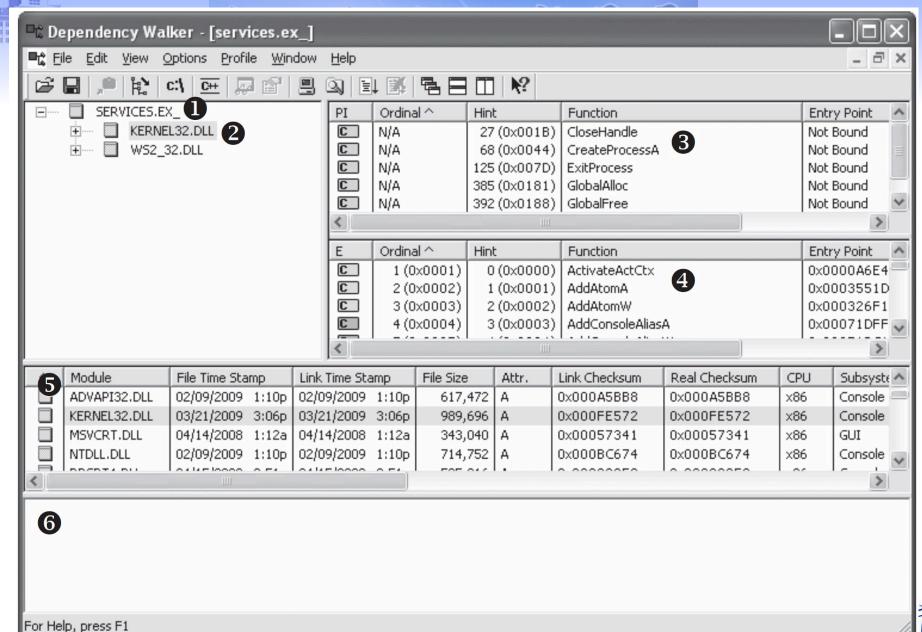




Dependency Walker







919

周大學



# Shows Dynamically Linked Functions

- Normal programs have a lot of DLLs
- Malware often has very few DLLs





#### Table 2-1. Common DLLs

DLL	Description
Kernel32.dll	This is a very common DLL that contains core functionality, such as access and manipulation of memory, files, and hardware.
Advapi32.dll	This DLL provides access to advanced core Windows components such as the Service Manager and Registry.
User32.dll	This DLL contains all the user-interface components, such as buttons, scroll bars, and components for controlling and responding to user actions.
Gdi32.dll	This DLL contains functions for displaying and manipulating graphics.





This DLL is the interface to the Windows kernel. Executables generally do not import this file directly, although it is always imported indirectly by *Kernel32.dll*. If an executable imports this file, it means that the author intended to use functionality not normally available to Windows programs. Some tasks, such as hiding functionality or manipulating processes, will use this interface.

WSock32.dll These are networking DLLs. A program that accesses and either of these most likely connects to a network or

Ws2\_32.dll performs network-related tasks.

Wininet.dll

This DLL contains higher-level networking functions that implement protocols such as FTP, HTTP, and NTP.







# **Exports**

- DLLs export functions
- EXEs import functions
- Both exports and imports are listed in the PE header





Kernel32.dll	User32.dll	User32.dll (continued)
CreateDirectoryW	BeginDeferWindowPos	ShowWindow
CreateFileW	CallNextHookEx	ToUnicodeEx
CreateThread	CreateDialogParamW	TrackPopupMenu
DeleteFileW	CreateWindowExW	TrackPopupMenuEx
ExitProcess	DefWindowProcW	TranslateMessage
FindClose	DialogBoxParamW	UnhookWindowsHookEx
FindFirstFileW	EndDialog	UnregisterClassW
FindNextFileW	GetMessageW	UnregisterHotKey
GetCommandLineW	GetSystemMetrics	
GetCurrentProcess	GetWindowLongW	GDI32.dll
${\tt GetCurrentThread}$	GetWindowRect	GetStockObject
GetFileSize	GetWindowTextW	SetBkMode
GetModuleHandleW	InvalidateRect	SetTextColor
GetProcessHeap	IsDlgButtonChecked	
GetShortPathNameW	IsWindowEnabled	Shell32.dll
HeapAlloc	LoadCursorW	CommandLineToArgvW
HeapFree	LoadIconW	SHChangeNotify
IsDebuggerPresent	LoadMenuW	SHGetFolderPathW
MapViewOfFile	MapVirtualKeyW	ShellExecuteExW
OpenProcess	MapWindowPoints	ShellExecuteW
ReadFile	MessageBoxW	
SetFilePointer	RegisterClassExW	Advapi32.dll
WriteFile	RegisterHotKey	RegCloseKey
	SendMessageA	RegDeleteValueW
	SetClipboardData	RegOpenCurrentUser
	SetDlgItemTextW	RegOpenKeyExW
	SetWindowTextW	RegQueryValueExW
	SetWindowsHookExW	RegSetValueExW





# Ex: A Packed Program

Table 2-3. DLLs and Functions Imported from PackedProgram.exe

Kernel32.dll User32.dll

GetModuleHandleA MessageBoxA

LoadLibraryA

GetProcAddress

ExitProcess

VirtualAlloc

VirtualFree





The PE File Headers and Sections



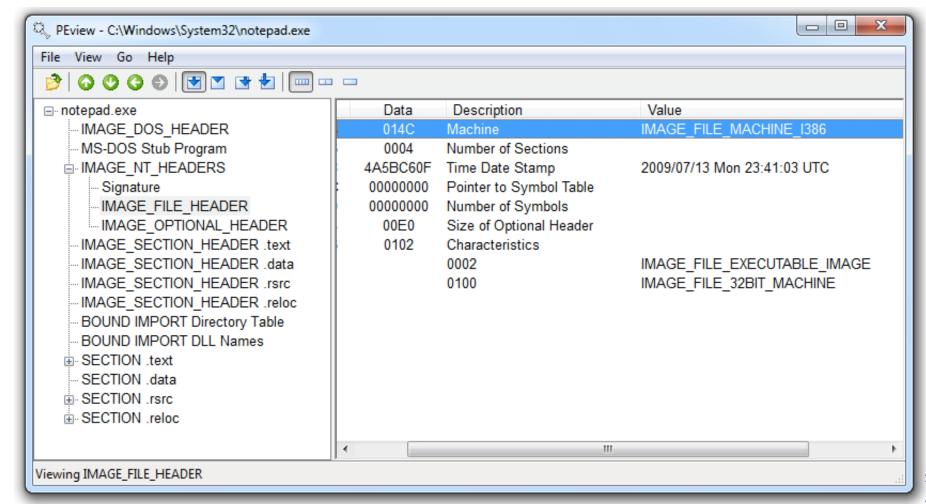
# Important PE Sections

- .text -- instructions for the CPU to execute
- .rdata -- imports & exports
- .data global data
- .rsrc strings, icons, images, menus





## **PEView**





## 允公允帐日新月异

# Time Date Stamp

- Shows when this executable was compiled
- Older programs are more likely to be known to antivirus software
- But sometimes the date is wrong
  - All Delphi programs show June 19, 1992
  - Date can also be faked





# IMAGE\_SECTION\_HEADER

- Virtual Size RAM
- Size of Raw Data DISK
- For .text section, normally equal, or nearly equal
- Packed executables show Virtual Size much larger than Size of Raw Data for .text section





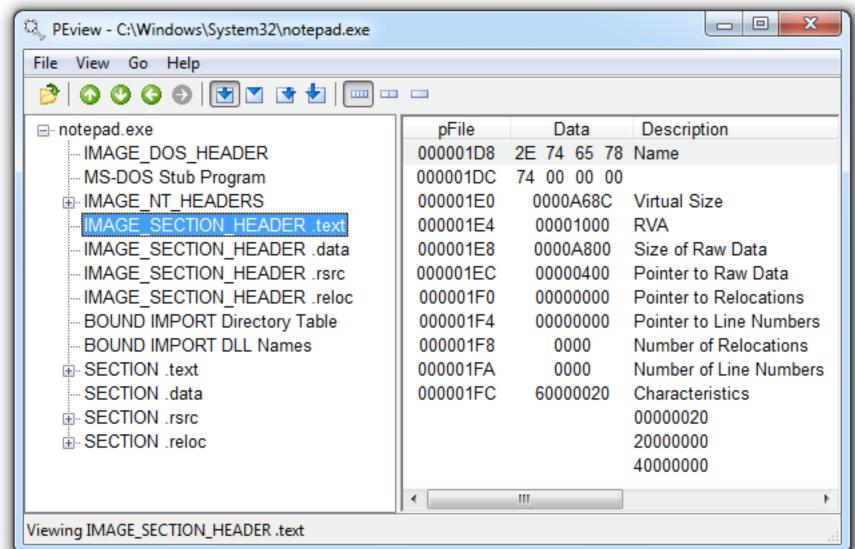






Table 2-6. Section Information for PackedProgram.exe

Name	Virtual size	Size of raw data
.text	A000	0000
.data	3000	0000
.rdata	4000	0000
.rsrc	19000	3400
Dijfpds	20000	0000
.sdfuok	34000	3313F
Kijijl	1000	0200







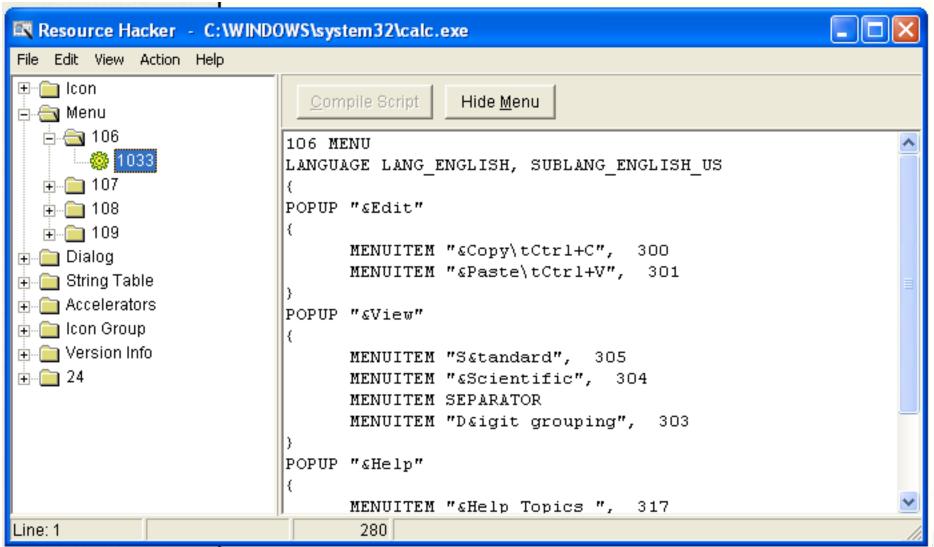
## Resource Hacker

- Lets you browse the .rsrc section
- Strings, icons, and menus





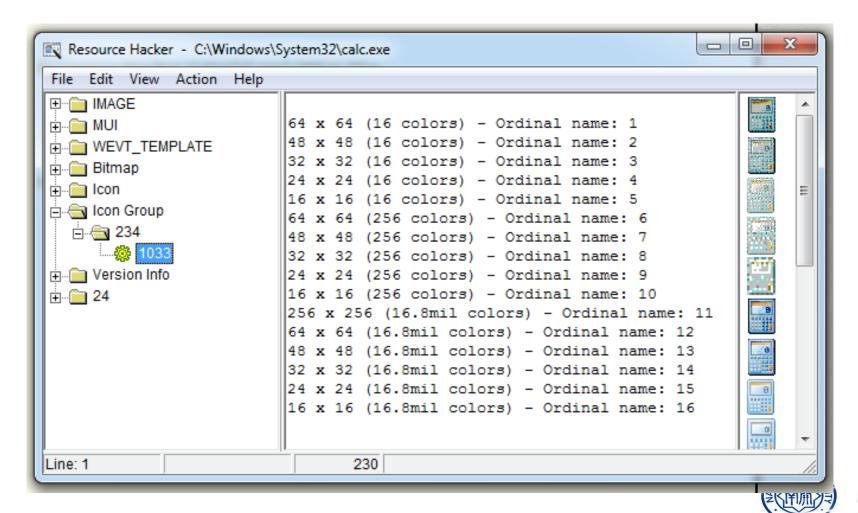
## Resource Hacker in Windows XP







## Resource Hacker in Windows 7









## Labs

- Practice our skills
- In order to simulate realistic malware analysis, little or no information about the program is given.
  - generic names
  - meaningless or misleading names







## Labs

- Each lab consists
  - a malicious file
  - a few questions
  - short answers to the questions
  - a detailed analysis of the malware
- The solutions are included in Appendix C





This lab uses the files *Lab01-01.exe* and *Lab01-01.dll*. Use the tools and techniques described in the chapter to gain information about the files and answer the questions below.

- 1. Upload the files to *http://www.VirusTotal.com/* and view the reports. Does either file match any existing antivirus signatures?
- 2. When were these files compiled?
- 3. Are there any indications that either of these files is packed or obfuscated? If so, what are these indicators?
- 4. Do any imports hint at what this malware does? If so, which imports are they?
- 5. Are there any other files or host-based indicators that you could look for on infected systems?
- 6. What network-based indicators could be used to find this malware on infected machines?
- 7. What would you guess is the purpose of these files?





## Lab 1-2

- 1. Upload the *Lab01-02.exe* file to *http://www.VirusTotal.com/*. Does it match any existing antivirus definitions?
- 2. Are there any indications that this file is packed or obfuscated? If so, what are these indicators? If the file is packed, unpack it if possible.
- 3. Do any imports hint at this program's functionality? If so, which imports are they and what do they tell you?
- 4. What host- or network-based indicators could be used to identify this malware on infected machines?





## Lab 1-3

Analyze the file *Lab01-03.exe*.

- 1. Upload the *Lab01-03.exe* file to *http://www.VirusTotal.com/*. Does it match any existing antivirus definitions?
- 2. Are there any indications that this file is packed or obfuscated? If so, what are these indicators? If the file is packed, unpack it if possible.
- 3. Do any imports hint at this program's functionality? If so, which imports are they and what do they tell you?
- 4. What host- or network-based indicators could be used to identify this malware on infected machines?





## Lab 1-4

Analyze the file *Lab01-04.exe*.

- 1. Upload the *Lab01-04.exe* file to *http://www.VirusTotal.com/*. Does it match any existing antivirus definitions?
- 2. Are there any indications that this file is packed or obfuscated? If so, what are these indicators? If the file is packed, unpack it if possible.
- 3. When was this program compiled?
- 4. Do any imports hint at this program's functionality? If so, which imports are they and what do they tell you?
- 5. What host- or network-based indicators could be used to identify this malware on infected machines?
- 6. This file has one resource in the resource section. Use Resource Hacker to examine that resource, and then use it to extract the resource. What can you learn from the resource?





# 实验报告提交

•实验报告以附件的形式在雨课堂上提交。





恶意代码分析与防治技术 第2章 基本静态分析技术

王志

zwang@nankai.edu.cn

updated on 2022-09-25

南开大学 网络空间安全学院 2022-2023学年