# Malware

#### **CS-576 Systems Security**

Instructor: Georgios Portokalidis

Spring 2018

### Malware

#### **Sample definition**

"a program that is inserted into a system, usually covertly, with the intent of compromising the confidentiality, integrity, or availability of the victim's data, applications, or operating system or otherwise annoying or disrupting the victim."



### **Evil Takes Many Forms**

Viruses	Dialers
Worms	Droppers
Rootkits	Spyware
Keyloggers	Adware
Trojan Horses	
Ransomware	Backdoors

### Main Classification

#### **Infection vector**

The type of vulnerability the malware exploits to infect a host

> Software vulnerability, download, design flaw, social engineering, ...

The method the malware uses to propagate

 Disks, USB, network, website, ad, ...

#### Payload

The actions the malware takes after infecting

 DDoS, encrypting disk, stealing data, backdoor,

### Also Classified By

Code of malware

Binary, interpreted (JS, VB), macros, ...

Infection point

File, boot sector, firmware, memory-only, BIOS, ...

Activation

Through user interaction, automatic, mixed

#### **Evasion strategies**

Packing, polymorphism, obfuscation, anti-VM/debugging, ...

### **Infection Vectors**

### Installed by User

🗅 New tab	×	▼ 2 5:00
$\leftarrow \rightarrow G$	Search or type URL	☆ 🎙 :

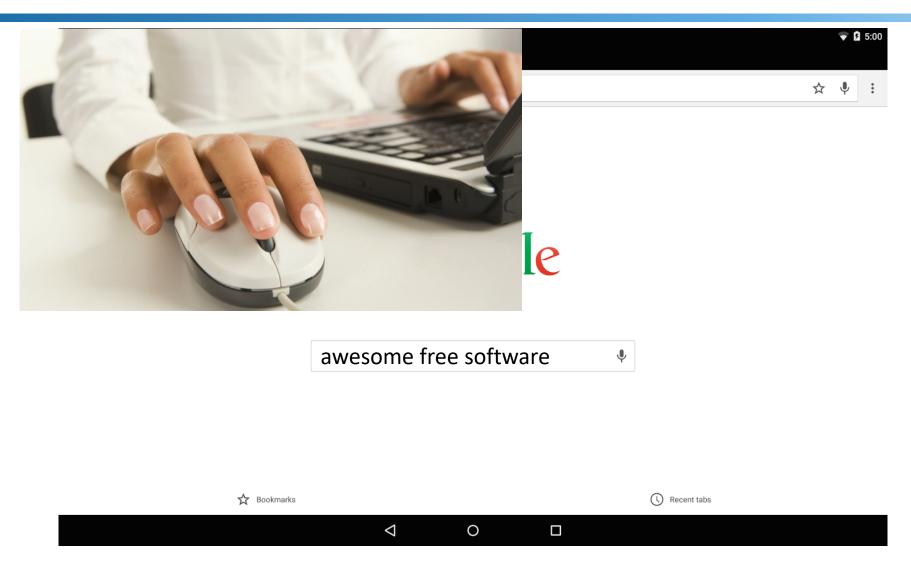


awesome free software

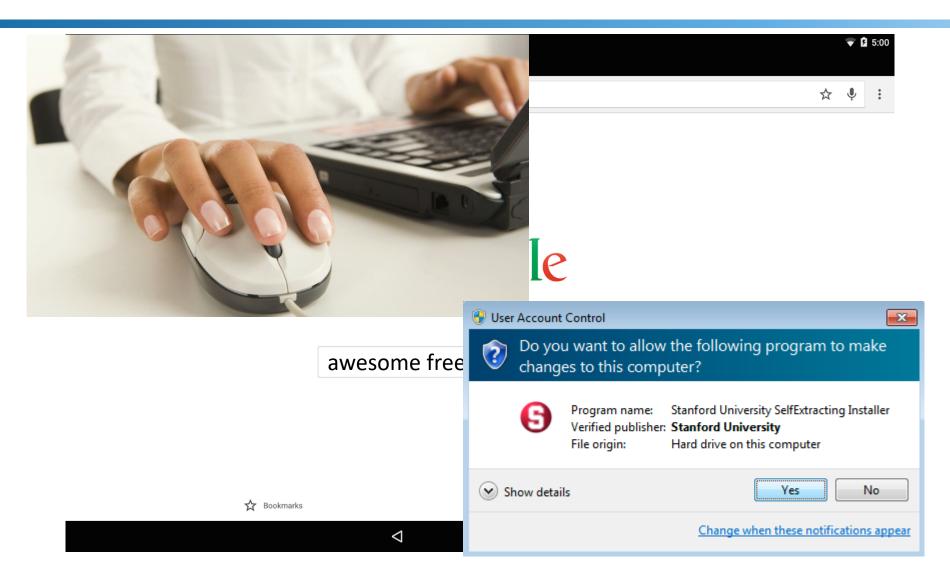
Ļ

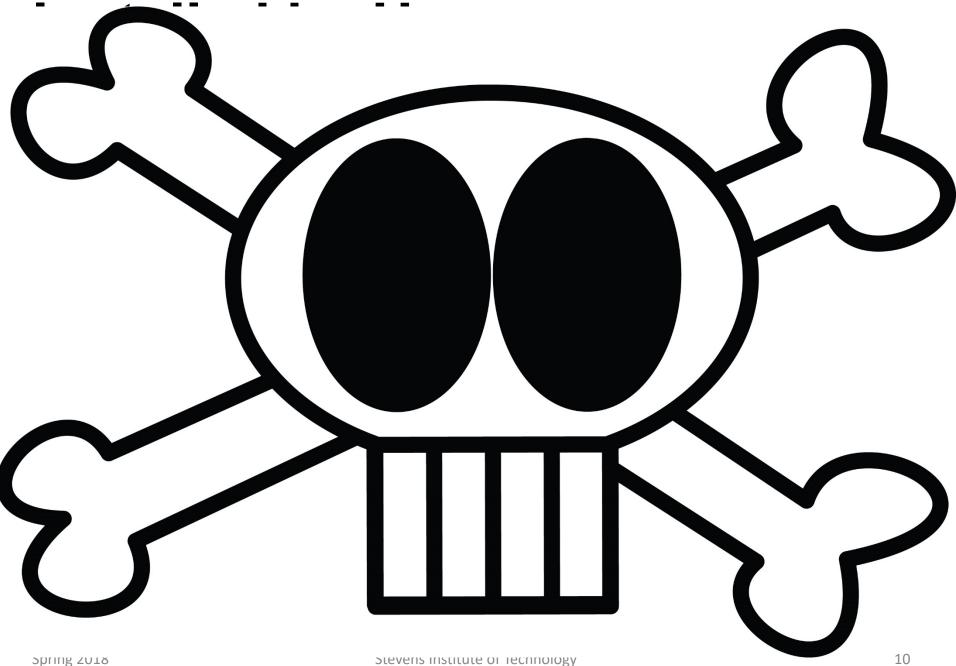


### Installed by User

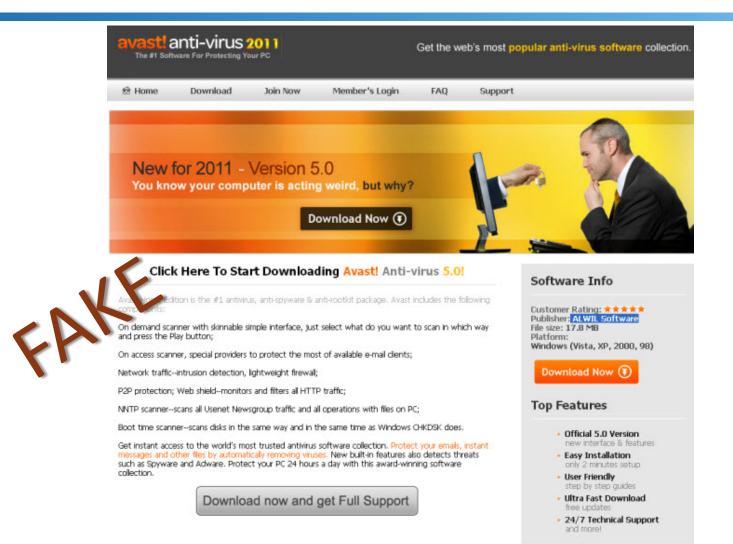


### Installed by User



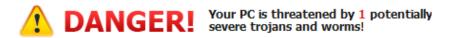


#### Malware Anti-Malware



#### **ScareWare**

#### **Personal Antivirus**



Viruses are programmed to damage the computer by damaging programs, deleting files, or reformatting the hard disk. As a result, they cause erratic behavior and can result in system crashes. In addition, many viruses are bug-ridden, and these bugs lead to system crashes and data loss.

#### Optimize and protect your system with advanced antivirus technology.

#### Before you register this program, please read the following carefully:

This is a one-time charge. Your credit card will never be rebilled and you will receive UPGRADES FOR FREE! Registration is immediate, and once registered, Personal Antivirus will remove all viruses, spyware, adware and other security risks and block them from accessing your system.

Our best-solution software has been already registered by 876,130 US citizens

#### YOU CAN ALSO MAKE YOUR PC UP-TO-DATE!



You have an exclusive **40% discount**, since US citizens are our most frequent buyers.

# The Malware Is Hidden In the Application



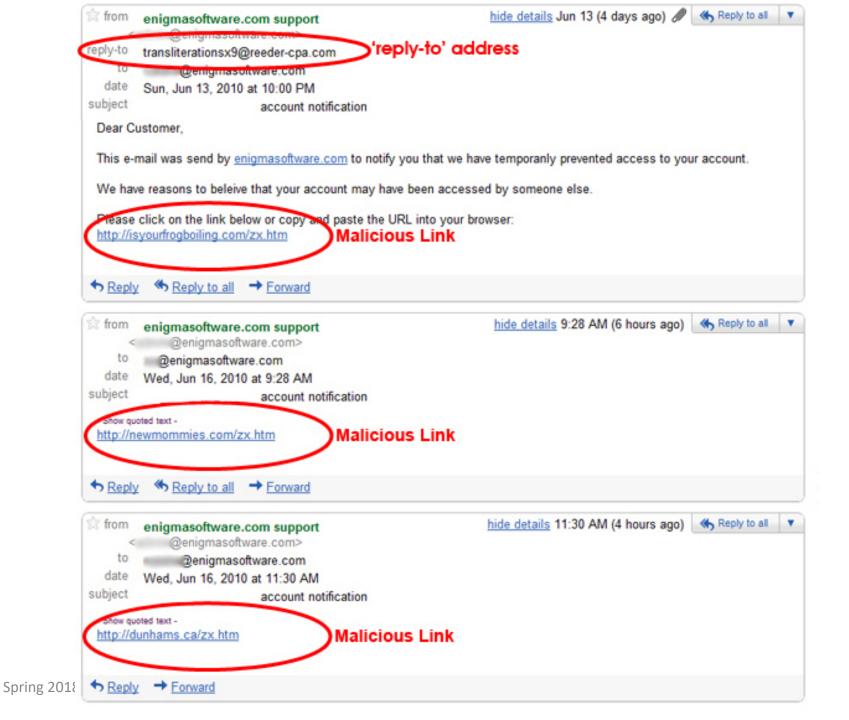
Obtained awesome game from shady source

### You Got Mail

Receiving a malicious executable or link over email

The obvious

- A malicious executable as an attachment
  - Usually compressed
- A link to a website serving malicious executables



Notice of appearance in Court #000215737 - Mozilla Thunderbird _ 💷 🗙					
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>G</u> o <u>M</u> essage <u>T</u> ools <u>H</u> elp					
From State Court <adam.kaiser@c1291.colo.hc.ru>☆ Reply Meply All ➤ Forward More ➤ Subject Notice of appearance in Court #000215737 To bizdev@rackspace.co.uk☆</adam.kaiser@c1291.colo.hc.ru>					
Notice to Appear,					
This is to inform you to appear in the Court on the September 14 for your case hearing. Please, prepare all the documents relating to the case and bring them to Court on the specified date. Note: If you do not come, the case will be heard in your absence.					
You can review complete details of the Court Notice in the attachment.					
Yours faithfully, Adam Kaiser, District Clerk.					
▽ @1 attachment: Notice_to_Appear_000215737.zip 3.6 KB					
Notice_to_Appear_000215737.zip 3.6 KB					

### You Got Mail

Receiving a malicious executable or link over email

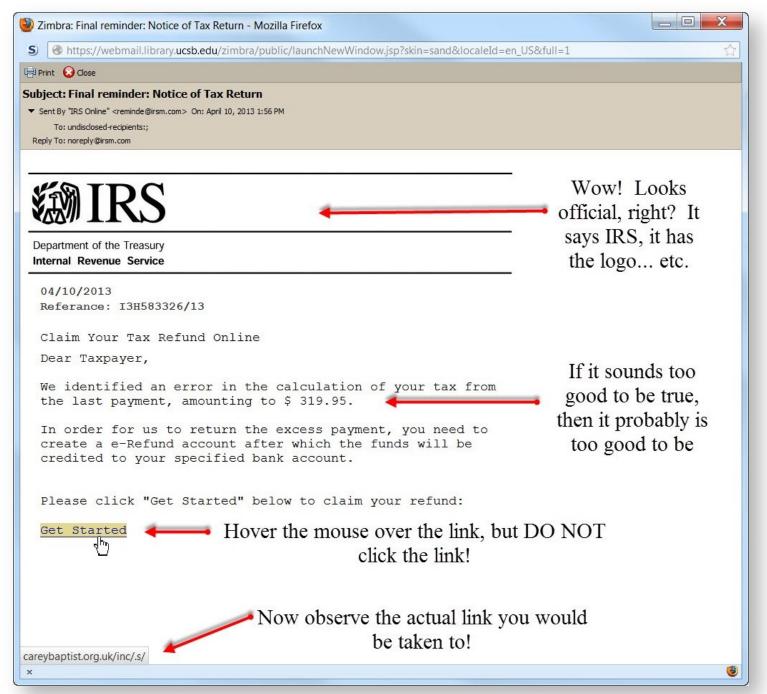
The obvious

- A malicious executable as an attachment
  - Usually compressed
- A link to a website serving malicious executables

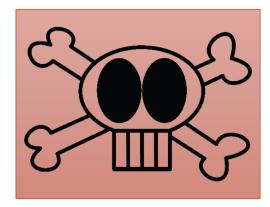
#### Less obvious

- The attachment has non-executable extension (e.g., .gif), but will be executed when opened
  - The file magic number is used instead of the extension
- Two file extensions are used and system hides known extensions
  - Example: "Image.gif.exe"
- HTML emails with "hidden" URLs

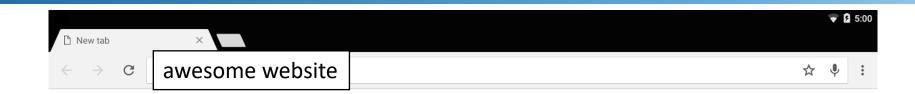
🏘 Re: Your archive - Spam Filtered - Netscape Folder										
<u>File Edit View Go M</u>	lessage <u>C</u> o	mmunicati	or <u>H</u> elp							
Get Msg New Msg	斗 Reply F	Reply All	Forward	💕 File	Vext	et al. Print	) Delete	Stop		
Name	U	Total 4		Subject					Sender	Date
- Metwork	00501246-0004	135 :			Your archiv	/e			kennmcd@pewterauv.com	7:40 AM
Subject: Re: Your archive Date: Tue, 20 Apr 2004 07:40:24 -0600 From: kennmcd@pewterguy.com To: support@westernet.net Please have a look at the attached file.										
Name: your_archive.pif Type: Shortcut to MS-DOS Program (application/x-unknown-content-type-piffile) Encoding: base64										

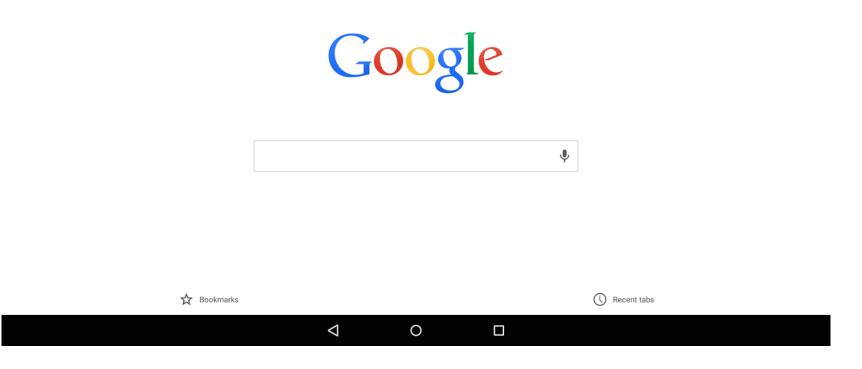


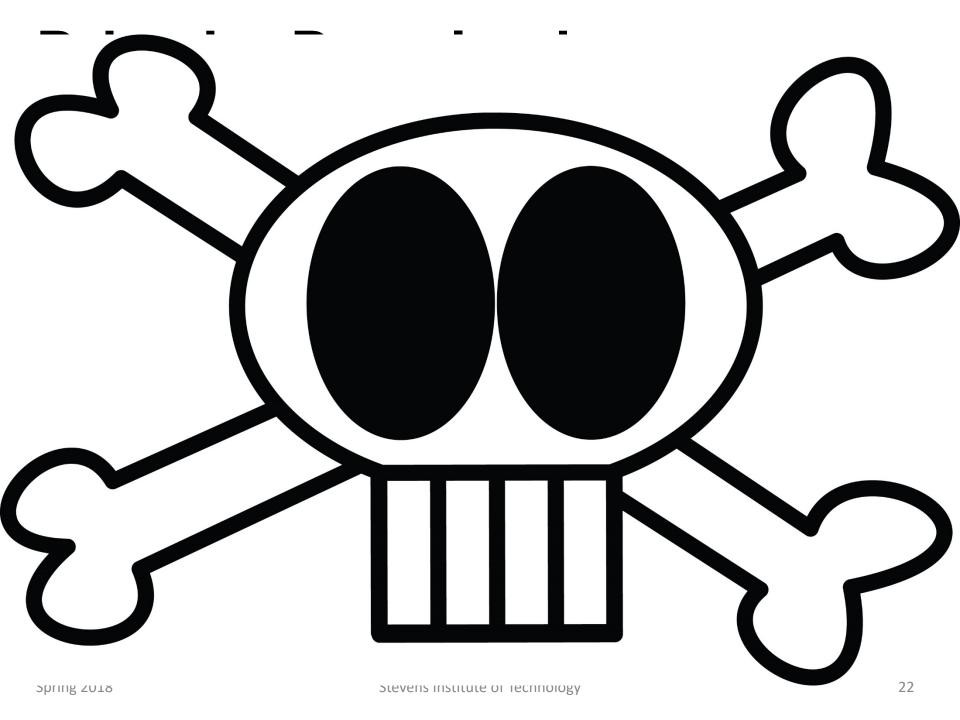
#### The Malware Is the Downloaded and Executed Application



### **Drive-by-Downloads**







## **Typical Drive-by-Download**

User visits a website

 Site does not have to be malicious, but could include a malicious ad

The website exploits a vulnerability in one of the browser's components

- The malware infects and takes over the browser
  - For example, through a memory corruption vulnerability
- Less frequently, other vulnerabilities that trick the browser to download and execute a program

Usually, no interaction is required for the second step

Hence, drive-by

### **Over Email**

Email includes document carrying exploit

• PDF, doc, etc.

The document exploits a memory corruption vulnerability

Or it can include malicious interpreted code

Scripts, macros, etc.

	- 10 4 V =	registration form for journalists - Message (HTML)	_ = ×
	Message Adobe PDF		۲
From:	Construction and a second	Sent:	Thu 3/25/2010 4:16 PM
To:	<ul> <li>The Depart (R. d);</li> </ul>		
Co		MALICIO	US PDF
Subject:	registration form for Jour	nalists	
🖂 Mes	ssage 👔 🔁 englishversinon.pdf	(143 KB)	
(See	attached file: englishver	rsinon.pdf)	
-			
Dea			
	-	EXCEL forms but you have chan	
	-	inge with these two forms and :	fill out
	thank you.		
-	ards.		
bur	eau of snangnal W	orld expo coordination	
L			¥

### **Computer Worms**

Malware that exploits software vulnerabilities in client or server programs

It actively seeks out vulnerable hosts to infect  $\rightarrow$  self replicates

Propagation vectors include:

- The network (network shared, server vulnerabilities, email)
- Physical media (USB drives, CD, DVD data disks)

#### Random

 Generate random IP addresses and probe them

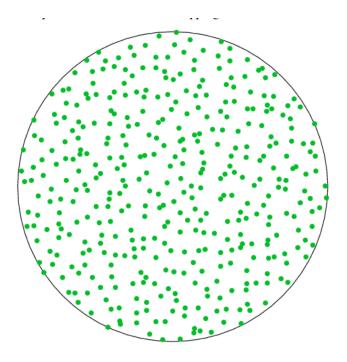
Local subnet

- First scan the local subnet for targets
  - Bypasses network-periphery defenses
  - Faster than random

#### Hit-list

- Compile a long list of potentially vulnerable machines
- Parts of the list are distributed to "siblings" of the worm
- Fast!

#### Topological



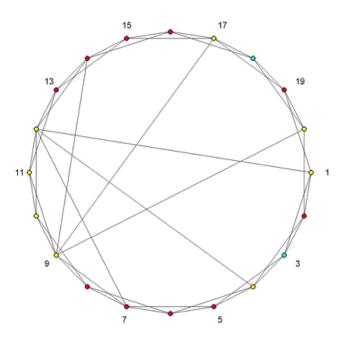
#### Random

- Generate random IP addresses and probe them
- Local subnet
  - First scan the local subnet for targets
    - Bypasses network-periphery defenses
    - Faster than random

#### Hit-list

- Compile a long list of potentially vulnerable machines
- Parts of the list are distributed to "siblings" of the worm
- Fast!

#### Topological



#### Random

- Generate random IP addresses and probe them
- Local subnet
  - First scan the local subnet for targets
    - Bypasses network-periphery defenses
    - Faster than random

#### Hit-list

- Compile a long list of potentially vulnerable machines
- Parts of the list are distributed to "siblings" of the worm
- Fast!

#### Topological



#### Random

 Generate random IP addresses and probe them

Local subnet

- First scan the local subnet for targets
  - Bypasses network-periphery defenses
  - Faster than random

#### Hit-list

- Compile a long list of potentially vulnerable machines
- Parts of the list are distributed to "siblings" of the worm
- Fast!

#### Topological



### Case Study: Morris Worm

Earliest significant worm infection

Released by Robert Morris in 1988



- The Internet consisted of about 60,000 computers
- They were connected using TCP/IP
- Mostly run BSD Unix



### **Vectors of Infection**

Attempted to crack local password file to use login/password to logon to other systems

Exploited a bug (buffer overflow) in the finger daemon

Exploited a trapdoor in the debug option of the remote process that receives and sends mail

- DEBUG function enables one to run a program at the host
  - Added by the author of sendmail to remotely troubleshoot systems
- Who needs shellcode! Morris sent a C program in an email, compiled it, and executed it to fetch the rest of the worm from the already infected host

### **Target Acquisition**

Internet too sparse for random scanning

On SUN and VAX architectures other hosts can be found by looking in

- /etc/hosts.equiv
- /.rhosts
- .forward
- ~/.rhosts
- Routing tables
- End points of point-to-point interfaces



### What It Never Did

Gain privileged access

**Destroy data** 

Install time bombs or backdoors

Brute force the root account



#### Flaws

Never checked if a host is already infected

Routines would exit with errors while leaving a copy of the virus running:

- When multiple instances of the worm attempted to infect a clean host concurrently
- When multiple instances of the worm attempted to infect and already compromised host
- When a machine is heavily loaded (remember it is actually compiling)

#### Infection rate was proportional to number of instances of the worm running on a host

### **Bad Target Finding**

# Worm checked for telnet or rsh ports to determine if a host is running UNIX

Some systems only run sendmail!

It did not utilize DNS

### **First Reactions**

Administrators cut off the sendmail service

**Big mistake!** 

It shut off communication channels necessary to fix the error

The worm had alternate ways to propagate

## The Author

Robert Morris was a student at Cornell University

He was identified, tried, and convicted in 1990

3 years of probation, fines, etc.

Received a PhD from Harvard

Now a professor at MIT





## More Examples: Code Red

### 13 July 2001 – CodeRed: Buffer overflow in Microsoft IIS

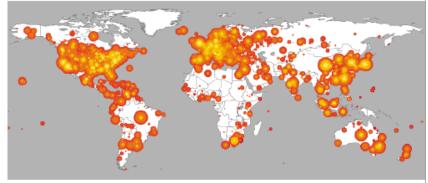
#### Defaced affected website:

HELLO! Welcome to http://www.worm.com! Hacked By Chinese!

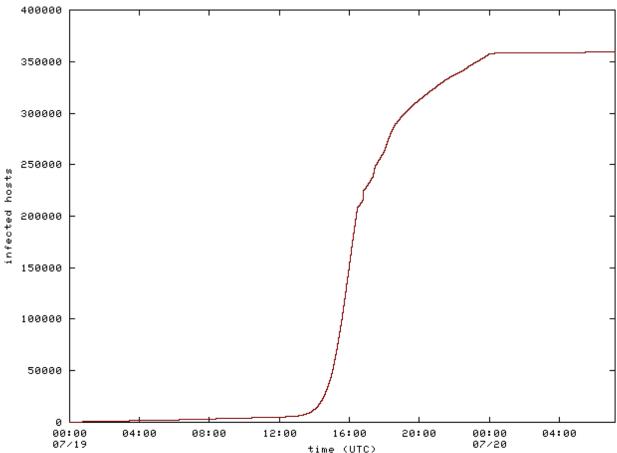
Days 1–19: propagate through random scanning Days 20–27: DoS attack against www.whitehouse.gov

### 4 August 2001 – CodeRed II

Localized scanning



### **Code Red Propagation**



Code Red Worm - infected hosts

## **Example: Nimda Worm**

### 18/9/2001 – Nimda

### Many infection vectors

Code Red IIS buffer overflow

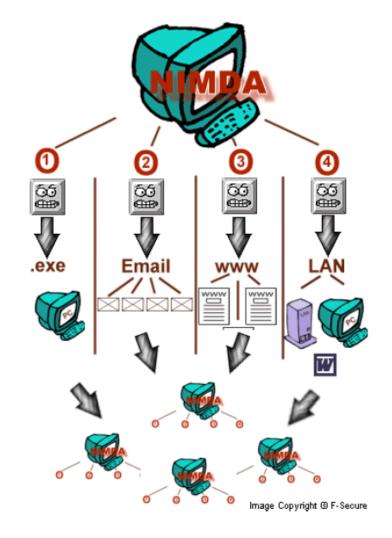
Bulk email to harvested addresses from victim host

Open network shares

Infect visitors of compromised web sites

Microsoft IIS 4.0/5.0 directory traversal vulnerabilities

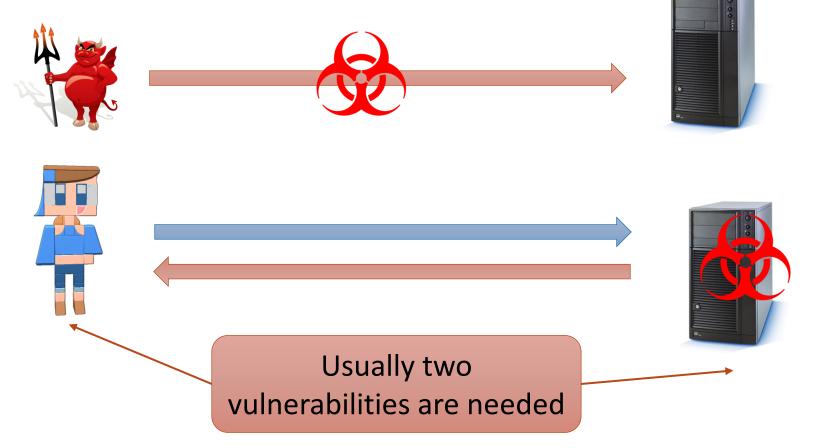
Backdoors left behind by the Code Red II and Sadmind/IIS worms



## **Passive Propagation**

Infect service and wait for clients to connect

Infect connected clients

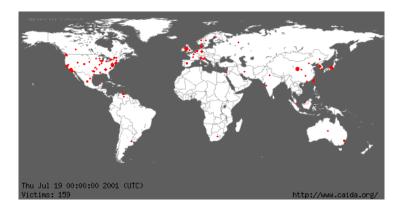


## **Recent Worm Attacks**

Melissa	1998	e-mail worm first to include virus, worm and Trojan in one package
Code Red	July 2001	exploited Microsoft IIS bug probes random IP addresses consumes significant Internet capacity when active
Code Red II	August 2001	also targeted Microsoft IIS installs a backdoor for access
Nimda	September 2001	had worm, virus and mobile code characteristics spread using e-mail, Windows shares, Web servers, Web clients, backdoors
SQL Slammer	Early 2003	exploited a buffer overflow vulnerability in SQL server compact and spread rapidly
Sobig.F	Late 2003	exploited open proxy servers to turn infected machines into spam engines
Mydoom	2004	mass-mailing e-mail worm installed a backdoor in infected machines
Warezov	2006	creates executables in system directories sends itself as an e-mail attachment can disable security related products
Conficker (Downadup)	November 2008	exploits a Windows buffer overflow vulnerability most widespread infection since SQL Slammer
Stuxnet	2010	restricted rate of spread to reduce chance of detection targeted industrial control systems
Spring 2019		Stavans Institute of Tashnalagy AF

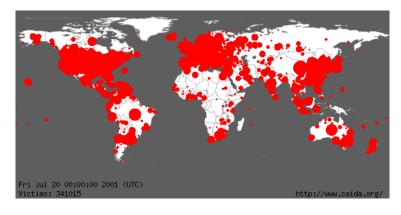
### **Fast Worms**

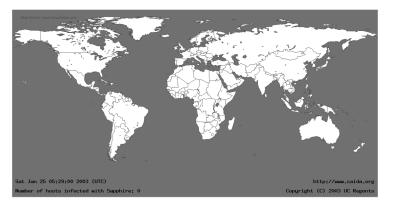




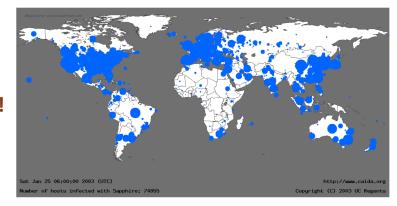
July 19, 2001 spread of CODE RED in 24 hours!







Jan 29, 2003 spread of SLAMMER in 30 minutes!

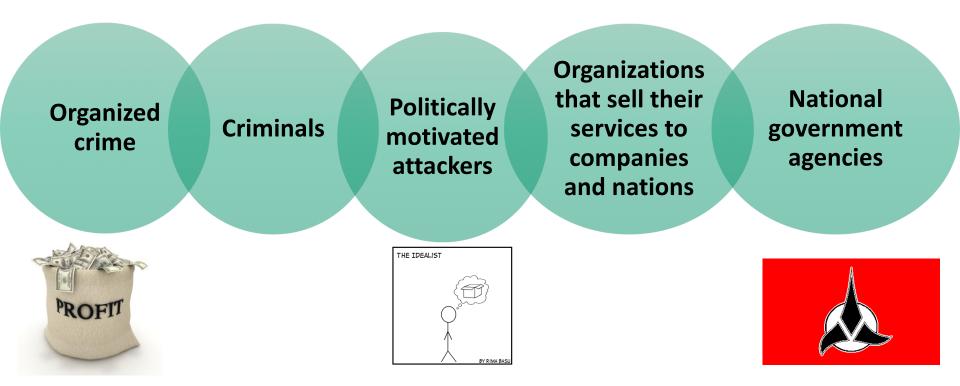


### **Payloads**

### Payloads

Click fraud	Data theft
Bank fraud	Identity theft
Phishing	Extortion
	DDoS
Spamming	Identity theft Extortion
Spreading malware	Sabotage

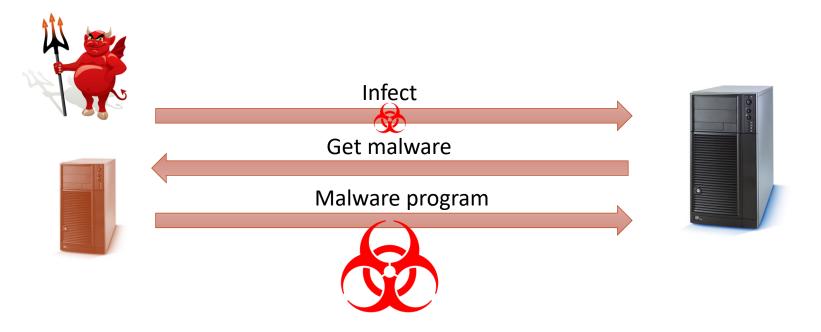
### **Sources/Motivation**



## **Two Stage Payloads**

The more complex the payload the larger the malware

Maybe too large to send along with the infection



## **Identifying Known Malware**

Looking for exact matches of known malware



### Looking for known byte/instruction sequences

$\bigcirc$	4
Sigg	glyre
$\langle   ^{\circ}$	U

GET /default.ida?\*%u9090%u6858%ucbd3%u7801%u9090 %u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090 %u9090%u8190%u00c3%u0003%u8b00%u531b%u53f f%u0078%u0000%u00=a HTTP/\*

## Polymorphism

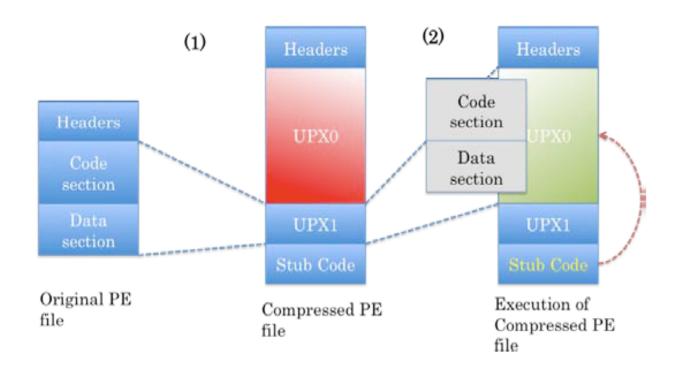
Methods for changing the content of malware to evade detection

### Methods:

- Appending/prepending garbage/data
- Encrypt payload using different key for each infection
  - Compression is also an option

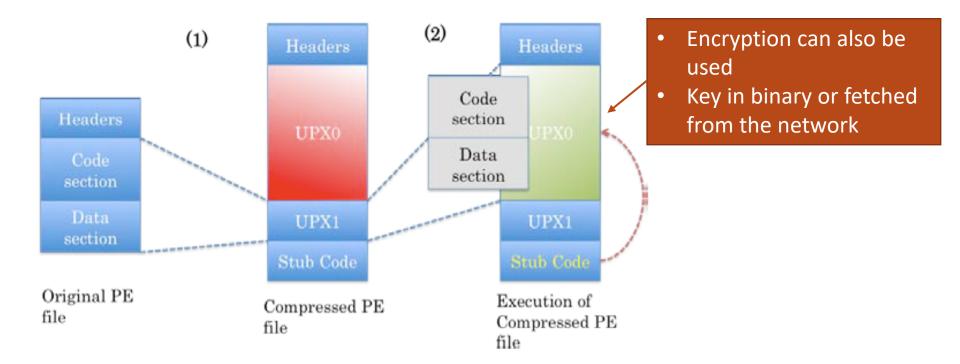
## **Creating Variants**

### Compression (packers) Examples: UPX, Aspack, FSG, PE Compact, ...



## **Creating Variants**

### Compression (packers) Examples: UPX, Aspack, FSG, PE Compact, ...



## Metamorphism

Create different "versions" of the program code that look different but have the same semantics (i.e., do the same thing)

### **Example techniques:**

Dead-code insertion Instruction reordering Instruction substitution Register substitution

. . .

### **Dead Code Insertion**

5B	pop ebx
8D 4B 42	lea ecx, $[ebx + 42h]$
51	push ecx
50	push eax
90	nop
50	push eax
40	inc eax
OF 01 4C 24 FE	sidt [esp - 02h]
48	dec eax
5B	pop ebx
83 C3 1C	add ebx, 1Ch
FA	cli
8B 2B	mov ebp, [ebx]
	0 90 50 40 OF 01 4C 24 FE 48 5B 83 C3
1C FA 8B 2B	

### **Instruction Reordering**

5B EB 09	pop ebx jmp <s1></s1>	1
50 0F 01 4C 24 FE 5B EB 07	S2: push eax sidt [esp - 02h] pop ebx jmp <s3></s3>	3
8D 4B 42 51 50 EB F0	S1: lea ecx, [ebx + 42h] push ecx push eax jmp <s2></s2>	2
83 C3 1C FA 8B 2B	S3: add ebx, 1Ch cli mov ebp, [ebx]	4
5B         EB         091         50         0F           4B         42         51         50         EB         F	4C 24 FE 5B EB 0738D 283 C3 1C FA 8B 2B4	

### Instruction Substitution

pop ebx lea ecx, [ebx + 42h] push ecx push eax push eax sidt [esp - 02h]

pop ebx add ebx, 1Ch cli mov ebp, [ebx]

5B				·
8D	4B	42		
51				
50				
83	EC	04		
89	04	24		
ΟF	01	4C	24	FE
83	04	24	<b>1C</b>	
5B				
FA				
8B	2B			

pop ebx
lea ecx, [ebx + 42h]
push ecx
push eax
sub esp, 04h
mov [esp], eax
sidt [esp - 02h]
add [esp], 1Ch
pop ebx
cli
mov ebp, [ebx]

#### 5B 8D 4B 42 51 50 83 EC 04 89 04 24 0F 01 4C 24 FE 83 04 24 1C 5B 8B 2B

# Behavior-based Malware Detection

Focus on the malicious behaviors demonstrated by malware

Click fraud	Identity theft
Bank fraud	Extortion
Phishing	DDoS
Spamming	
Spreading malware	Espionage
Data theft	Sabotage

### **Exploit Kits**

## **Exploit Kits**

Initially the development and deployment of malware required considerable technical skill by software authors

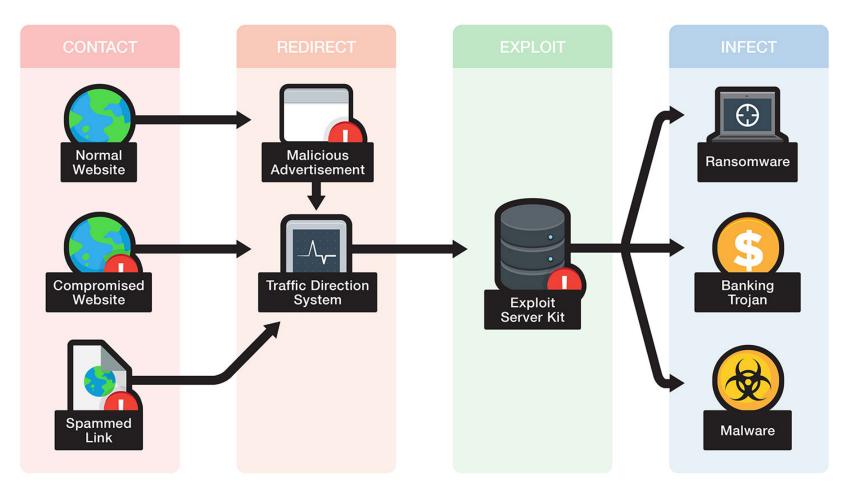
The development of virus-creation toolkits in the early 1990s and then more general attack kits in the 2000s greatly assisted in the development and deployment of malware

- Toolkits are often known as "crimeware"
- Widely used toolkits include: Zeus, Blackhole, Sakura, Phoenix

Include a variety of propagation mechanisms and payload modules that even novices can deploy

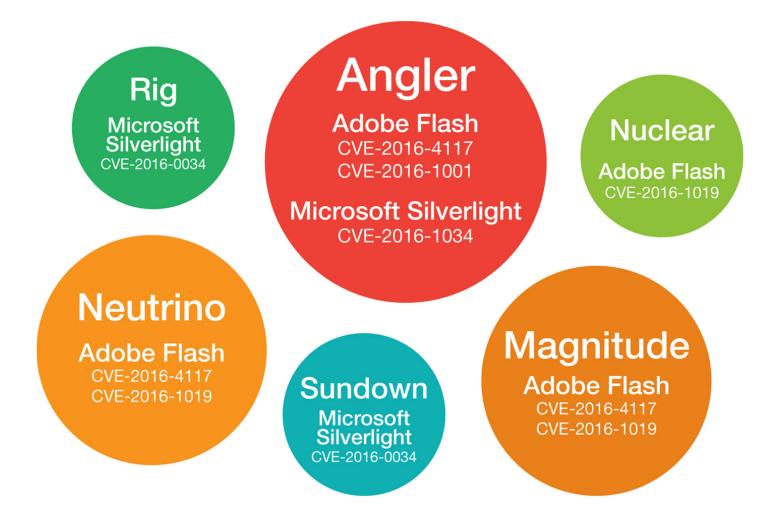
Variants that can be generated by attackers using these toolkits creates a significant problem for those defending systems against them

## **Exploit Kit Infection Chain**



#### http://www.trendmicro.com/vinfo/us/security/definition/exploit-kit

## **Exploit Arsenal in Popular Kits**



http://www.trendmicro.com/vinfo/us/security/definition/exploit-kit

## **Malware Analysis**

## Why Do We Analyze Malware?

To access damage

To identify infection signs for discovering other compromised hosts

To determine how it can be removed

To generate a "vaccine" for the infection

### To learn

## **Types of Analysis**

### Static

Analyze the code without executing it

Disassemble

Higher coverage, prone to obfuscation

### Dynamic

Analyze the code while executing it

Monitor execution

Lower coverage, what I see is what is executed

### Examples of extracted data: Function call graph Control flow graph Invoked APIs

## Obfuscation

### Make program hard to analyze

• For example, polymorphism

### **Common obfuscation approaches:**

Use indirect control transfers

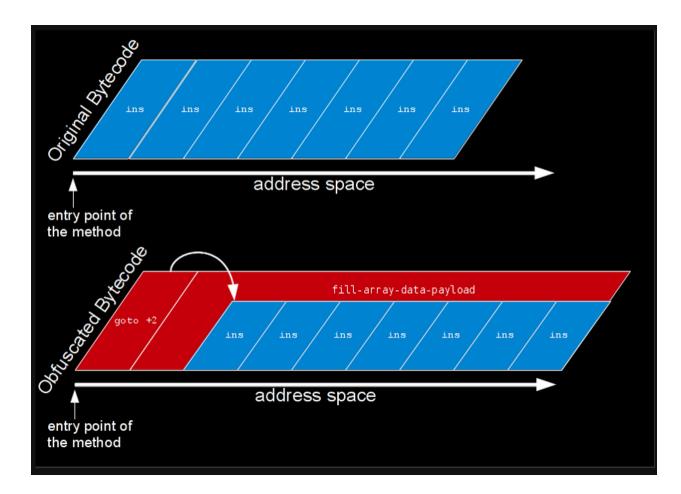
### **Overlapping instructions**

- **x86**
- Dalvik -- http://www.dexlabs.org/blog/bytecode-obfuscation

### Return-oriented programming

https://www.usenix.org/system/files/conference/usenixsecu rity13/sec13-paper\_wang-updated-8-23-13.pdf

0	1	2										
	ins ins			ins ins		ins		ins				
			ins		ins	ir	ıs	ins	ins		ins	
				ins		ins		ins	ins	ins	ins	



Combining both static and dynamic analysis is frequently necessary

# Detecting Virtual Machines and Debuggers

If a program is not run natively on a machine, chances are high that it

- is being analyzed (in a security lab)
- scanned (inside a sandbox of an Antivirus product)
- debugged (by a security specialist)

Modern malware detect execution environment to complicate analysis

- Hide or alter its functionality
- Attempt to breakout

## **Detecting the Matrix VM**

Look for VM artifacts in processes, file system, and/or registry

- Special files and processes (e.g., VMtools)
- Look for VM artifacts in memory
  - How do you peak into memory?
- Look for VM-specific virtual hardware
  - Device names, device parameters (e.g., MAC address)

Look for VM-specific processor instructions and capabilities

- Extra instructions added by VM for guest-host communication
- VMs frequently don't support obscure instructions or have limited support for instructions
- `The red pill'

### Remotely by inspecting IP packets

## **Detecting VMs with the Red Pill**

### Proposed by Joanna Rutkowska

Runs a single instruction

- SIDT  $\rightarrow$  Store Interrupt Descriptor Table
  - Store IDT in memory



- Facts and observations
  - On VMware guest machines, the IDT is typically located at 0xffXXXXX
  - On VirtualPC guests, it is located at 0xe8XXXXXX
  - On host operating systems, it is located lower than that, typically around 0x80ffffff (Windows) and 0xc0ffffff (Linux)
- It is sufficient to look at the first byte of the address
  - If it's greater than 0xd0, you've got a virtual machine
  - If it is less than or equal to 0xd0, you are in a real machine

### Red Pill++

Other OS features can be obtained with other instructions

- The Global Descriptor Table (GDT), measured by the SGDT instruction
- The Local Descriptor Table (LDT), measured by the SLDT instruction

## **Anti-Debugging**

Detecting a debugger is easy

- Windows: IsDebuggerPresent()
- Linux: ptrace()

```
if (ptrace(PTRACE_TRACEME, 0, NULL, 0) == -1)
    printf("traced!\n");
```

 $\leftarrow$ 

trapkit.de/tools/scoopyng



Home - Books - Advisories - Blog - Tools - Twitter - Papers - About

#### ScoopyNG — The VMware detection tool

Latest Version: v1.0 from 2008 - Current Status: Not further maintained.

ScoopyNG combines the detection tricks of Scoopy Doo and Jerry as well as some new techniques to determine if a current OS is running inside a VMware Virtual Machine (VM) or on a native system.

ScoopyNG should work on all modern uni-, multi- and multi-core cpu's.

ScoopyNG is able to detect VMware even if "anti-detection-mechanisms" are deployed.

#### FAQ — Interpretation of results

Q: How do I know if it's indeed a VMware Virtual Machine? A: If one or more of the test results state that VMware is detected.

Q: Do all the tests have to state that it's VMware? A: No, it is sufficient that **one** test result says that VMware is detected.

#### Sample results

Sample 1: native, Windows Vista SP1 32bit, Intel(R) Core(TM)2 Duo CPU

### http://www.trapkit.de/tools/scoopyng/

ഹ

□☆

=

1

×

## **Advanced Persistent Threats**

Well-resourced attacks targeting high-value targets

Aim to gain persistent presence in a system or network

- By utilizing multiple attack vectors
- By employing **hiding** techniques

# High profile attacks include Aurora, RSA, APT1, and Stuxnet

## How Would you Hide?

### Deception

Present a fake image of how things are (Potemkin village)

### Methods:

Masquerade

## Lie

- Modify programs to lie
- Modify the kernel to lie
- Modify VM to lie
- Modify the HW to lie?



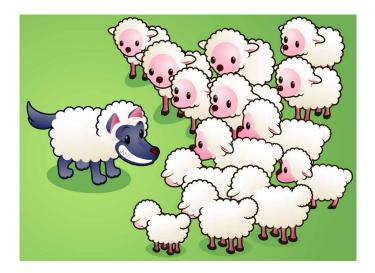
## Masquerading

Assume unsuspecting filename

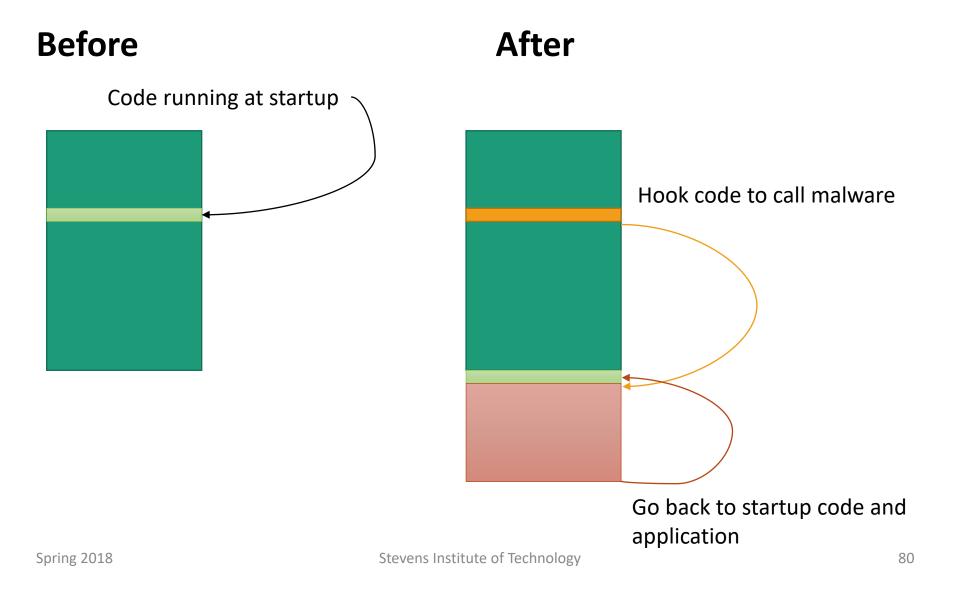
### Stealthy operation

- Piggyback communications
- Small footprint

Infect a legitimate application



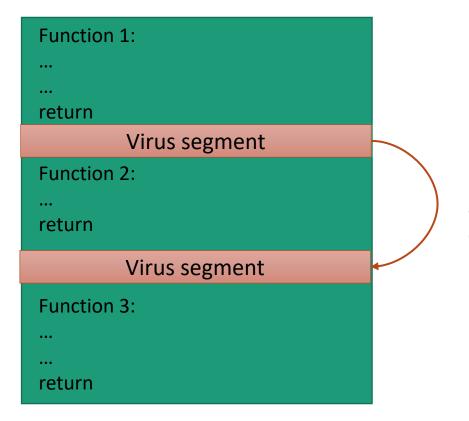
## **Application Infection**



## **Hiding Between Functions**

Function 1:
return
Gap left for alignment
Function 2:
return
Gap left for alignment
Function 3:
return

## **Hiding Between Functions**



Malware jumps from segment to segment to compose its functionality

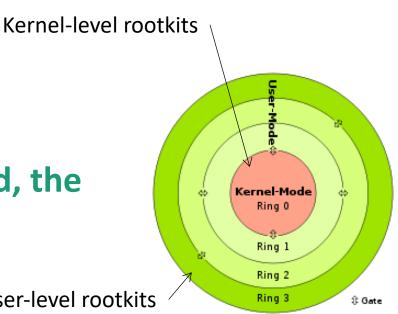
# Lying

### Rootkits $\rightarrow$ Malicious software designed to hide malware related data

- Files
- Processes
- Logins
- Network connections

### The inner the level controlled, the better!

### User-level rootkits



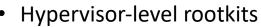
# Lying

## Rootkits $\rightarrow$ Malicious software designed to hide malware related data

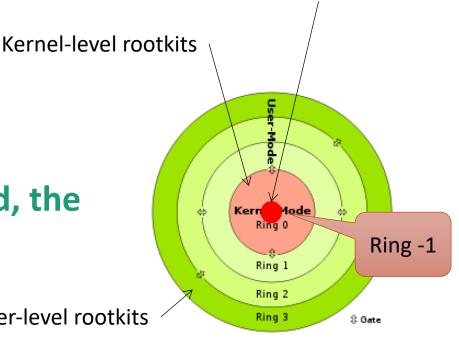
- Files
- Processes
- Logins
- Network connections

### The inner the level controlled, the better!

### User-level rootkits



- **Bootkits**
- Firmware-level bootkits



## **User-level Rootkits**

### Modify:

- Utilities  $\rightarrow$  ps, netstat, top, sshd
- Applications → Alter behavior (e.g., modify Windows Explorer to hide a file)
- API hooks  $\rightarrow$  replace system calls, etc.

## **Kernel-level Rootkits**

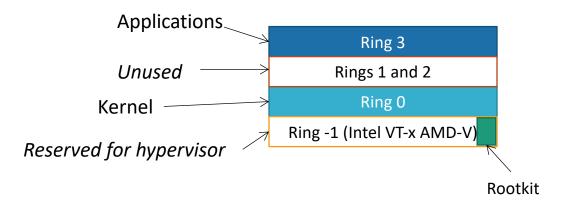
### Modify or add:

Kernel code (*Phantasmagoria* adds instructions in system calls) Kernel data structures (remove malware from process lists, *FU*) APIs (*Knark* adds entries in the proc file system, SuckIT adds new system calls)

### Mostly implemented as *Loadable Kernel Modules*

## **Hypervisor Rootkits**

Runs with higher privilege than the kernel Developed in academia (SubVirst, Blue Pill)



## **Firmware-level Rootkits**

Firmware is the lowest-level of software that controls certain operations of hardware

Till recently the integrity of firmware was not checked

• Companies have started using **signed** firmware updates

## Examples:

- Organized crime tampers with European card swipe device <u>http://www.theregister.co.uk/2008/10/10/organized crime doctors chip and pin machines</u>
- Attacks on BIOS anti-theft devices turn them into rootkits <u>http://www.blackhat.com/presentations/bh-usa-09/ORTEGA/BHUSA09-Ortega-DeactivateRootkit-PAPER.pdf</u>

## (Some) Defenses

Check for file integrity  $\rightarrow$  Tripwire, chkrootkit

Check for divergent results  $\rightarrow$  checkps

Protecting hooks  $\rightarrow$  system calls, internal kernel APIs

Code integrity checks  $\rightarrow$  page-level signing

## **File Integrity Testing**

Create checksums (hashes) of binaries on the system Periodically check installed binaries vs stored checksums Application signing

### **Challenges:**

Storing the checksums out of reach Keeping up with updates Storing the tools out of reach!

## **Looking for Divergent Results**

### Run binaries and collect results

- ps, top, netcat
- Collect results from other sources
  - Directly access /proc filesystem
- Compare results to find discrepancies

### **Challenges:**

- Find other sources of information
- False alerts, system state is dynamic
- Storing the tools out of reach!

## **Monitor API Hooks**

Store currently used, good set of hooks

Periodically read the values of hooks

Compare values to identify hooks being replaced

### **Challenges:**

- Which APIs should be monitored
- False alerts, hooks can be placed for legitimate reasons
  - That's usually the problem with running multiple antivirus engines on your PC

## **Code-integrity Checking**

Upon loading a page of code hash its contents

Periodically re-hash every page and check it against previously taken hash

Can be done

- By the kernel
- A hypervisor
- A coprocessor

### **Challenges:**

- Storing the hashes out of reach
- Keeping up with code updates
- Code provenance/generated code
- Pages containing both code and data