# Malware

Thanks to Ari Juels for most of this deck!

# Malware

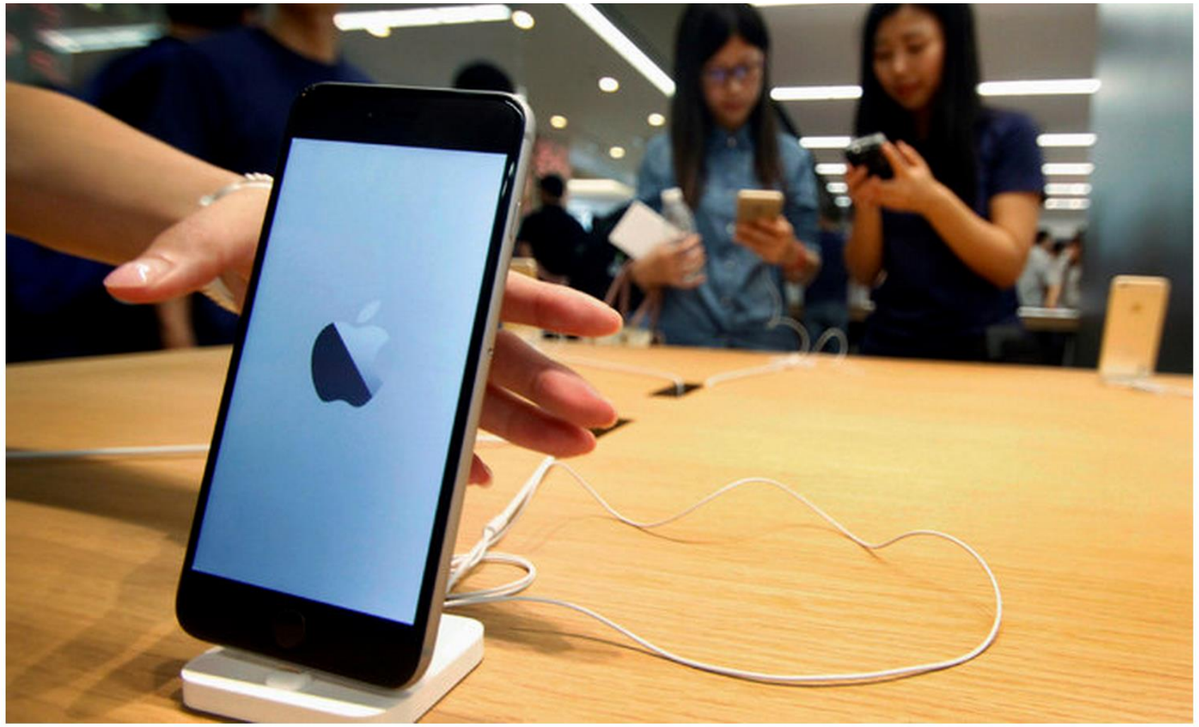
- **Mal**icious soft**ware** is a generic term that refers to hostile executable code. It can
- Be self-replicating and
  - Live in host code: virus
  - Or be freestanding: worm
- Or it can be non-self-replicating: Trojan
- Can masquerade as good software or install itself without user knowing
- Generally aims to disrupt or modify correct system operation or steal or modify secrets

# Can propagate in tricky ways

TECHNOLOGY

## *Apple Confirms Discovery of Malicious Code in Some App Store Products*

By KATIE BENNER SEPT. 20, 2015



Inside an Apple store in China over the weekend. The company said on Sunday that it was working with developers to make sure they were using the proper version of Xcode. CHINATOPIX, via Associated Press

# “Reflections on Trusting Trust”

- Ken Thompson’s 1983 Turing Award lecture
  - Showed how to add back door password Trojan to UNIX “login” command
    - Visible in “login” source
  - Then modified C compiler to hide Trojan in source and add at compile time
    - Visible in C compiler source
  - Then modified compiler to insert Trojan into newly compiled compilers
    - Now present only in binary
- “You can't trust code that you did not totally create yourself. (Especially code from companies that employ people like me.) No amount of source-level verification or scrutiny will protect you from using untrusted code.”

# Basic challenges

Malware history is game of hide and seek. Why?

- It's hard or impossible to determine how a program will behave.
  - Halting Problem: Whether or not a program will ever ***halt*** is an ***undecidable*** problem
    - Classic result by Turing
- It's not always clear what malware is, e.g.,
  - Is a useful mobile app malware if it steals your address book?
  - What if it removes your address book with permission, but the permission is buried in an unreadable user agreement?
  - Ambiguity started with Creeper vs. Reaper…

# Creeper vs. Reaper

- **Creeper** was created in 1971 (at BBN)
  - World's first worm: Jumped across systems (tried not to replicate)
  - "Enhanced Creeper" replicated
  - Displayed message "I'M THE CREEPER: CATCH ME IF YOU CAN"
  - "Infected" DEC PDP-10s running Tenex OS
- Became an annoyance
- The worm **Reaper** was created to hunt down Creeper and log it out.
- Was Creeper malware?
- Was Reaper malware?

# Viruses

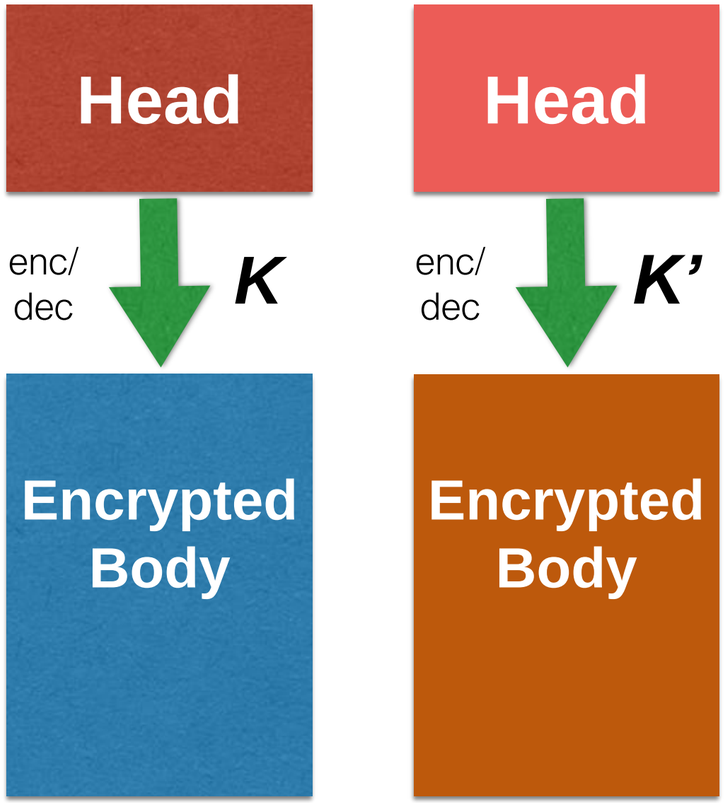
- A computer virus *infects other programs.*
  - Infected program must be executed to invoke virus
  - Virus *self-replicates* (as opposed to Trojan)
  - Installed without user consent
- Self-replication can exploit many different vectors, e.g.,
  - Infect stored executables
  - Infect OS routines to remain in memory
  - Infect disk boot sectors

# Anti-virus

- Anti-virus software identifies and attempts to purge malware of any type (not just viruses).
- Simple anti-virus software looks for ***signatures.***
  - (Remember signature-based intrusion detection?)
- A signature is a ***code fragment*** for a ***known*** virus.

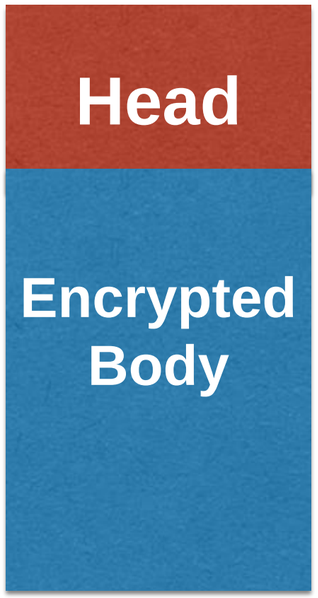
# Polymorphic viruses

- Code changes while virus functionality doesn't.
  - (Short) head + encrypted body or payload
  - When virus is executed, head decrypts payload.
  - Encryption key can change.
  - Head is short, so it can be easily changed too.
- Defeats basic signature checking

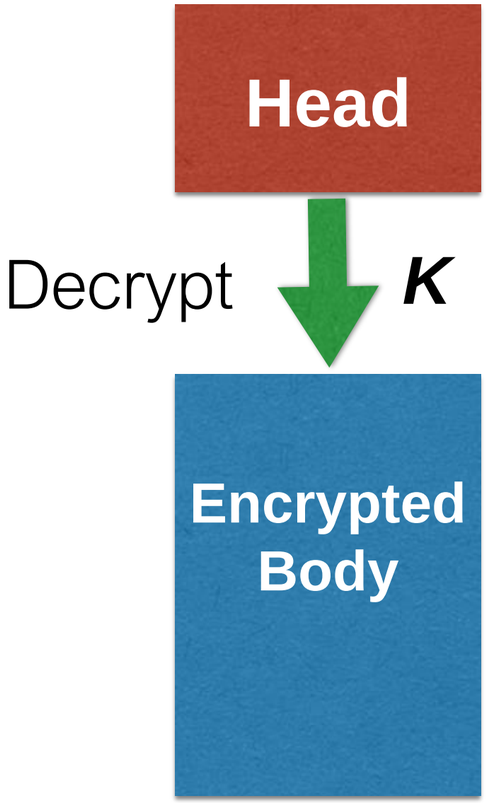


# Polymorphic virus

Encrypted virus



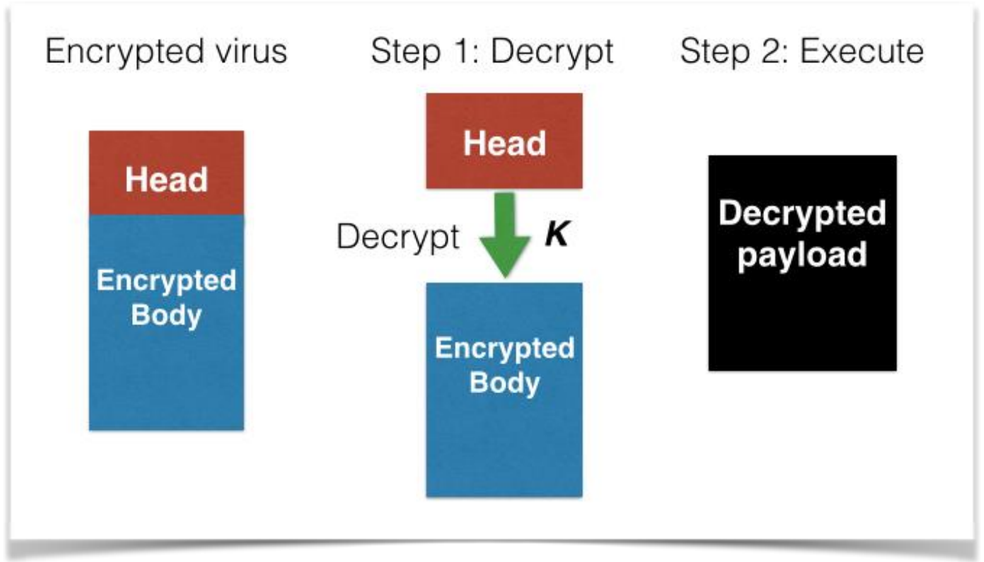

Step 1: Decrypt




Step 2: Execute

# Anti-virus defense: emulation

- Idea: Emulate execution of potentially infected code
- Look for signature of decrypted payload
- Defeats polymorphism!
- But…
  - Doesn't work if virus is deeply embedded in host executable; emulation too long
  - Can be defeated by next stage in virus evolution…



# Metamorphic viruses

- Mutate the payload too!
- Achieved by, e.g.,
  - Inserting NOPs
  - Changing registers
  - Reordering independent instructions or instruction-sequences
  - Swapping in equivalent code sequences

E.g., W95 / Bistro entry point

```
55        push    ebp
54        push    esp
5D        pop     ebp
8B7608    mov     esi, dword ptr [ebp + 08]
09F6      or      esi, esi
743B      je      401045
8B7E0C    mov     edi, dword ptr [ebp + 0c]
85FF      test    edi, edi
7434      je      401045
28D2      sub     edx, edx
```

```
55        push    ebp
8BEC      mov     ebp, esp
8B7608    mov     esi, dword ptr [ebp + 08]
85F6      test    esi, esi
743B      je      401045
8B7E0C    mov     edi, dword ptr [ebp + 0c]
09FF      or      edi, edi
7434      je      401045
31D2      xor     edx, edx
```

# Metamorphic viruses

Anti-virus defense?

- Common subsequences in bold face
- Behavioral detection, but behavior AV incurs false positives and
- False positives are unacceptable for consumers

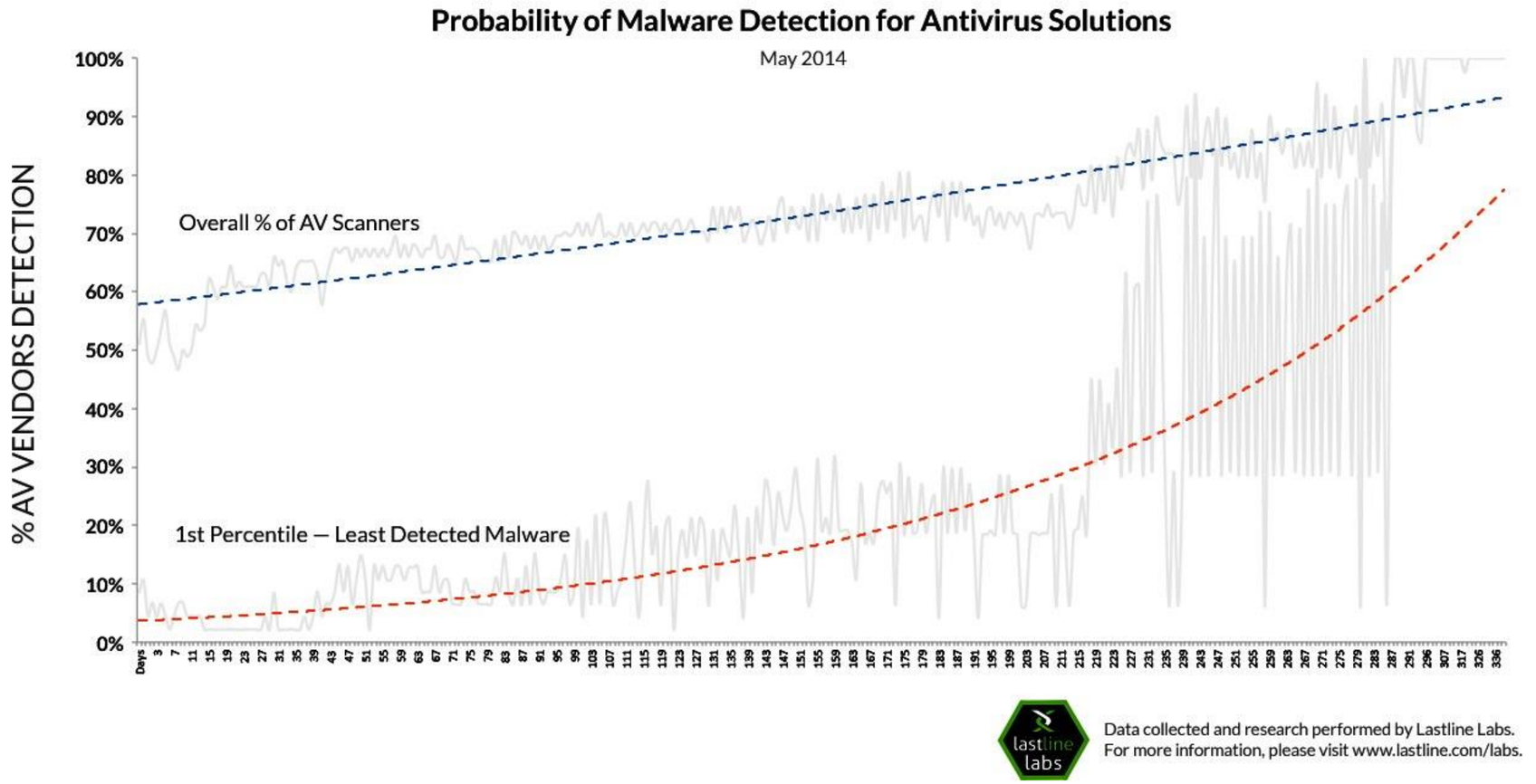
E.g., W95 / Bistro entry point

```
55        push    ebp
54        push    esp
5D        pop     ebp
8B7608    mov     esi, dword ptr [ebp + 08]
09F6      or      esi, esi
743B      je      401045
8B7E0C    mov     edi, dword ptr [ebp + 0c]
85FF      test    edi, edi
7434      je      401045
28D2      sub     edx, edx
```

```
55        push    ebp
8BEC      mov     ebp, esp
8B7608    mov     esi, dword ptr [ebp + 08]
85F6      test    esi, esi
743B      je      401045
8B7E0C    mov     edi, dword ptr [ebp + 0c]
09FF      or      edi, edi
7434      je      401045
31D2      xor     edx, edx
```

# Limitations of AV

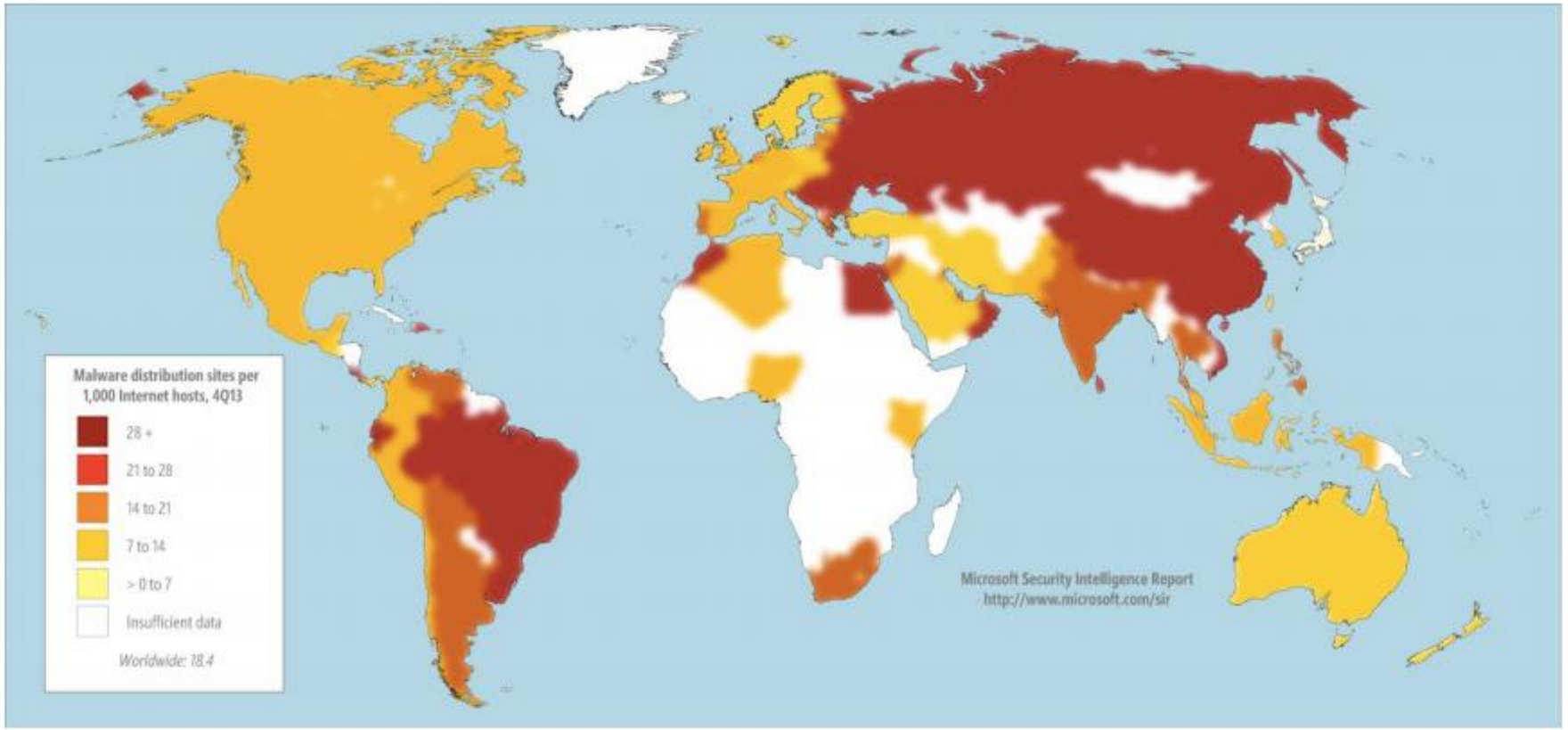
Tested against 47 vendors in VirusTotal from May 2013 to May 2014




Lastline Labs: http://labs.lastline.com/lastline-labs-av-isnt-dead-it-just-cant-keep-up

# Where is malware coming from?

## Global distribution of malware hosting sites




[Microsoft Security Intelligence Report, http://www.microsoft.com/security/sir, vol. 16]

# Where is it going?

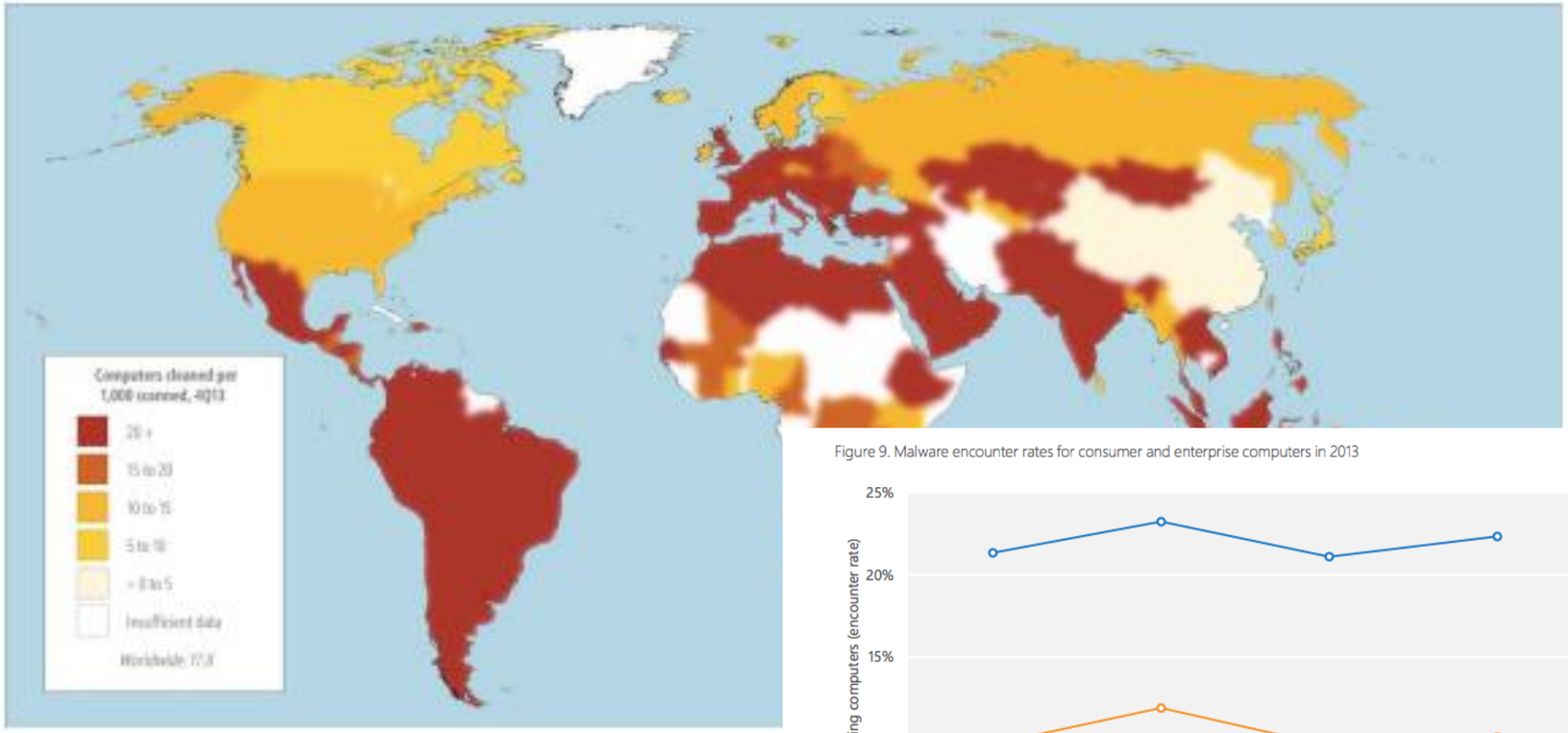
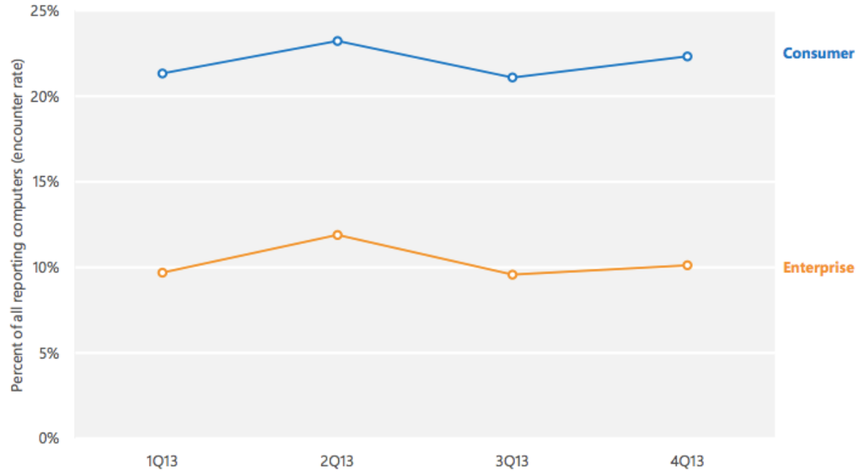



Figure 9. Malware encounter rates for consumer and enterprise computers in 2013

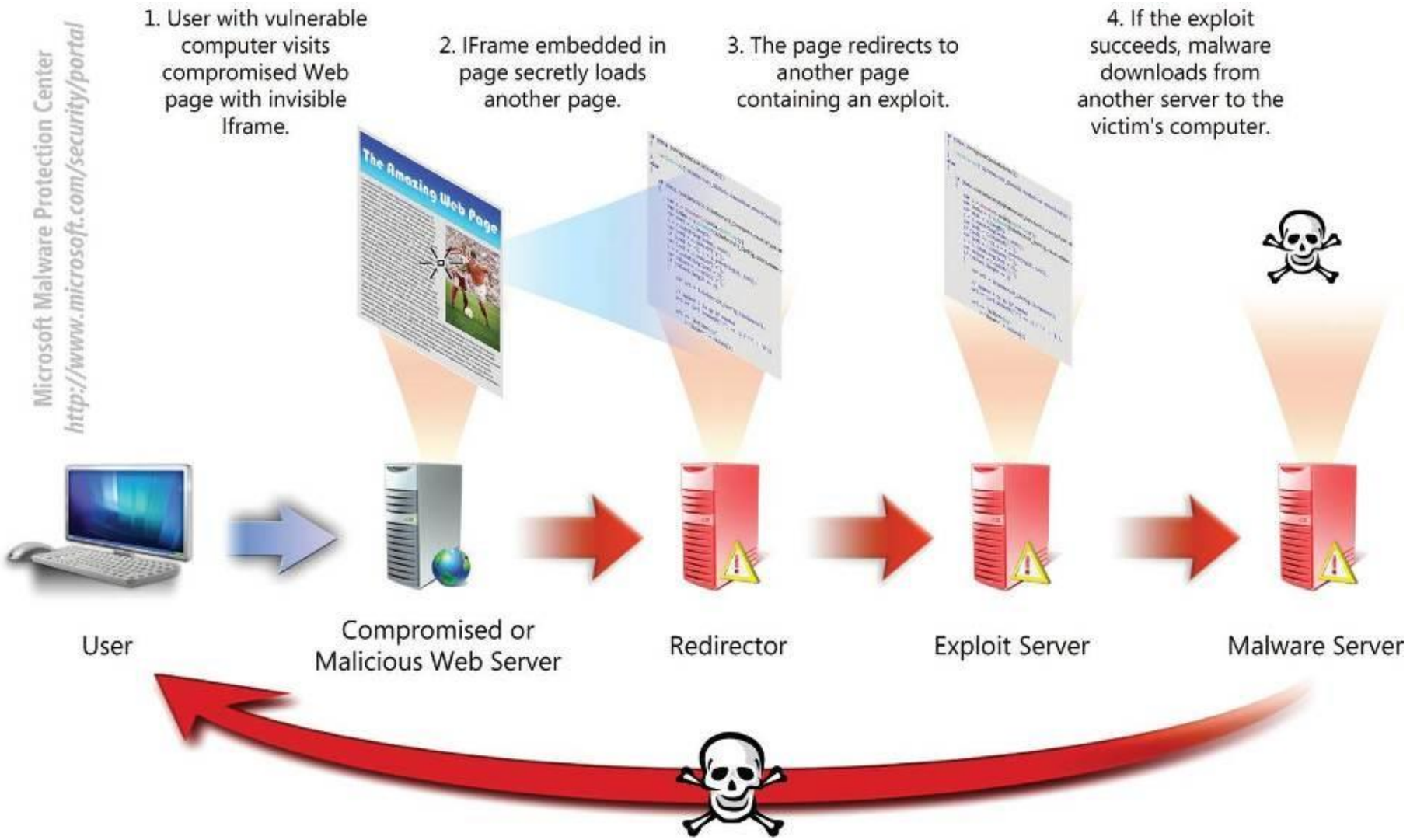



[Microsoft Security Intelligence Report, http://www.microsoft.com/security/sir]

# Drive-by downloads / installations

- A drive-by download site targets vulnerabilities in web browsers and browser add-ons (e.g., ActiveX)
- Users with vulnerable computers can be infected with malware ***automatically*** ***by visiting such a website***; they don't need to attempt to download anything.
  - 1.3% of Google queries - "All Your IFRAMEs Point to Us"
  - 1.5% of URLs - Moshchuk et al.
  - 5.3% of URLs - "Ghost Turns Zombie"
- Usually hosted on ***legitimate web sites*** compromised by
  - Intrusion
  - User-contributed content (cross-site scripting (XSS) very common)
- Exploit code usually hosted on a different website and exposed through compromised webpage using a technique like a URL embedded in malicious script code or an iFrame (HTML document embedded in another HTML document).

# Drive-by downloads



# Malicious websites

- Domains infected:
  - Games: 20%
  - Wallpaper: 9.6%
  - Celebrity: 7.6%
  - Kids: 1.6%
  - News: 0%
- Small number of spyware programs account for most infection
- Mostly adware, but 9.1% (May 2005) and 13% (Oct. 2005) Trojan downloader

| site | # infected executables |
|---|---|
| ***scenicreflections.com*** | 1,776 |
| ***screensaver.com*** | 191 |
| celebrity-wallpaper.com | 136 |
| ***screensavershot.com*** | 118 |
| download.com | 116 |
| ***gamehouse.com*** | 111 |
| ***galttech.com*** | 38 |
| ***appzplanet.com*** | 37 |
| megspace.com | 36 |
| download-game.com | 30 |

May 2005 crawl

[A. Moshchuk, T. Bragin, S.D. Gribble, and H.M. Levy. A Crawler-based Study of Spyware on the Web, NDSS, 2006.]

# 2013 MOST DANGEROUS CELEBRITIES™

## Lily Collins is the Most Dangerous Celebrity™

Lily Collins is this year's most dangerous celebrity to search for on the web. Daughter of rock musician Phil Collins, Lily quickly rose to fame as a talented actress, teenage journalist, and Hollywood trendsetter. She landed in the number-four spot in People magazine's 2012 World's Most Beautiful Women list, and has the lead in *The Mortal Instruments: City of Bones*.

## The Dangers of Online Searching

McAfee researched pop culture's most famous people to reveal the riskiest celebrity athletes, musicians, politicians, comedians, and Hollywood stars on the web. When you search for pictures and downloads of Lily Collins you have about a 14.5% chance of landing on a page that tested positive for spam, adware, spyware, viruses, or other malware.

**To better protect yourself on the web:**

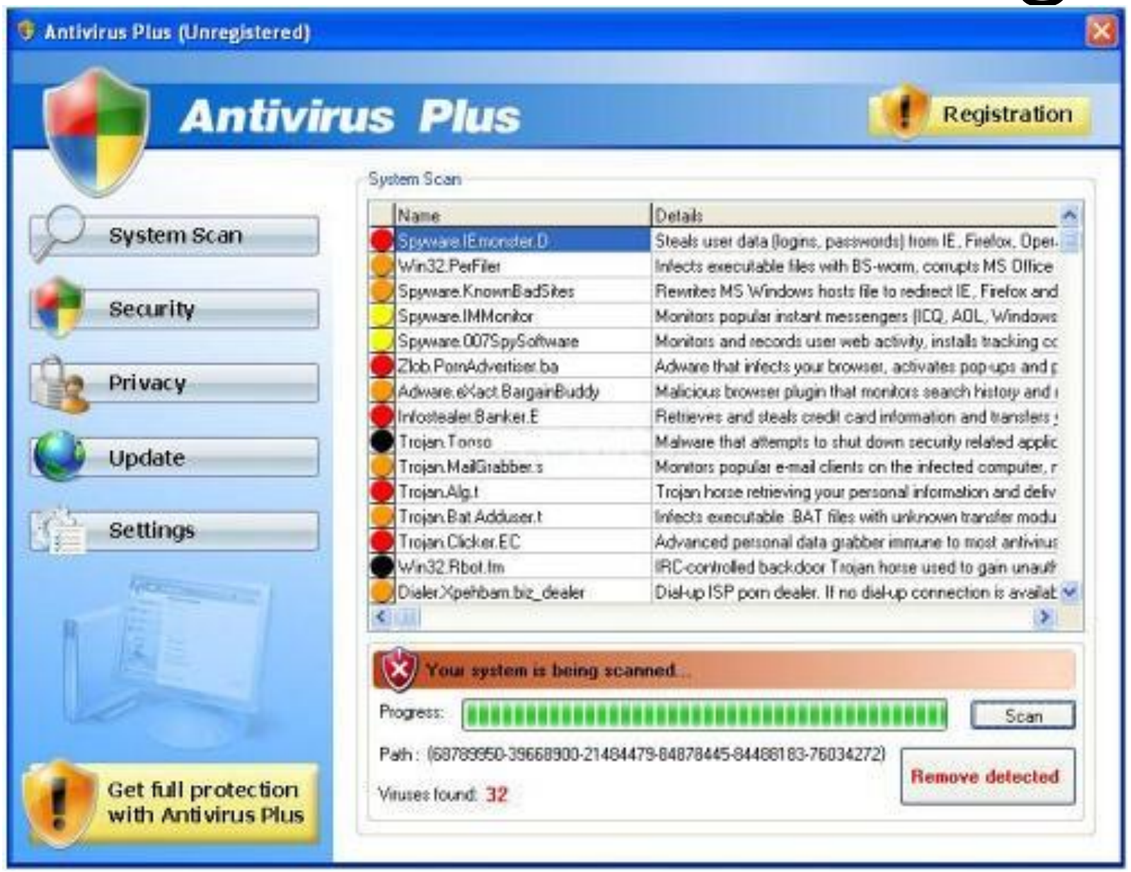
- Be wary of links to free content or too-good-to-be-true offers
- Be extra cautious when searching on hot topics, which often lead to fake and malicious sites created by cybercriminals
- Check the web address for misspellings or other clues that the link might lead to a phony website
- Protect yourself with comprehensive security, including a tool that identifies risky websites in search results

| | | |
|---|---|---|
| #1 | Lily Collins | 14.5 |
| #2 | Avril Lavigne | 12.7 |
| #3 | Sandra Bullock | 10.8 |
| #4 | Kathy Griffin | 10.6 |
| #5 | Zoe Saldana | 10.5 |
| #6 | Katy Perry | 10.4 |
| #7 | Britney Spears | 10.1 |
| #8 | Jon Hamm | 10.0 |
| #9 | Adriana Lima | 9.9 |

Cybercriminals' finger on the pulse of pop culture…

**Breaking news!!**

# So you'd better install anti-virus software, right?




- 2010 Google study found 11,000 domains hosting rogue anti-virus software
  - M.A. Rajab et al. The Nocebo Effect on the Web: An Analysis of Fake Anti-Virus Distribution. LEET, 2010.

# Mobile malware

- Still trailing PC-based malware in prevalence and sophistication
  - 0.26%-0.28% infection rate on Android (Truong et al., 2013)
- Device breakdown:
  - About 60% Android
  - About 40% Windows PCs connecting through mobile networks
  - <1% Windows Mobile, iPhones, Blackberrys, and Symbian devices
- Four of the 10 top threats are spyware, e.g.,
  - SMSTracker, which allows the attacker to remotely track and monitor all calls, SMS/MMS messages, GPS locations, and browser histories of an Android device

# When malware goes for breadth

# Slammer (Sapphire) worm

- January 24/25, 2003: UDP worm exploiting buffer overflow in Microsoft's SQL Server (port 1434)
  - Overflow was already known and patched by Microsoft… but not everybody installed the patch
- Entire code fits into a single 376-byte UDP packet
  - Could propagate extremely fast
- Random scanning: once control is passed to worm code, it randomly generates IP addresses and sends a copy of itself to port 1434

[Slide adapted by permission from Vitaly Shmatikov]

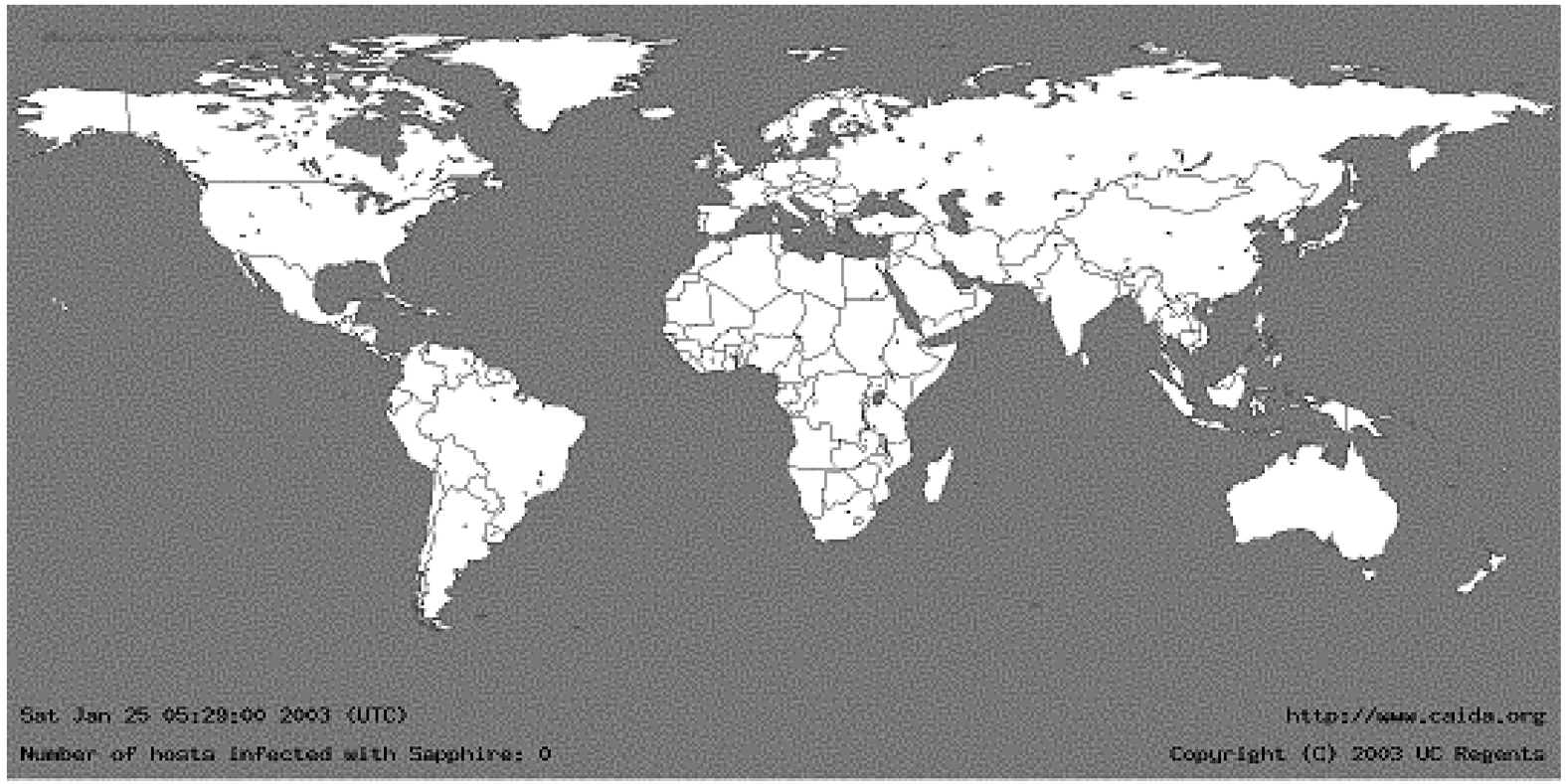
# Slammer Propagation

- Scan rate of 55,000,000 addresses per second
  - Scan rate = rate at which worm generates IP addresses of potential targets
  - Up to 30,000 single-packet worm copies per second
- Initial infection was doubling in 8.5 seconds (!!)
- Worm-generated packets saturated carrying capacity of the Internet in 10 minutes

[Slide adapted by permission from Vitaly Shmatikov]

# 05:29:00 UTC, January 25, 2003

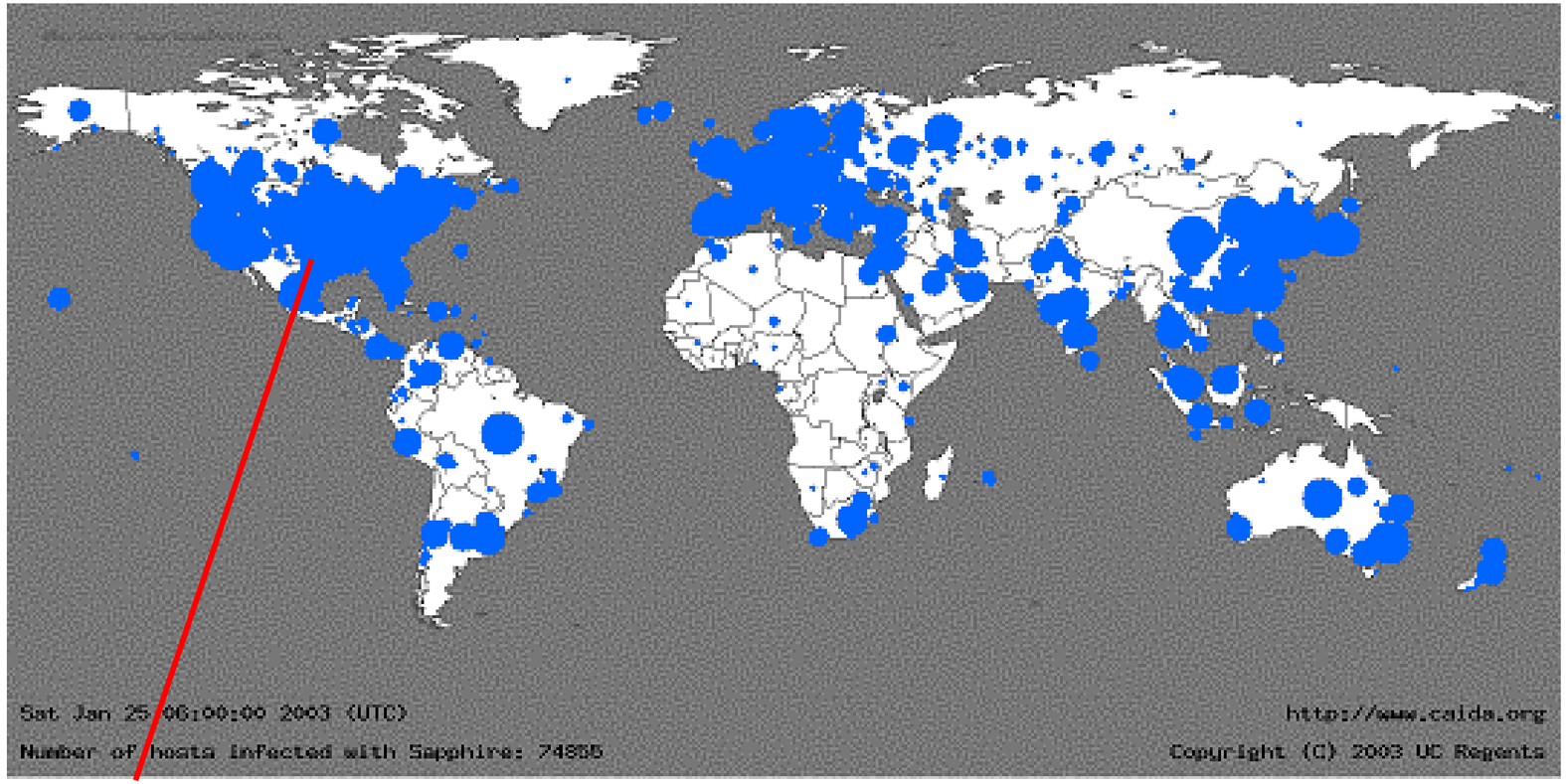
[from Moore et al. “The Spread of the Sapphire/Slammer Worm”]




[Slide adapted by permission from Vitaly Shmatikov]

# 30 minutes later

[from Moore et al. "The Spread of the Sapphire/Slammer Worm"]




Size of circles is **logarithmic** in the number of infected machines

[Slide adapted by permission from Vitaly Shmatikov]

# Impact of Slammer

- $1.25 billion of damage
- Temporarily knocked out many elements of critical infrastructure
  - Bank of America ATM network
  - Entire cell phone network in South Korea
  - Five root DNS servers
  - Continental Airlines' ticket processing software
- The worm did not even have malicious payload!
- It simply exhausted bandwidth on the network and CPU on infected machines.

[Slide adapted by permission from Vitaly Shmatikov]

# When malware goes for stealth

# Example: RSA attack

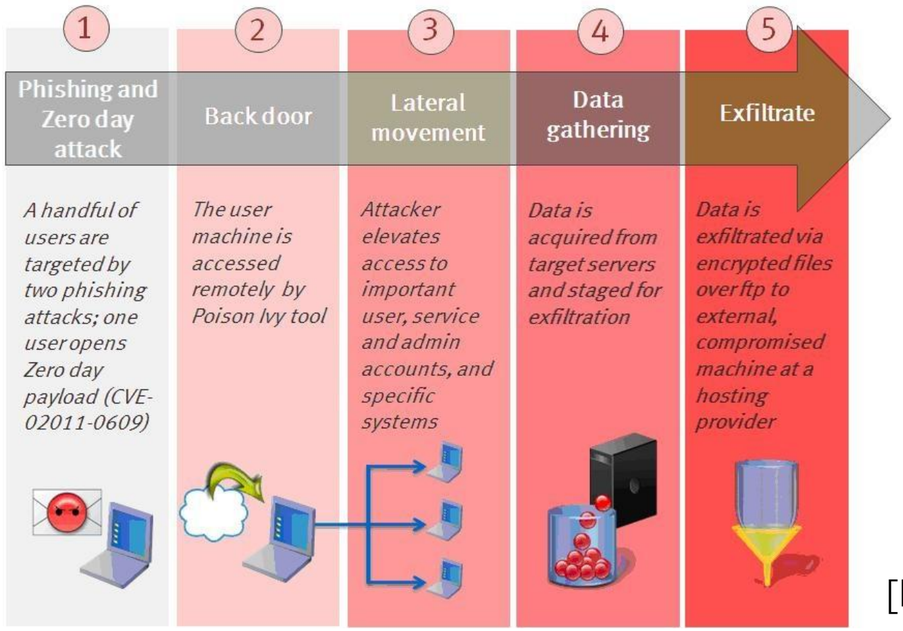
- Security company RSA attacked in 2011
  - Can't say what attackers took, but RSA had to reissue millions of tokens…
- Attack involved "spear-phishing"
  - Phishing (malicious e-mail) tailored specially to victim
  - Subject line: "2011 Recruitment Plan"
  - Contained spreadsheet "2011 Recruitment plan.xls"
  - Employee retrieved it from junk mail folder and opened it
- Payload mounts zero-day exploit against Adobe Flash (CVE-2011-0609)
  - Zero-day ***previously unknown*** vulnerability

RSA®
The Security Division of EMC


[From U. Rivner. "Anatomy of an Attack" blog, 1 April 2011]

# Example: RSA attack

- Exploit installed Poison Ivy Remote Administration Tool (RAT)
  - Remotely controllable by command and control (C&C) server
  - "Reverse-connect": Reaches out to C&C for extra stealth
- In general, classic Advanced Persistent Threat (APT)
  - Highly targeted, highly resourced "low" and "slow" attack against specific target




[From U. Rivner. "Anatomy of an Attack" blog, 1 April 2011]

# In-class exercise

How might you gather intelligence needed to craft spear phishing e-mail for a Stony Brook administrator? What would the e-mail look like?

# Who was responsible?

- Hint: Attacks typically took place between 9 a.m. and 5 p.m. Beijing time
- Domains used in attack
  - www. usgoodluck .com
  - obama .servehttp .com
  - prc .dynamiclink .ddns .us

# Who was responsible?

- China's People's Liberation Army (PLA) Unit 61398 near Shanghai
- Hacker "UglyGorilla"
  - Linked to Chinese national named Wang Dong
  - Blogged about experience as PLA hacker from 2006-09
  - "Low pay, long hours and instant ramen meals"
    - Sounds like a graduate student…
- Note the investigation spans computer code to classic espionage
  - We'll see more of this in studying cybercrime… next lecture

**NSA Chief: China Behind RSA Attacks**

**Chinese steal a "great deal" of military-related intellectual property, and were responsible for last year's attacks on cybersecurity company RSA, Gen. Keith Alexander tells Senators.**

[D. E. Sanger and N. Perlroth. NYTimes. Hackers From China Resume Attacks on U.S. Targets, 19 May 2013.]