Lecture 09: More Malware Defenses

Professor Adam Bates

CS 461 / ECE 422

Fall 2019

Goals for Today

- Learning Objectives:
  - Understand different methods of malware detection/defense
  - Articulate the tradeoffs between these systems
  - Consider heuristic defenses to ransomware
- Announcements, etc:
  - Difficulty access course slides? Switch browsers!
  - MP1 Checkpoint #2: **Due Sept 18 at 6pm**
  - MP2 will be released… **Sept 18 at 6pm**
    - Checkpoint #1 will be due on Sept 25 at 6pm

Reminder: Please put away devices at the start of class

Malware Detection

- Several strategies to consider:
  1. **Signatures**: What does the malware look like (i.e., code, data)?
  2. **Heuristics / Rules**: What does the malware do (i.e., runtime behaviors)?
  3. **Anomaly-based**: What does *normal behavior* look like?

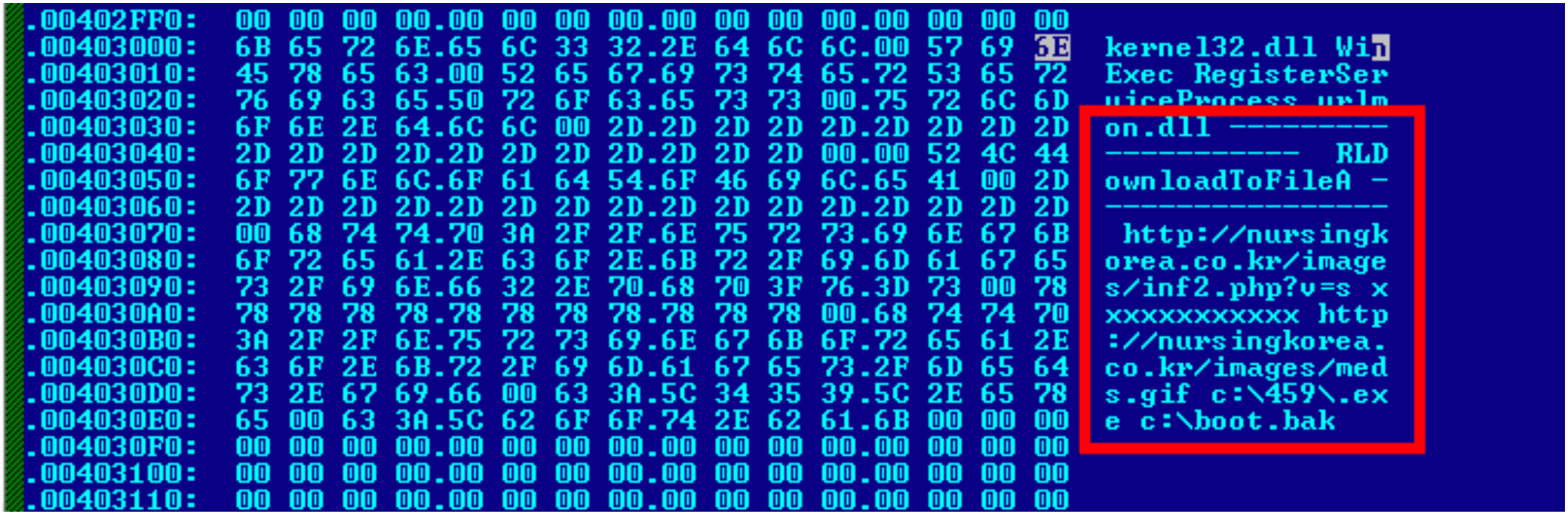
HIDS, AV

- Terminology
  - IDS: Intrusion detection system
  - IPS: Intrusion prevention system
  - HIDS/NIDS: Host/Network Based IDS
- Difference between IDS and IPS
  - Detection happens after the attack is conducted (i.e. the memory is already corrupted due to a buffer overflow attack)
  - Prevention stops the attack before it reaches the system (i.e. shield does packet filtering)
- Anomaly, heuristic, behavior-based vs. Misuse, Rule-based, Signature-based

Signatures: A Malware Countermeasure

- Scan compare the analyzed object with a database of signatures
- A signature is a virus fingerprint
  - E.g.,a string with a sequence of instructions specific for each virus
  - Different from a digital signature
- A file is infected if there is a signature inside its code
  - Fast pattern matching techniques to search for signatures
- All the signatures together create the malware database that usually is proprietary

Signatures: A Malware Countermeasure



Signatures: A Malware Countermeasure

The following is a list of updates and amendments tothe Virus Bulletin Table of Known IBM PC Viruses asof 15 December 1998. Each entry consists of the virusname, its aliases (if any) and the virus type. This isfollowed by a short description (if available) and a24-byte hexadecimal search pattern to detect thepresence of the virus with a disk utility or a dedicatedscanner which contains a user-updatable pattern library

Prodigy.268

CN: An overwriting, 268-byte direct infector containing the texts '*.COM', 'Pr0diGy VeEr0oZ (c) 1995' and 'HaPpY nEw YeAR! SeE U iN HeLL...'.

```
Prodigy.268          890E C301 8916 C101 BA00 01B4 40B9 0C01 CD21 90B8 0157 8B16
```

Saha.2382

CR: An overwriting, 2382-byte virus which infects COM files and modifies EXE files with the same name as COM targets. The virus appends nine bytes (the string ' Sahand ') to EXE programs.

```
Sahand.2382          B918 048D 1618 058B 1EBF 09B4 40CD 21B9 8F00 8D16 3009 8B1E
```

Simple.331

CN: An encrypted, appending, 331-byte, direct infector with the text '*.COM'. Infected programs ends with the string 'SIMPLE'.

```
Simple.331           60E8 0000 5E81 EE32 01B9 2E01 2EF6 1446 E2FA 61C3 5349 4D50
```

Spartak.1360

CEN: An encrypted, appending, 1360-byte virus containing the texts 'COWIIBAIAVDRWEADHICH', 'Spartak Virus by Crazy Punk (C) | v1.0 beta', 'Moscow, Russia, 06/10/1998', 'F_C_S_M.COM', '*.com' and '*.exe'. The last two bytes are XOR-ed together give the value of 0FFh.

```
Spartak.1360         B919 031E 06E8 0000 FA5D 81ED 0E01 0E1F BA40 0052 07B8 ????
```

Spartak.1453

CEN: An encrypted, appending, 1453-byte virus with the texts 'COWIIBAIAVDRWEADHICH', 'Spartak-II Virus by Crazy Punk (C)', 'Moscow, Russia, 06/10/1998', '*.com' and '*.exe'. The last two bytes are XOR-ed together give the value of 0FFh.

```
Spartak.1453         1EB9 4804 06BA 4000 E800 00FA 5D81 ED11 010E 1F52 07B8 ????
```

Spy.447

CN: A 447-byte appender with the texts 'host.com', 'Opening file:', 'Unable to open file.', 'Storing first three bytes:', 'Storing file size...', 'Appending virus code...' and 'Setting jump to virus code...'.

```
Spy.447              B440 B9BF 018D 9600 01CD 21C3 8D9E 9E02 E82D 008B 8601 012D
```

Variola

MDR: An boot sector virus which infects MBRs on hard disks and DOS Boot Sectors on diskettes. It has the encrypted text 'PeaceMaker by VaRiOLa'. The virus stores the original boot sectors encrypted.

```
Variola              8BD9 D1E9 4B8A 248A 0032 E132 C126 8805 2688 2146 474B E2EC
```

Wild.2406

CER: An appending, 2406-byte virus.

```
Wild.2406            B873 0BBB 7373 CD21 80FC 7374 03E9 6B08 0E58 1E5B 2BC3 7518
```

XM.2401

CE: A polymorphic, 2401-byte appender with the texts '[XyeBo_MHe], (c)MidnighÅPr0wler -=Version', 'COMMAND.COMDOS4GW.EXEIBMBIO.COMCOMEXEcomexe' and 'c:\autoexec.bat'. Infected files have the word E958h at offset 0000h (COM) and the word FAFAh at offset 0010h (EXE). The following template detects the virus in memory only.

```
XM.2401              B961 0953 E80E FE5B 2E89 0E1E  0406 1FB4 BFBA 1C00 E862 FAE
```

https://www.virusbulletin.com/uploads/pdf/magazine/1999/199901.pdf

Coarse Signatures: White/Black Listing

- Maintain database of cryptographic hashes for
  - Operating system files
  - Popular applications
  - Known infected files
- Compute hash of each file
- Look up into database
- Needs to protect the integrity of the database

Heuristic / Rule-Based Analysis

- Based on what it DOES
- Useful to identify new and "zero day" malware
- Code analysis; what code MIGHT do
  - Based on the instructions, the antivirus can determine whether or not the program is malicious, i.e., program contains instruction to delete system files,
- Execution emulation; what code ACTUALLY does
  - Run code in isolated emulation environment
  - Monitor actions that target file takes
  - If the actions are harmful, mark as virus
- Heuristic methods can trigger false alarms
  - E.g., Ransomware versus compression, full disk encryption

Example Heuristics / Rules

Network	Process	Files	Registry
connects to 80	execs cmd.exe	writes winhlp32.dat	uses wininet.dll
connects to 25	execs IEXPLORE.EXE	writes tasklist32.exe	uses PRNG
connects to 6667	execs regedit.exe	writes change.log	modifies registered applications
connects to 587	execs tasklist32.exe	writes mirc.ini	modifies proxy settings
scans port 80	execs svchost.exe	writes svchost.exe	modifies mounted drives

Heuristic / Rule-Based Analysis

- Extremely prevalent in commercial endpoint security monitors.

Sysdig splunk> Symantec

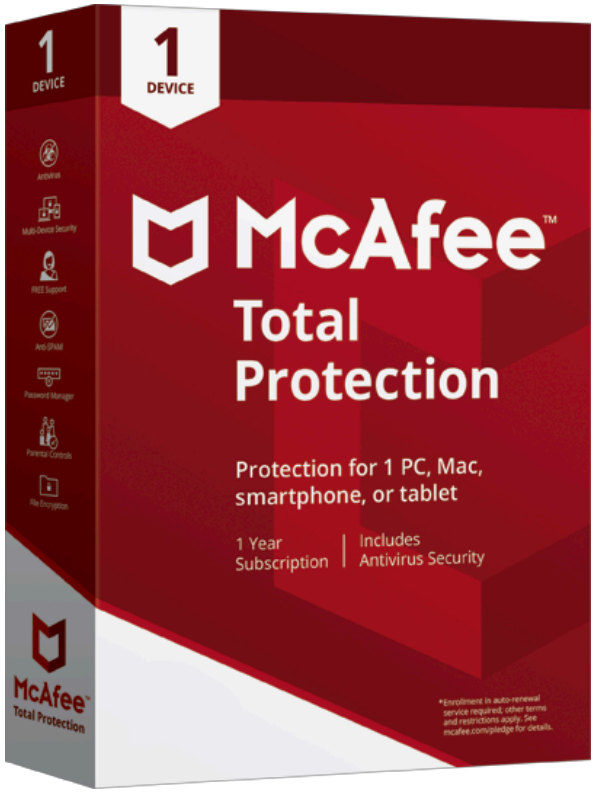
- Even the most advanced tools are prone to high false alert rates, e.g., …

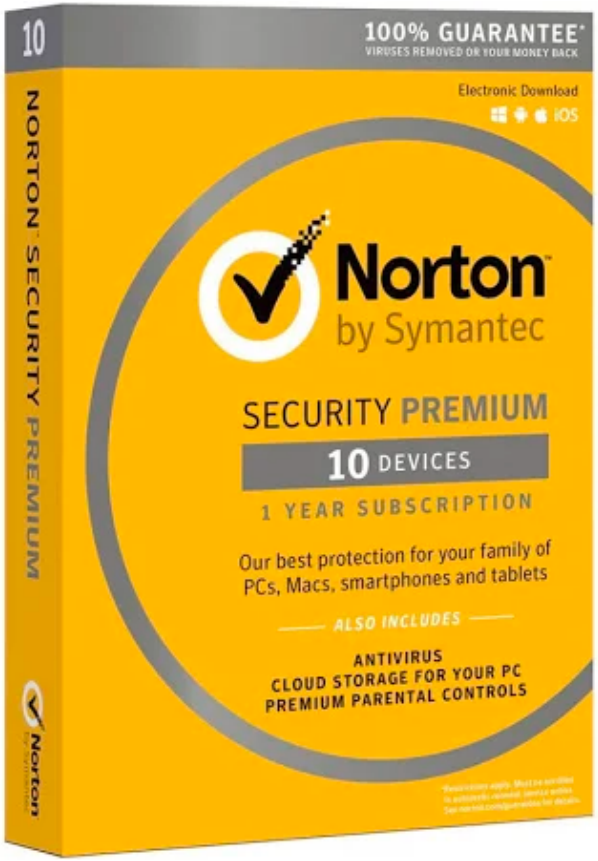
FIREEYE™

FireEye's "How Many Alerts is Too Many to Handle?" report:

- **US orgs receive 17,000 alerts per month on average**
- **51% false alarms**
- **Only 4% of alerts are properly investigated.**

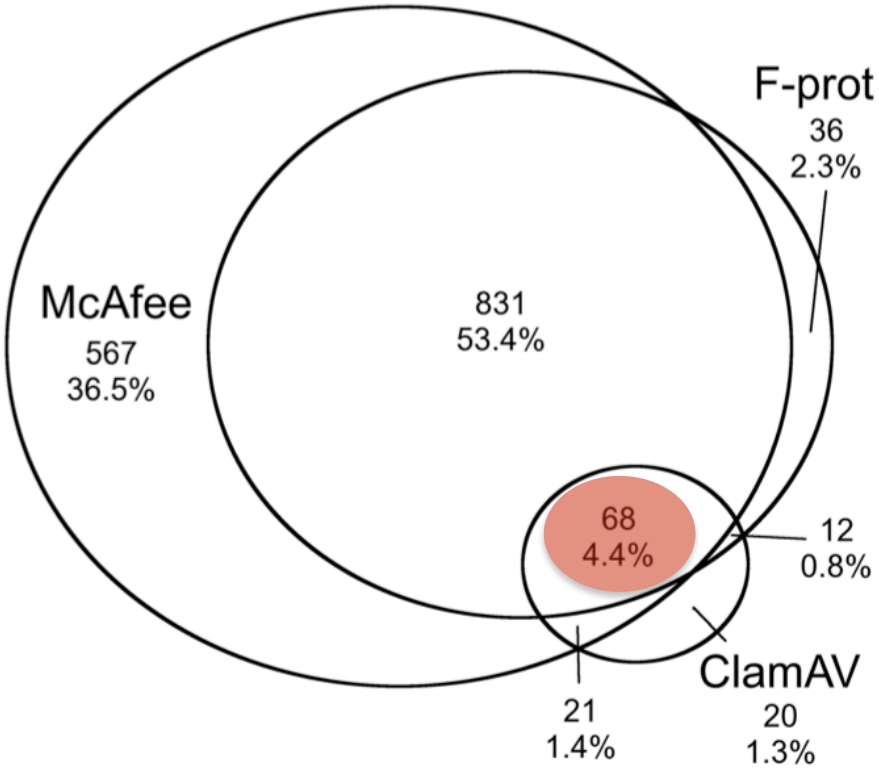
Antivirus


1
DEVICE
1
DEVICE
McAfee™
Total
Protection
Protection for 1 PC, Mac,
smartphone, or tablet
1 Year
Subscription
Includes
Antivirus Security
McAfee
Total Protection


10
100% GUARANTEE
VIRUSES REMOVED OR YOUR MONEY BACK
NORTON SECURITY PREMIUM
Electronic Download
iOS
Norton
by Symantec
SECURITY PREMIUM
10 DEVICES
1 YEAR SUBSCRIPTION
Our best protection for your family of
PCs, Macs, smartphones and tablets
ALSO INCLUDES
ANTIVIRUS
CLOUD STORAGE FOR YOUR PC
PREMIUM PARENTAL CONTROLS

SDBot

- Are different antivirus softwares making consistent claims about malware?
- SDBot variants as classified by McAfee, ClamAV, F-Prob

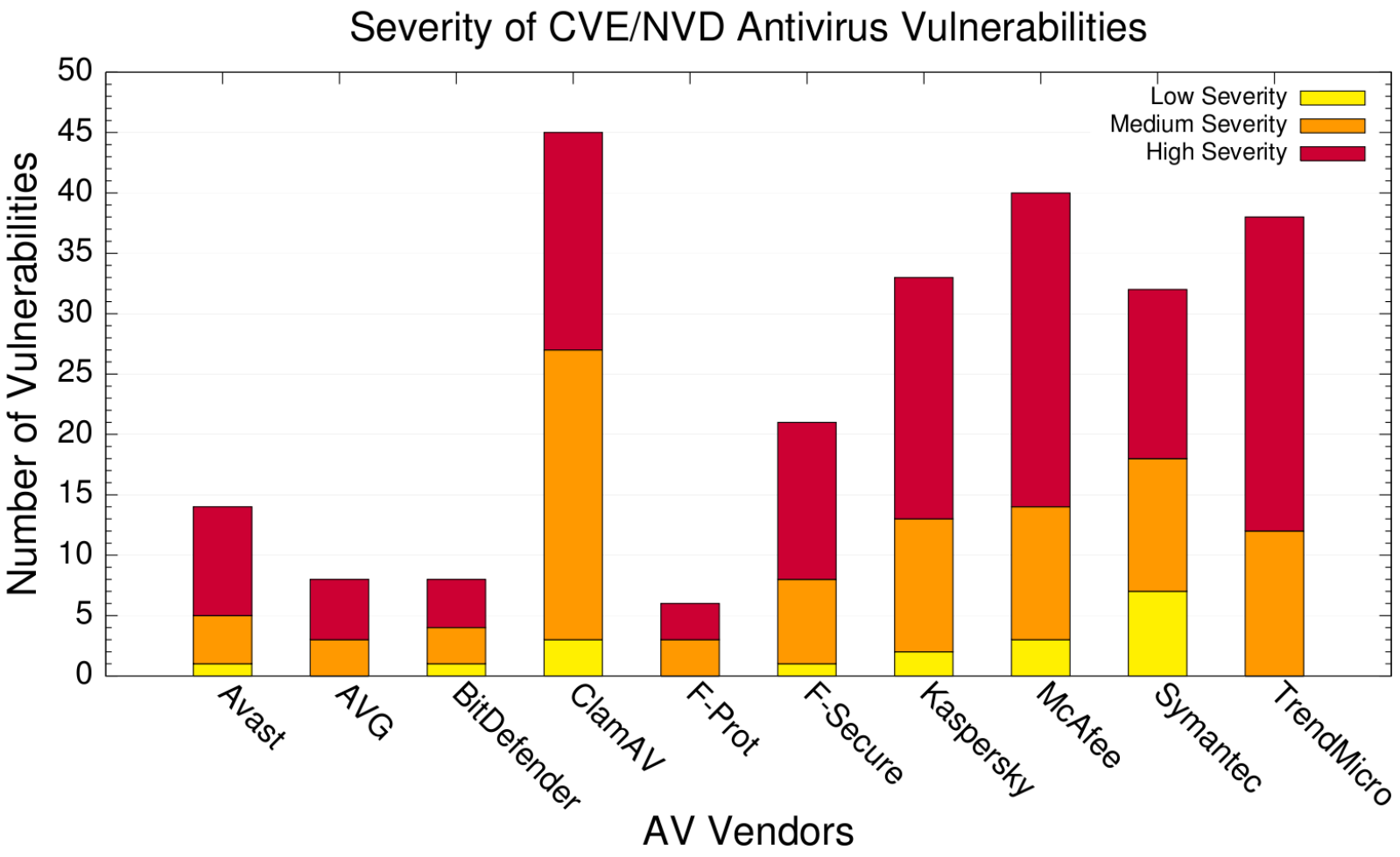


Completeness, Accuracy

- The percentage of malware samples detected across datasets and AV vendors
- ***AV system labels are incomplete***

Dataset	AV Updated	Percentage of Malware Samples Detected				
		McAfee	F-Prot	ClamAV	Trend	Symantec
legacy	20 Nov 2006	100	99.8	94.8	93.73	97.4
small	20 Nov 2006	48.7	61.0	38.4	54.0	76.9
small	31 Mar 2007	67.4	68.0	55.5	86.8	52.4
large	31 Mar 2007	54.6	76.4	60.1	80.0	51.5

Antivirus Vulnerabilities



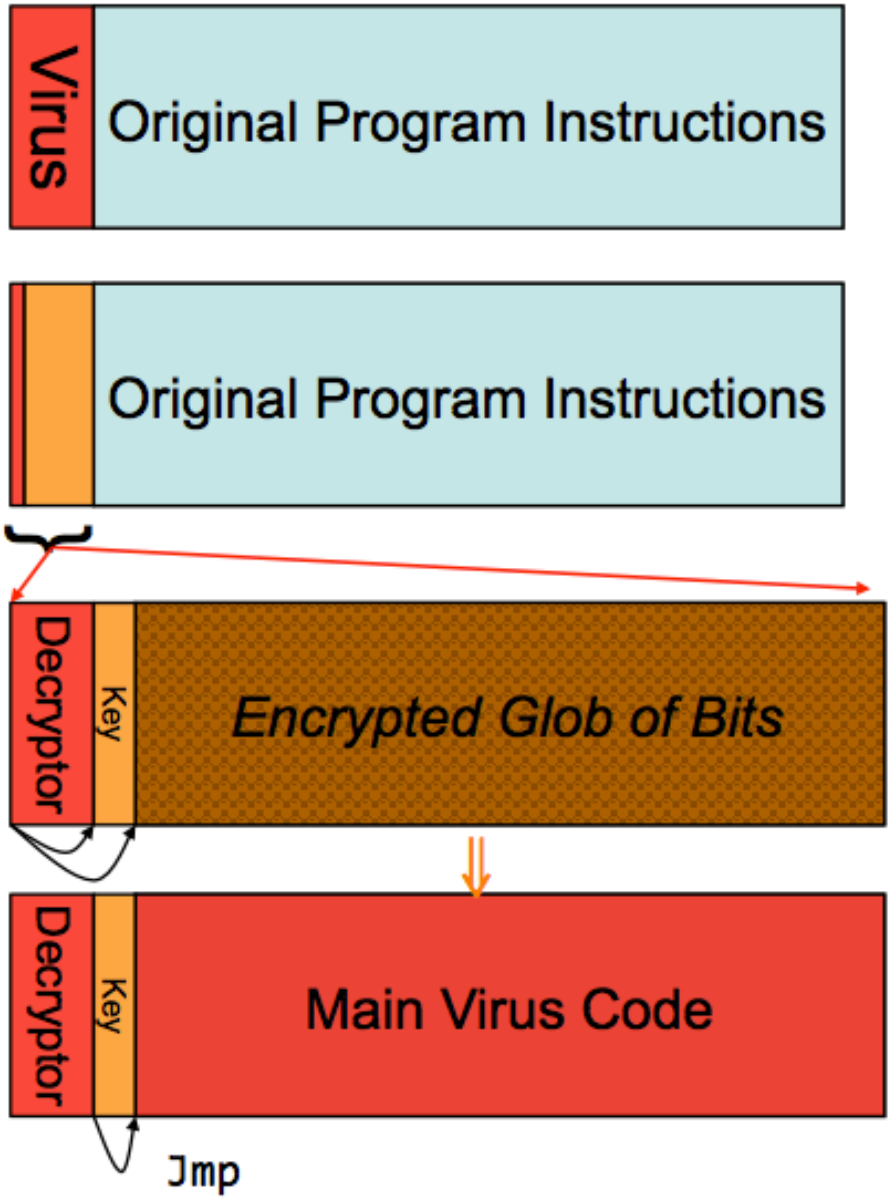
Antivirus engines vulnerable to numerous local and remote exploits

(number of vulnerabilities reported in NVD from Jan. 2005 to Nov. 2007)

Concealment

- Encrypted virus
  - Decryption engine + encrypted body
  - Randomly generate encryption key
  - Detection looks for decryption engine
- Polymorphic virus
  - Encrypted virus with random variations of the decryption engine (e.g., padding code)
  - Detection using CPU emulator
- Metamorphic virus
  - Different virus bodies
  - Approaches include code permutation and instruction replacement
  - Challenging to detect

Encrypted Virus



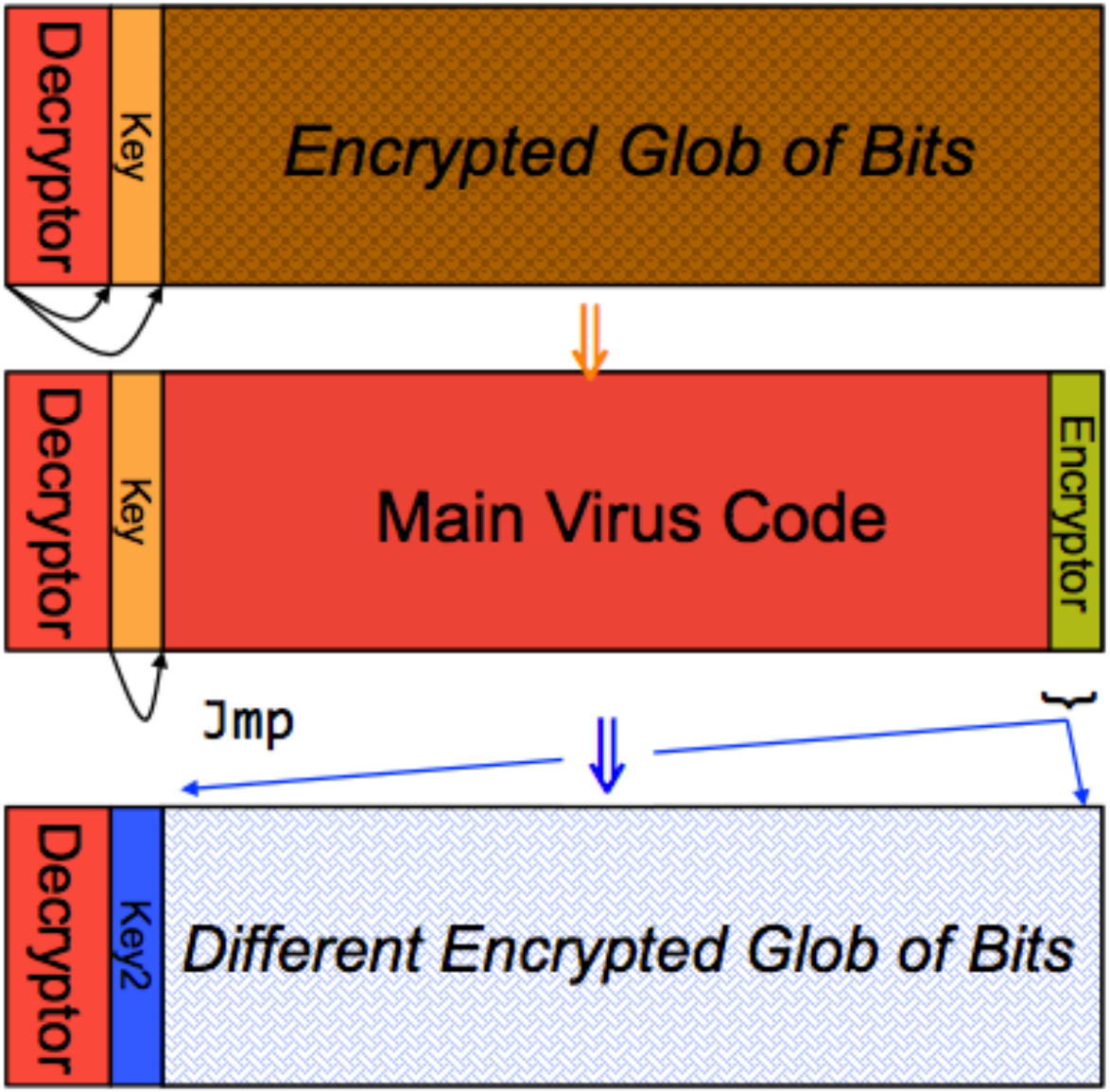
Instead of this …

Virus has *this* initial structure

When executed, decryptor applies key to decrypt the glob …

… and jumps to the decrypted code once stored in memory

Polymorphic Virus



Once running, virus uses an *encryptor* with a new key to propagate

New virus instance bears little resemblance to original

Arms Race: Polymorphic Code

- Given polymorphism, how might we then detect viruses?
- Idea #1: use narrow sig. that targets decryptor
  - Issues?
    - Less code to match against " more false positives
    - Virus writer spreads decryptor across existing code
- Idea #2: execute (or statically analyze) suspect code to see if it decrypts!
  - Issues?
    - Legitimate "*packers*" perform similar operations (decompression)
    - How long do you let the new code execute?
      - If decryptor only acts after lengthy legit execution, difficult to spot

Metamorphic Code

- Idea: every time the virus propagates, generate *semantically* different version of it!
  - Different semantics only at immediate level of execution; higher-level semantics remain same
- How could you do this?
- Include with the virus a code rewriter:
  - Inspects its own code, generates random variant, e.g.
  - Renumber registers
  - Change order of conditional code
  - Reorder operations not dependent on one another
  - Replace one low-level algorithm with another
  - Remove some do-nothing padding and replace with different do- nothing padding ("chaff")

Detecting Metamorphic Viruses?

- Need to analyze execution behavior
  - Shift from syntax (*appearance* of instructions) to semantics (*effect* of instructions)
- Two stages: (1) AV company analyzes new virus to find behavioral signature; (2) AV software on end systems analyze suspect code to test for match to signature
- What countermeasures will the virus writer take?
  - Delay analysis by taking a long time to manifest behavior
    - Long time = await particular condition, or even simply clock time
  - Detect that execution occurs in an analyzed environment and if so behave differently
    - E.g., test whether running inside a debugger, or in a Virtual Machine
- Counter-countermeasure?
  - AV analysis looks for these tactics and skips over them
- Note: attacker has edge as *AV products supply an* ***oracle!***

Anomaly-Based HIDS

- Idea behind HIDS
  - Define normal behavior for a process
    - Create a model that captures the behavior of a program during normal execution.
  - Monitor the process
    - Raise a flag if the program behaves abnormally

Why System Calls?

- The program is a layer between user inputs and the operating system
- A compromised program cannot cause significant damage to the underlying system without using system calls
- i.e Creating a new process, accessing a file etc.

Syscall N-Grams

- Forrest et. al. A Sense of Self for Unix Processes, 1996.
- Tries to define a normal behavior for a process by using sequences of system calls.
- As the name of their paper implies, they show that fixed length short sequences of system calls are distinguishing among applications.
- For every application a model is constructed and at runtime the process is monitored for compliance with the model.
- *Definition:* The list of system calls issued by a program for the duration of it's execution is called a *system call trace*.

N-Gram: Building the Model by Training

- Slide a window of length N over a given system call trace and extract unique sequences of system calls.

Example:

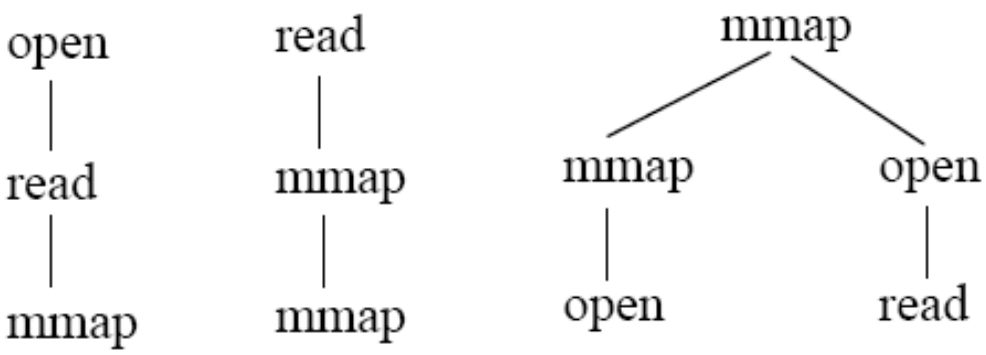
open, read, mmap, mmap, open, read, mmap

System Call trace

Unique Sequences

open, read, mmap
read, mmap, mmap
mmap, mmap, open
mmap, open, read

Database



N-Gram: Monitoring

- Monitoring
  - A window is slid across the system call trace as the program issues them, and the sequence is searched in the database.
  - If the sequence is in the database then the issued system call is valid.
  - If not, then the system call sequence is either an intrusion or a normal operation that was not observed during training (false positive) !!

Experimental Results for N-Gram

- Databases for different processes with different window sizes are constructed
- A normal sendmail system call trace obtained from a user session is tested against all processes databases.
- The table shows that sendmail's sequences are unique to sendmail and are considered as anomalous by other models.

	5		6		11	
Process	%	#	%	#	%	#
sendmail	0.0	0	0.0	0	0.0	0
ls	6.9	23	8.9	34	13.9	93
ls -l	30.0	239	32.1	304	38.0	640
ls -a	6.7	23	8.3	34	13.4	93
ps	1.2	35	8.3	282	13.0	804
ps -ux	0.8	45	8.1	564	12.9	1641
finger	4.6	21	4.9	27	5.7	54
ping	13.5	56	14.2	70	15.5	131
ftp	28.8	450	31.5	587	35.1	1182
pine	25.4	1522	27.6	1984	30.0	3931
httpd	4.3	310	4.8	436	4.7	824

The table shows the number of mismatched sequences and their percentage wrt the total number of subsequences in the user session

Anomaly-Based Analysis

- False alarm rates are even worse than rule-based!
- Like rule-based, alerts for unexpected activity…
  - e.g., administrator diagnostics?
- Unlike rule-based, also problems of concept drift and training coverage!!
  - i.e., what happens when the way we use our systems slowly changes?
  - i.e., what happens if fail to properly train our models?
- Deployed in more limited fashion in security products

What about ransomware?

Maine Police Pay Ransomware Demand in Bitcoin

BY STEPHANIE MLOT APRIL 14, 2015 12:55PM EST 8 COMMENTS

The Lincoln County Sheriff's Office and four town police departments were infected with the "megacode" virus.

0 SHARES

In an effort to keep their computer files from being destroyed, a group of cooperative police departments in Maine paid a $300 ransom demand—in bitcoin.

According to local news station WCSH-TV, the

Police pay ransom after cyberterror attack on network

Story Comments (1) Print Font Size:

Posted: Saturday, April 4, 2015 10:27 am

By Jayne W. Miller News Editor
Jayne@YourTownCrier.com | 1 comment

Chief: "Paying ransom was the last resort"

TEWKSBURY – Last December Tewksbury Police confronted a new, and growing, frontier in cyberterrorism when the CryptoLocker ransomware virus infected the department's network, encrypting essential department files until the town paid a $500 bitcoin ransom. In total, police systems were down between four and five days as the department worked with the FBI, Homeland Security,

Hospital pays $17k for ransomware crypto key

Hollywood Presbyterian says systems were restored after 10-day lockout.

by **Sean Gallagher** - Feb 18, 2016 10:17am EST

Share Tweet Email 141

University of Calgary paid $20K in ransomware attack

No evidence cyberattackers released personal or university data to public

CBC News Posted: Jun 07, 2016 2:27 PM MT | Last Updated: Jun 08, 2016 8:26 AM MT

What about ransomware?

Kansas Heart Hospital hit with ransomware; attackers demand two ransoms

Credit: Shutterstock

Kansas Heart Hospital was hit with a ransomware attack. It paid the ransom, but then attackers tried to extort a second payment.

RELATED

Paying ransomware is what ills some hospitals

4 reasons not to pay up in a ransomware attack

Got ransomware? These tools may help

on IDG **Answers**

What is a 'watering hole' attack?

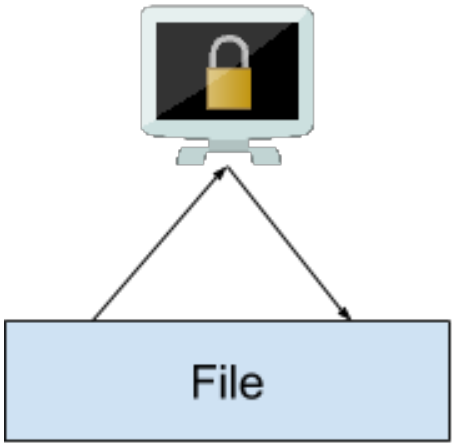
Efficacy of Malware Defenses?

- Signature-based?
  - Malware knows how to defeat…
- Rule-based?
  - Ransomware targets commonly-accessed files
  - Other programs have similar access patterns
- Anomaly-based?
  - Other programs have similar access patterns
  - Anomaly detection is reactive; how much damage is done prior to detection?

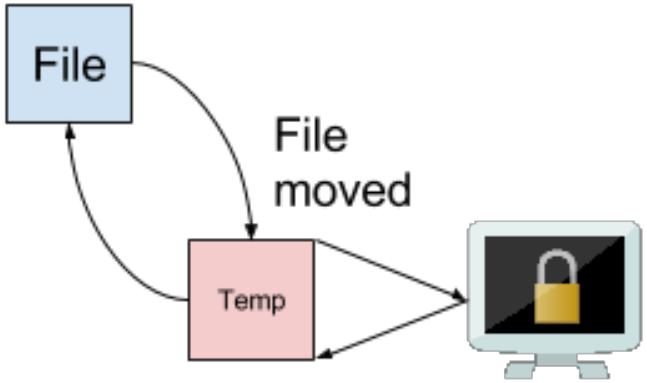
It's about the data!

- Define matching rules not on access patterns (i.e., sys calls), but on the data manipulations themselves.
- Ransomware must:
  - Read data
  - Write obfuscated data
  - Dispose of original data
- Ransomware may:
  - Use network access
  - Display ransom messages
  - Access files linearly or quickly

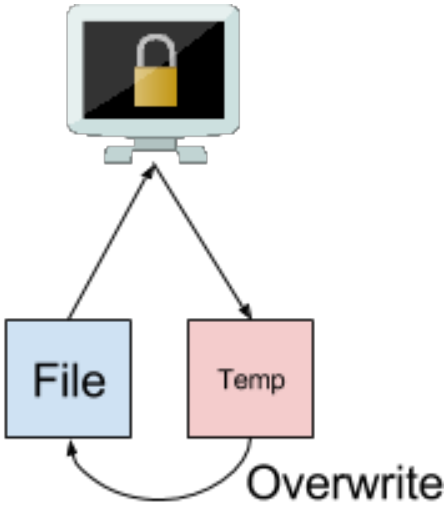
Ransomware Workflows



Class A



Class B



Class C

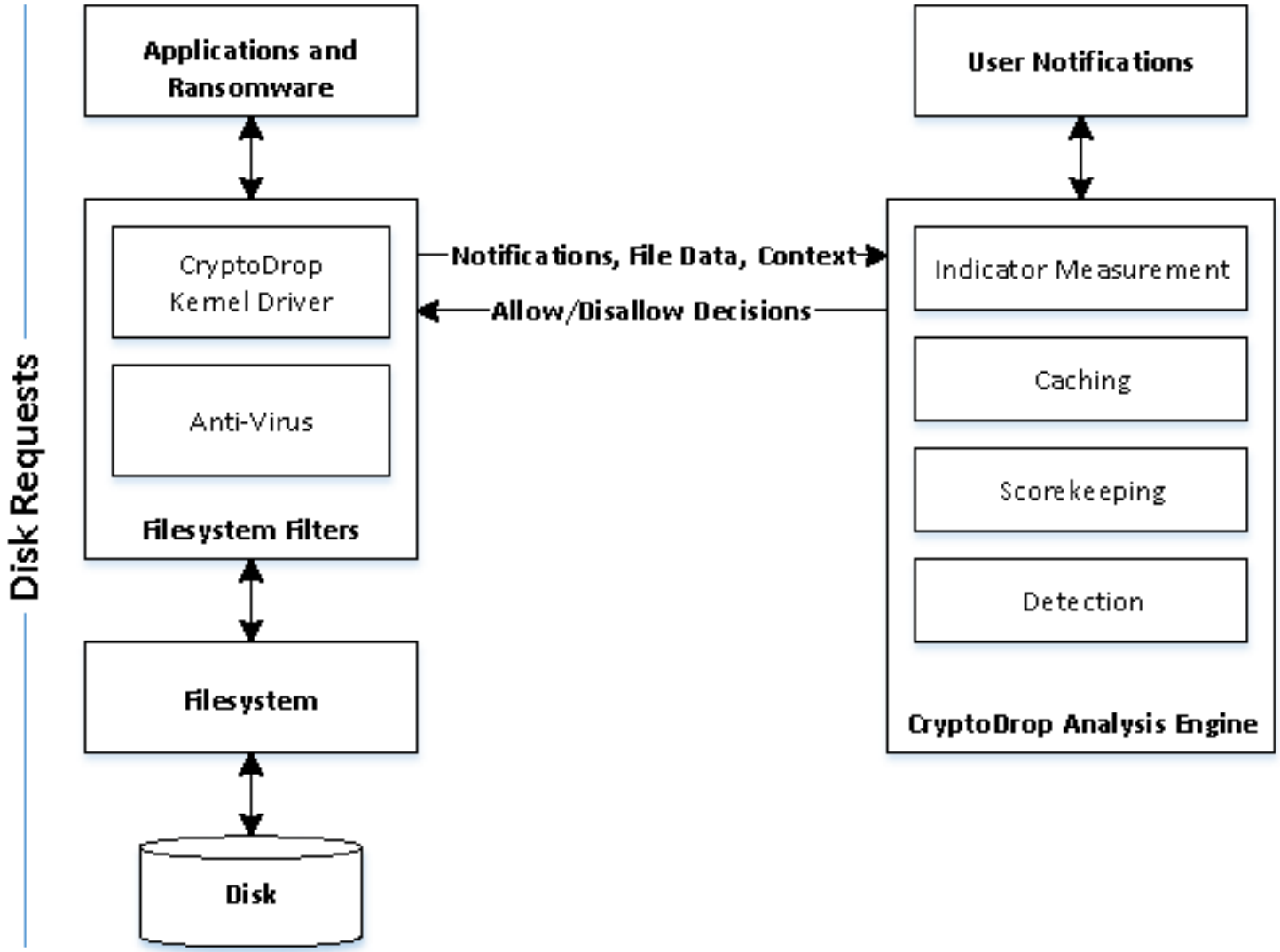
Case Study: CryptoDrop

- CryptoDrop is a research artifact **turned** commercial ransomware detection system.
- Provides early-warning ransomware detection by…
  - Mediating filesystem reads/writes
  - Monitoring I/O data for transformative changes
  - Tracking when changes exceed thresholds

CryptoDrop

[Scaife et al., ICDCS'16]

CryptoDrop Overview



[Scaife et al., ICDCS'16]

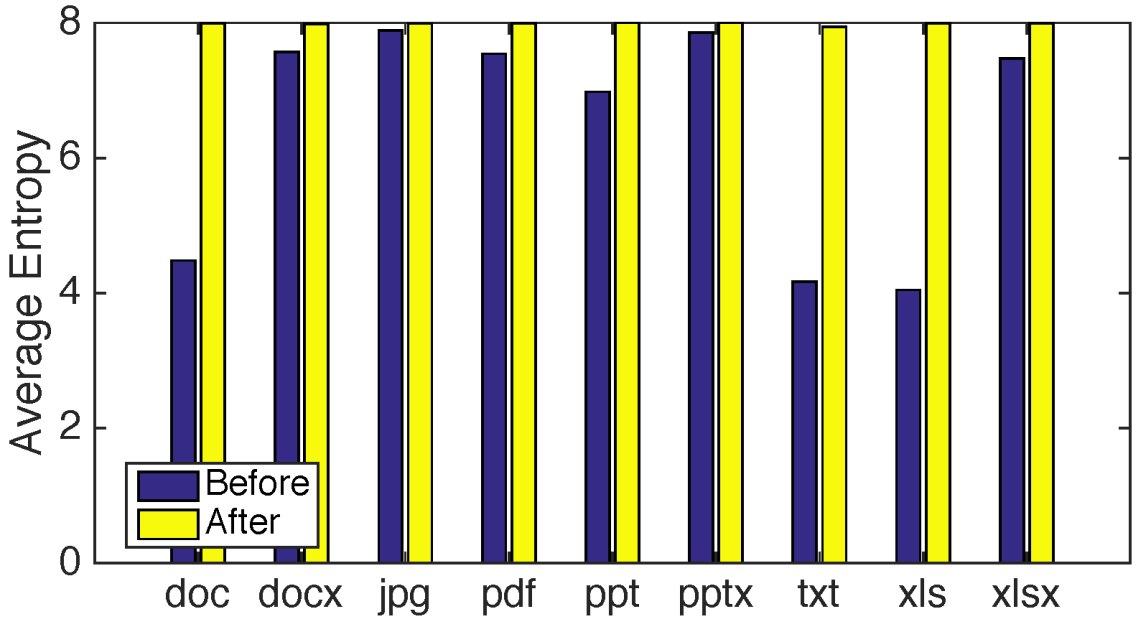
Ransomware Indicators

- Ransomware must encrypt/obfuscate data, must erase/ overwrite original data.
- **Observation**: Original data has MUCH lower entropy than encrypted data!!
- CryptoDrop measures a weighted arithmetic mean of the Shannon entropy over file I/O.
- An indicator flag marks programs as suspicious if they consistently *write* more entropy than they *read*.

[Scaife et al., ICDCS'16]

Ransomware Indicators

- Entropy measurement sounds perfect! How effective?
- File Type entropies before/after encryption:



- Why do so many file types have high 'before' entropy??

[Scaife et al., ICDCS'16]

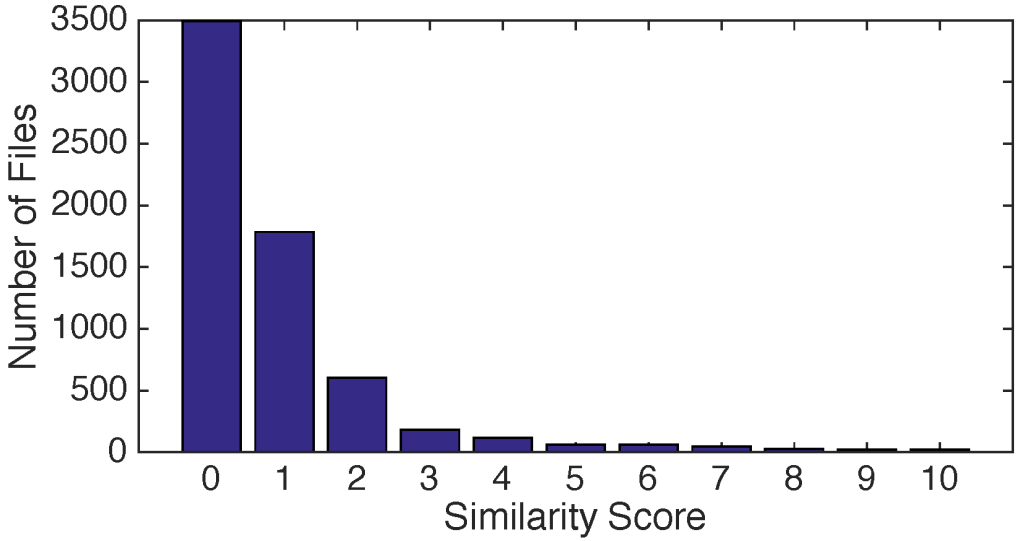
Ransomware Indicators

- **Observation**: File types often imply a data format; ransomware may produce data that does not match this format.
- CryptoDrop checks specific byte values to see if they match a signature for a the expected file type
- An indicator flags is flipped any time the file's measured format deviates from the expected format.

[Scaife et al., ICDCS'16]

Ransomware Indicators

- **Observation**: Many programs (e.g., text editor) modify files incrementally, leaving much of the data unchanged from session to session.
- CryptoDrop leverages similarity-preserving hashes of files before and after I/O sessions to detect wild variations in content similarity.



[Scaife et al., ICDCS'16]

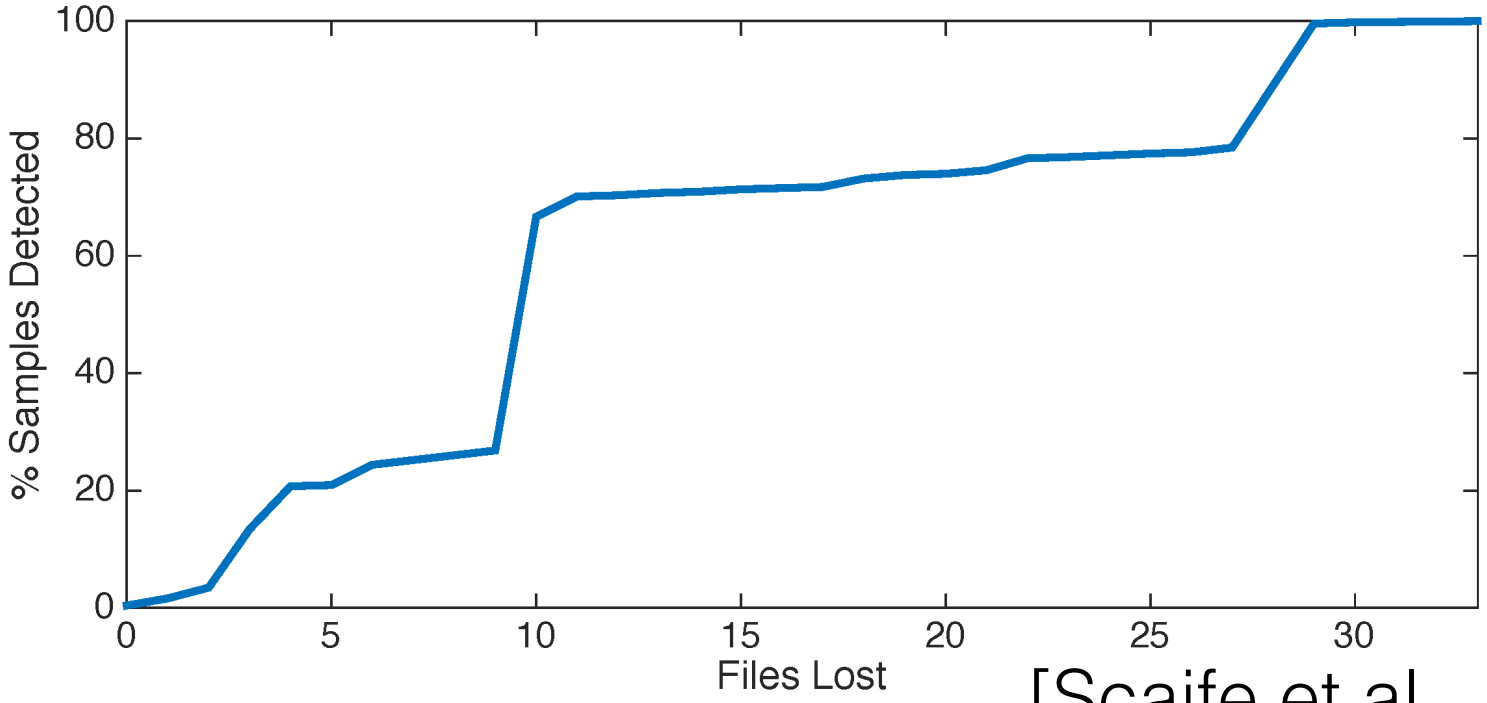
Scoring & Detection

- CryptoDrop maintains a score for each process based on a composite of the individual indicators.
- If the total score exceeds a provided threshold –or– all three primary indicators have contributed to the process's score:
  - Pauses all filesystem activity on user documents
  - Automatically block program, or prompt user for authorization decision

[Scaife et al., ICDCS'16]

Efficacy (ICDCS'16 Version)

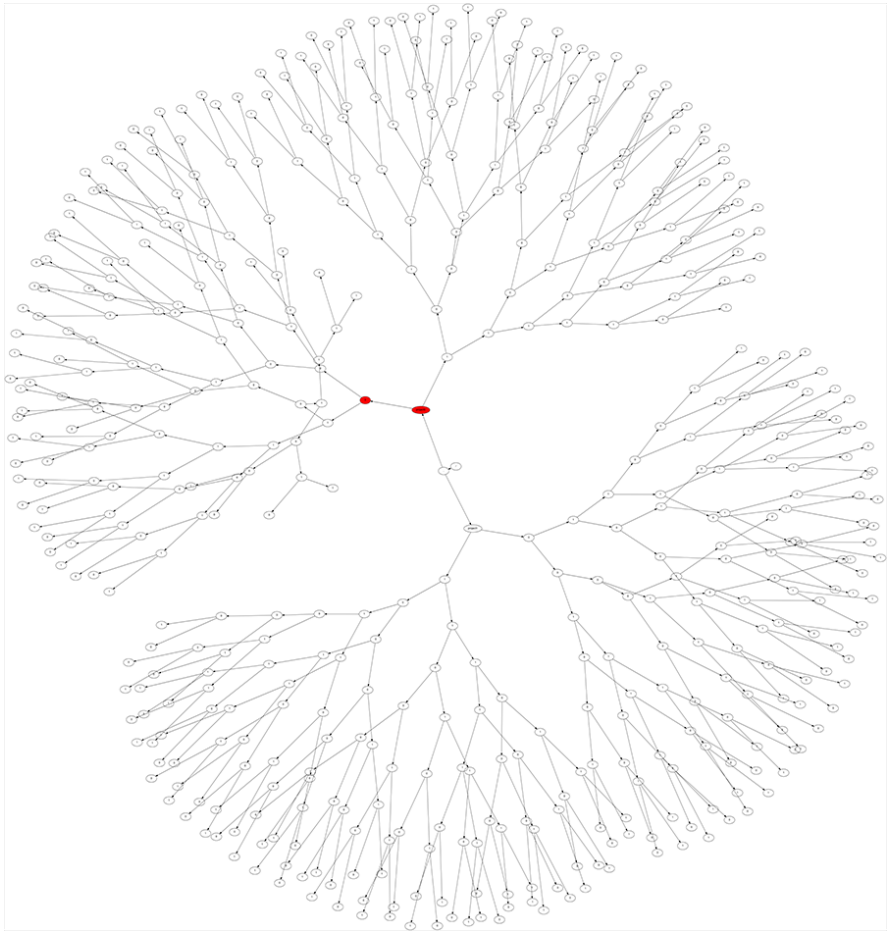
- Obtained and launched 492 ransomware samples
- CryptoDrop successfully detected all 492 samples
- Damage: Median of just 10 files lost before detection



[Scaife et al., ICDCS'16]

Ransomware Variance

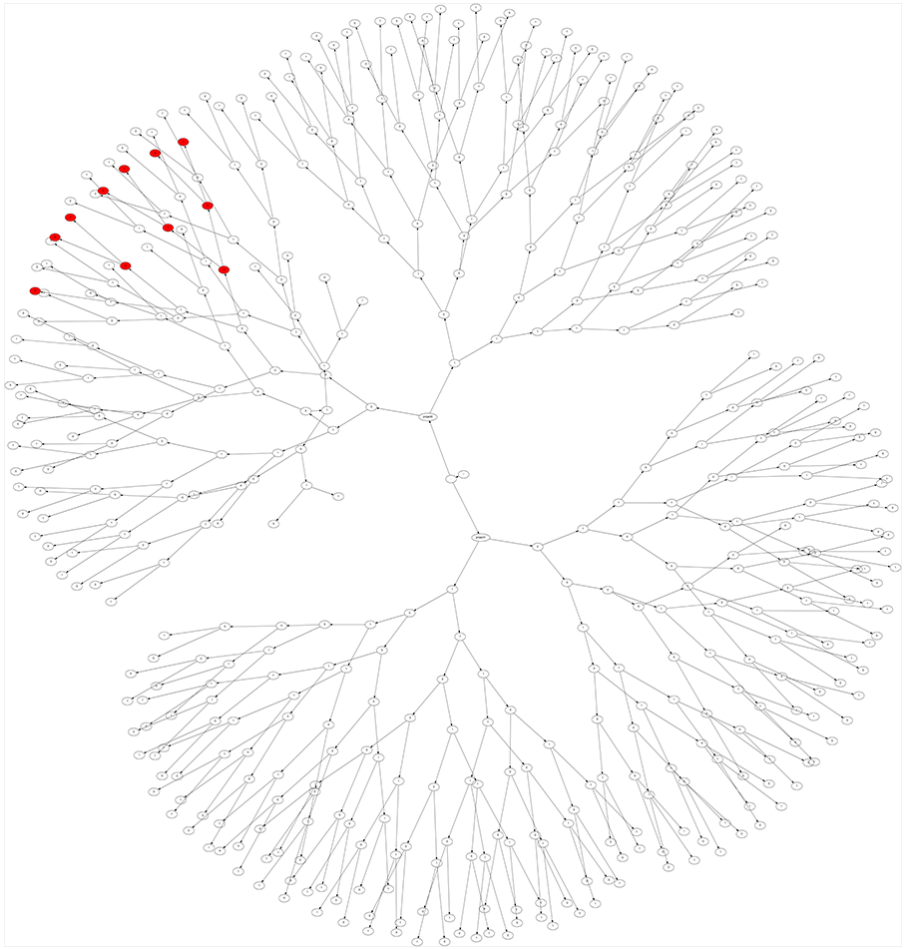
- Directory tree and files encrypted prior to detection:



`GPcode` [Scaife et al., ICDCS'16]

Ransomware Variance

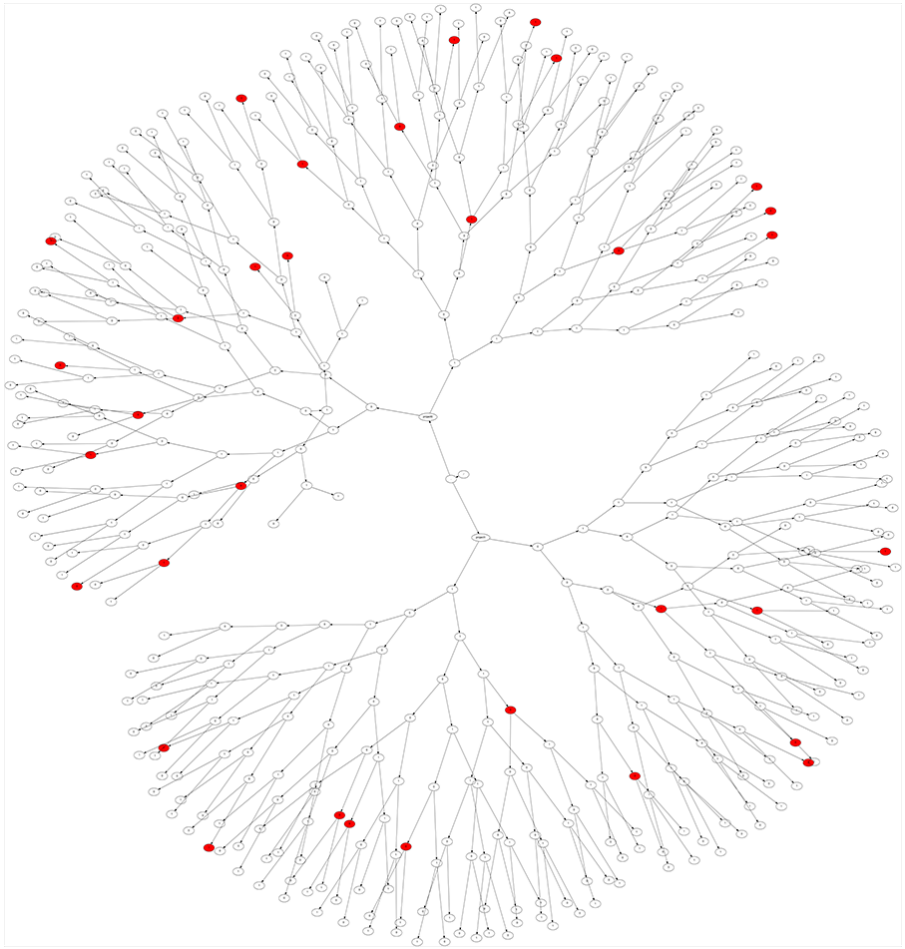
- Directory tree and files encrypted prior to detection:



`TeslaCrypt` [Scaife et al., ICDCS'16]

Ransomware Variance

- Directory tree and files encrypted prior to detection:



`CTB-Locker` [Scaife et al., ICDCS'16]

False Alerts?

- Application Corpus (30): **7-zip**, Adobe Lightroom, Avast Anti-Virus, ChocolateDoom, Chrome, Dropbox, F.lux, GIMP, ImageMagick, iTunes, Launchy, LibreOffice Calc, LibreOffice Writer, Microsoft Excel, Microsoft Office…
- Only 7-Zip (compression utility) triggers false alerts.
- Fundamental Limitation — CryptoDrop can't determine *intent* of changes it observes.
- Possible Mitigations?
- Possible evasion strategies for malware?

[Scaife et al., ICDCS'16]