CMSC 426
Principles of Computer Security

Malware Lifecycle and Analysis

Last Class We Covered

- Types of malware
- Well-known malware families
  - Gratuitous examples of malware

Any Questions from Last Time?

Today’s Topics

- Malware lifecycle
- Intro to malware analysis
  - Indicators of Compromise

Malware Lifecycle

Infection Lifecycle

- Timeline between when malware gets delivered to a system and when it gets done running

- Everyone has their own spin, but here's a simple one:
  1. Initial infection of victim occurs
     - First-stage malware on victim's computer
  2. Payload is delivered
     - Malware takes action
  3. Malware makes contact with actor
     - "Command & Control"

Infection Vector Example: Phishing

- Using email to convince a victim to click a link or download an attachment
- Initial infection occurs via this act

- Spearphishing
  - Phishing of specific, chosen victims
  - Higher rate of success

Infection Vector Example: Exploit Kit

- Compromised website redirects to a malicious website that is hosting the exploit kit
- Exploit kit does what it says on the box:
  - Scans the victim’s computer for vulnerabilities
  - Sends an appropriate exploit to the victim’s computer
    - Allows delivery of malware
- Patching exploits (allowing updates) is incredibly important
  - When patched, redirects can still happen, but exploit kit won’t have anything to exploit

First-Stage Malware

- A full malware payload is rarely delivered directly through the initial infection vector
- The "first-stage" malware gets execution on the victim's computer, then downloads and runs the payload
  - May be referred to as droppers, loaders, downloaders, etc.

- Most of the time, only first-stage malware is delivered
  - What purpose does this serve?
    - Most email clients don't allow executable attachments
    - First-stage can be smaller in size, with its limited functionality

First-Stage Example: Malicious Macros

- Files that contain macros that are attached to phishing emails
  - With the intention of the user running the macro and downloading/running the full payload
  - Often Microsoft Office documents, RTF files, or PDFs

- Office documents used to automatically run macros when a user opened the file
  - Now a notification (often including a warning) is shown to the user requiring them to manually enable macros
  - (Many users just click “Enable Content” anyway)

Payloads

- The actual file(s) that perform the malicious actions and achieve the actor’s end goal

- We talked about the different categories of payloads last time
  - Direct actions, like ransomware and cryptojacking
  - End goals, like making the machine part of a botnet, or setting up long-term monitoring with a RAT

- The parts of the malware that actually do the “cool stuff”

Command & Control

- Malware's communication of information with the actor
  - Banking Trojan – send login credentials when seen
  - RAT – constant possible interaction
  - Botnet – centralized C&C (master)
- End of the lifecycle (but this "end" can be very extended)

- Often referred to as C2 or C&C

- Payloads often connect back to a C&C IP address or domain in order to receive instructions from the malware actor

Missing Command & Control

- Not every malware has a C&C stage
  - Depends on malware’s actions and end goal

- Ransomware
  - Victim communicates “directly” with the actor
- Wiper
  - No communication necessary

Indicators of Compromise

Review: Indicators of Compromise

- Evidence that malware was on a system/network
- Can be used for attribution to a malware family, actor, and/or campaign

- Examples:
  - IP addresses and domain names
  - Email addresses
  - Cryptocurrency wallets
  - Hashes

IP Addresses and Domain Names

- Can show up in different instances:
  - IP address or domain name the malware downloaded from
  - IP address or domain name that the malware uses for C&C

- Quick reminder:
  - IP address:
    - 192.168.0.1
  - Domain name:
    - google.com

Email Addresses

- Can show up in different instances:
  - Email address used to send a phishing email
    - (May be spoofed, however)
  - Email address used to register a domain name
    - Not actually provided in the malware, but possible to look up who registered the domain name
    - With that information, possible to find out what other domains have been registered by that actor

Cryptocurrency Wallets

- Can show up in different instances:
  - Wallet listed in a ransomware note
    - Easy to find, for obvious reasons
  - Wallet that a cryptocurrency miner “deposits” into

Hashes

- Unique large number calculated by a hashing algorithm
  - In other words, the output of the hashing algorithm
  - Sometimes called the “digest,” often just called the “hash”

- If two files share the same hash, there is an *exceedingly* high probability that the files are identical

- Hashing algorithm may be run on any malware file
  - Files in payload, in first stage, etc.

Side Note: Hashing

- If two files have the same hash, they are functionally identical
  - Sort of allows a “diff” without having both files together

- If even one small change is made, the hash will change *drastically* (may be entirely different)

- Different hashing algorithms generate different sizes of hash
  - MD5, SHA1, and SHA256 are most common algorithms
  - (16, 20, and 32 byte hashes are generated, respectively)

Import Table Hashing

- Import address table is metadata within payload files
  - Contains list of all library functions used, in order they appear in code
  - Created by the original compiler/linker as the file is compiled/linked
- Hashing the import table gives you an imphash
  - “import hash”
- If hashing the whole file, a single change ➔ different hash
  - If an imphash, changes would have to be more substantial
  - But still unique-ish – variants will likely have different import tables

Fuzzy Hashing

- Official name is “context triggered piecewise hashing”
  - Most common program used for this is called ssdeep

- Details of how it works are complex, but essentially:
  - More robust against changes than traditional hashing
  - Can compare two fuzzy hashes and get a similarity score

DEMO TIME!

Announcements

- Homework 1 is up on the course Blackboard
  - Due at midnight on Wednesday, October 3rd

- Lab 2 will come out that same Wednesday

- Midterm 1 is on Tuesday, October 9th

Image Sources

- Bitcoin wallet (adapted from):
  - https://www.flickr.com/photos/30478819@N08/24874103608