SECRET TOOLS

Learning About Government Surveillance Software You Can't Ever See

Peyton "Foofus" Engel
DEF CON 25

Hurley, Burish & Stanton, S.C.

Non-Public Technology

- Breath analyzers: source code analysis reveals bugs
  - These are used for deciding whether to arrest and charge people.
    - https://arstechnica.com/tech-policy/2009/05/buggy-breathalyzer-code-reflects-importance-of-source-review/
    - https://www.wired.com/2009/05/minnesota-court-release-source-code-of-breath-testing-machines/
- The Clipper Chip: the NSA-designed encryption device with a built-in back door
    - https://en.wikipedia.org/wiki/Clipper_chip
- Trust us: we know what we're doing

Surveillance Is Pervasive

- CALEA
  - Restrictions on installing un-tappable phone systems
    - https://en.wikipedia.org/wiki/Communications_Assistance_for_Law_Enforcement_Act
- NSA Call Metadata Collection
  - Traffic analysis can be just as useful as content analysis
- Surveillance is secret, also: most people didn't find out about the extent of surveillance until Snowden told us
    - http://www.npr.org/sections/thetwo-way/2015/11/29/457779757/nsa-ends-sept-11th-era-surveillance-program

More Than Just Surveillance

- Some of the “Playpen” cases have been dropped over the use of a NIT
  - » FBI implants malware on a website that it seized in order to obtain IP addresses of TOR users
    - https://arstechnica.com/tech-policy/2017/03/doj-drops-case-against-child-porn-suspect-rather-than-disclose-fbi-hack/
- Government exploit code leaks
    - https://www.theregister.co.uk/2017/03/08/cia_exploit_list_in_full/
- Where is the boundary between aggressive investigation and violating rights?

Investigating Peer-to-Peer Networks

- Gnutella, BitTorrent, Ares, etc.
- These have been around for a while (the Gnutella variant has been in use since 2009)
- Forks of open-source software
  - Make use of under-the-hood aspects of the peer-to-peer protocol that aren't usually accessible to users
  - Add in some features that would not be of interest to ordinary users

Who Develops These?

- Joseph Versace: Programmer/analyst with the Ontario Provincial Police
  - » Roundup Ares: .Net-based client for the Ares network
- Collaboration of CS departments at Univ. MA Amherst, Georgetown; PA and MA state police:
  - » Roundup: Java-based fork of the Phex Gnutella client
  - » Roundup Torrential Downpour: for BitTorrent

New Uses for Existing Features

- Gnutella
  - Search query hits include SHA-1 values of files shared
  - “Swarming” information provided by download source hosts includes IP addresses and GUIDs of systems sharing the same file
  - Direct browsing of peers

New Uses for Existing Features

- BitTorrent
  - Tracker Messages: which peers are interested in which torrents?
  - Torrent Segment Data: peers announce what pieces of files they possess, when they connect for downloads and when they acquire new segments
  - Peer Exchange: like swarming info for Gnutella

New Features

- Known file lists: a database of hashes of known files of interest
- IP Geolocation
- Single-Source Downloading: attempt to get all segments of a file from a single host (*i.e.*, defeat the purpose of the peer-to-peer protocol)
- Anti-feature: uploading is disabled
- “Tagging” individual systems – more on this later...

So What Do They Do?

- Impersonate regular peers
- Engage in activity (*e.g.*, queries, announcements) designed to attract connections
- Do queries of their own to find peers sharing files of interest
- Inspect the systems that they connect with
- Perform single-source downloads
- Log their activity

The Code Must Remain Secret #1

- It would divulge our database of contraband
- The database is part of the software? Unlikely.
- And disclosing would be disruptive to the trading of illegal materials:
  - Everyone would go and flip a bit or two in their files, and that would mean they wouldn’t be identified as the same on the network

The Code Must Remain Secret #2

- It will disclose the undercover investigators
- Interesting: suggests there may be a shared list of static Ips or reverse DNS, so that investigators don’t accidentally target one another
- The list is probably not part of the software, but the software probably does refresh its copy from time to time
- But then from time to time, they show logs from the software, which include public IPs

Potential Problems: Reliability

- False Positives: does the software ever erroneously report what it has done?
  - » Investigators frequently don’t find the files their warrant affidavits say they downloaded
  - » But, they also generally don’t execute warrants until months after the downloads
- Are there conditions under which the software malfunctions?

Potential Problems: Warrants

- Beyond the technology the public has
- ***Kyllo v. United States***, 533 U.S. 27 (2001): use of FLIR system to visualize activities with home required a warrant
  - » “But this is just modified open-source software; any user could do the same thing.”
  - » How would we know we were doing the same thing?
- Tagging: GUIDs and log files
  - » Shared vs. non-shared areas

Potential Problems: Testimony

- What are the chances a judge will be able to evaluate the reliability of statements about:
  - How IP addresses can be correlated to ISP subscriber identity?
  - How peer-to-peer networks work?
  - How a government tool based on open-source software works?
- Who is qualified to testify about how these tools work in court?
  - Requires more than just knowledge of their use

Potential Problems: Exploitation

- The software may inherit flaws from the components out of which it was constructed (*e.g.*, Java, Phex, the BitTorrent protocol, *etc*.)
- The software may have bugs of its own
- Exploitation would likely go undetected
  - » Lack of transparency/availability
  - » Mostly used by investigators, not security professionals

THANKS!

Brought to you by the kindly folk of

Hurley, Burish & Stanton, S.C.

and

foofus.net