How to Spy with Python

So easy, the NSA can do it!

Lynn Root | @roguelynn

Agenda

- Introduction
- Presentation: Some Context with what the NSA is doing (15 min)
- Installation! (30 min)
- Break! (10 min)
- Introduction to packet sniffing with Scapy (40 min)
- Break! (10 min)
- 6 Queries inspired by the NSA (60 min)
- Protecting oneself (15 min)

whoami

Backend Engineer @ Spotify

Vice Chair of the PSF Board of Directors

PyLadiesSF Founder

Historical Context

TL;DR: it’s nothing new

1946

Five Eyes Group

- Comprised of the US, Canada, Australia, UK, and New Zealand
- Purpose: share signal intelligence
- Each country to surveil a set of other countries

1952

NSA Established

- Originally started within the US Army in 1917 after US declared war on Germany
- Morphed into the Armed Forces Security Agency (AFSA) after WWII
- AFSA redesignated itself as the NSA in 1952 after it failed to get its sh*t together

1973
Warrants Required

Finally – 20 years **after** the NSA was established – over 50 years **after** the US started its surveillance – the Supreme Court makes warrants a requirement for domestic surveillance

1978
Warrants Required

After the Senate's Church Committee revealed illegal domestic spying by the NSA in 1975, the Foreign Intelligence Surveillance Act (FISA) was signed to protect Americans.

A [not-so] “secret” court, the Foreign Intelligence Surveillance Court, was created for the purpose of hearing requests for warrants.

2001
Culture Shift

After the 9/11 World Trade Center attacks, the culture against spying begins to shift within the NSA. Within the first month after September 11th,

- White House asks NSA what more could be done against terrorism if the NSA had more authority.
- NSA resurfaces plan to perform contact chaining on metadata it collected, originally deemed illegal in 1999 by the FISA.
- US President gives NSA authority to begin targeting terrorist-associated foreign phone numbers.

2002-2003
Telecoms & Domestic Spying

- Unrevealed telecoms and internet providers in the US receive letters from NSA requesting data and support for its domestic spying program.

- AT&T employee discovers the NSA is working inside AT&T's San Francisco facility.

- Telecoms formally enter into a voluntary agreement with US to give data to the NSA.

- Installation of special technology to a "secret room", room number 641A, at AT&T's San Francisco facility that can read & analyze 10s of thousands of communications per second, and then send those communcations to a central database.

2003

Total Information Awareness (TIA)

- Program formally started in 2003, with development beginning in 2000.
- Aimed to gather detailed information about individuals to anticipate & prevent crimes.
- Congress stops TIA in late 2003, but program is quietly moved into the NSA's domestic spying program.

2005-2007
NSA Exposed

- In December 2005, the New York Times reveals that the NSA has been spying on Americans without warrants. Soon after, President Bush confirms NSA's warrantless eavesdropping.

- The New York Times also reveals that some of the NSA's spying is purely domestic with some telecoms giving backdoor access to communication streams.

- Soon after the NYT's articles, an unknown company requested that the NSA to issue court orders, rather than companies voluntarily handing over data.

- In 2007, the Protect America Act passed, allowing the NSA not to need warrants for collecting communications.

2007-Now
PRISM

- Data collection for PRISM starts with Microsoft in September 2007.
- July 9th, 2008, US Congress passed amendments to FISA that gives telecoms legal immunity for those that cooperated with the NSA's wiretapping.
- In 2012, the NSA started to build its biggest spy center in Utah for its collection of intercepted data.
- Also in 2012, the FBI pushes for wiretap-ready websites, asking internet companies to not oppose a law making backdoors mandatory.
- June 2013, the Washington Post exposes the PRISM program. Shortly after, XKeyScore was revealed.

What the NSA is actually doing

What is PRISM?

- Planning tool for Resource Integration, Synchronization, and Management
- Mines electronic data for the purpose of mass surveillance
- Collects intelligence that passes through US servers
- Targets foreigners, but is elusive about data on US citizens
- Only collects metadata (supposedly)

What the NSA is actually doing

What is XKeyScore?

- Digital Network Intelligence Exploitation System
- Federated Query System of completely unfiltered data
- 500 - 700 servers, as of 2008
- Gives users ability to query for email addresses, a target's activity, phone numbers, HTTP traffic, extract file attachments, etc.

What the NSA is actually doing

What is Hacienda?

- Data reconnaissance tool developed by the UK's GCHQ
- Port scans entire countries (27 listed but not revealed)
- Particularly interested in FTP, HTTP/S, SNMP, SSH, among others
- Looking for vulnerable services running on these ports
- (Ab)used by the Five Eyes group to launch exploits or steal data
- Can "infect" non-government machines to complete scans, building their own botnet essentially, and enabling to complete a scan for vulnerable devices within a subnet within 5 minutes
- It only takes a simple email request to access data!

Unanswered Questions

- What does metadata mean?
- How do companies not notice being backdoored? or are they lying when denying cooperation?
- How is a target's "foreignness" determined? How exactly are they identifying non-US citizens?
- What is done with data that's "accidentally" collected on Americans?
- How is the PRISM-collected data handled by the NSA? Does the NSA maintain rigorous security measures to protect against threats?

Why metadata matters

taken from the EFF presentation at 30C3 in December 2013

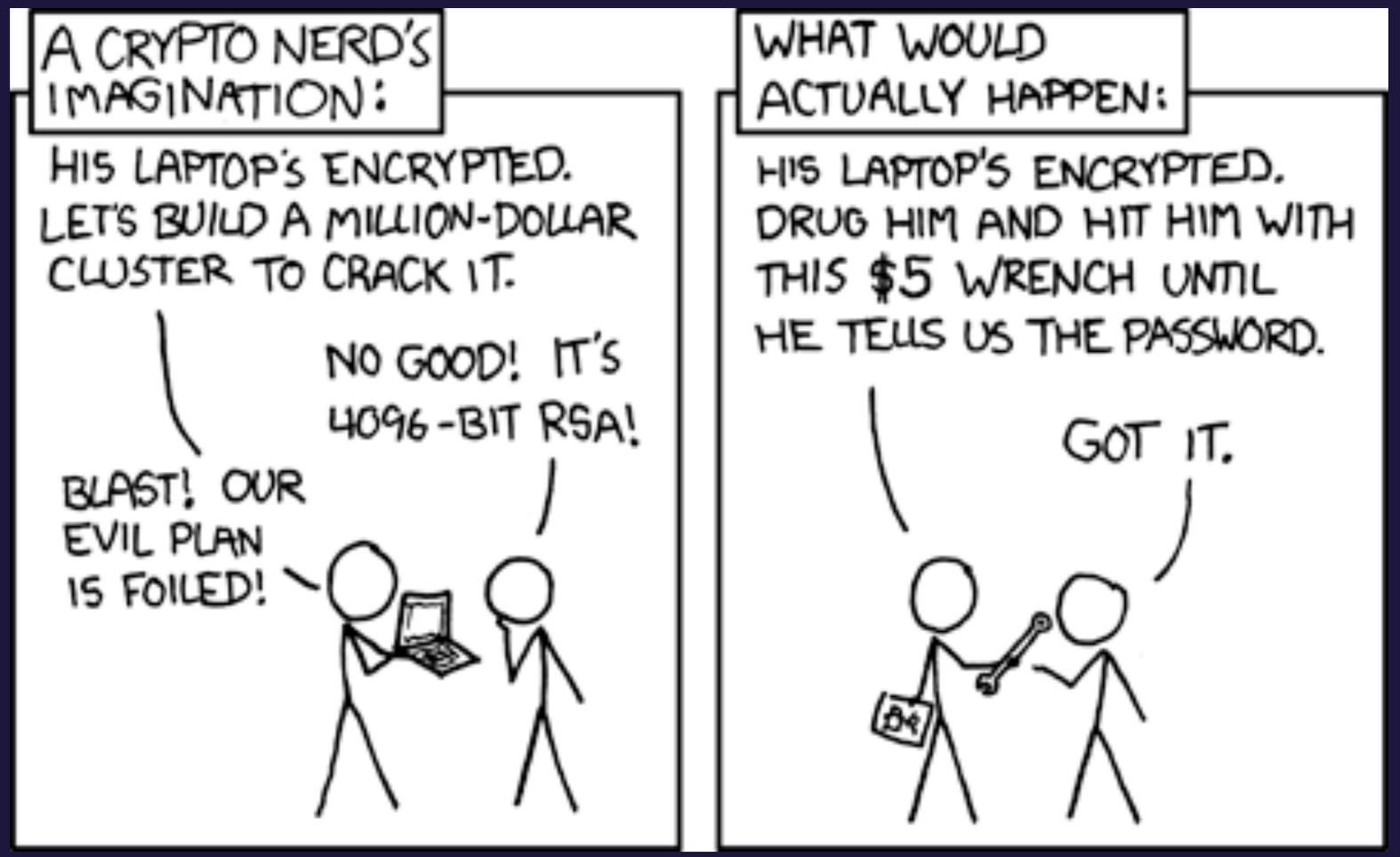
- They know you rang a phone sex service at 2:24 am and spoke for 18 minutes. But they don't know what you talked about.

- They know you called the suicide prevention hotline from the Golden Gate Bridge. But the topic of the call remains secret.

- They know you spoke with an HIV testing service, then your doctor, then your health insurance company in the same hour. But they don't know what was discussed.

How they're doing it

(actually it's more like)



How you can do it!

Tools we'll use:

- IPython Notebook
- Scapy – packet sniffing & manipulation
- pygeoip – API for GeoIP databases
- geojson – bindings & utilities for GeoJSON
- python-nmap –wrapper around nmap port scanner

How you can do it!

Follow the installation instructions here:

http://rogue.ly/spy-tutorial-setup/