Surveillance—How it Works

Steven M. Bellovin — https://www.cs.columbia.edu/~smb

February 5, 2024

CS@CU

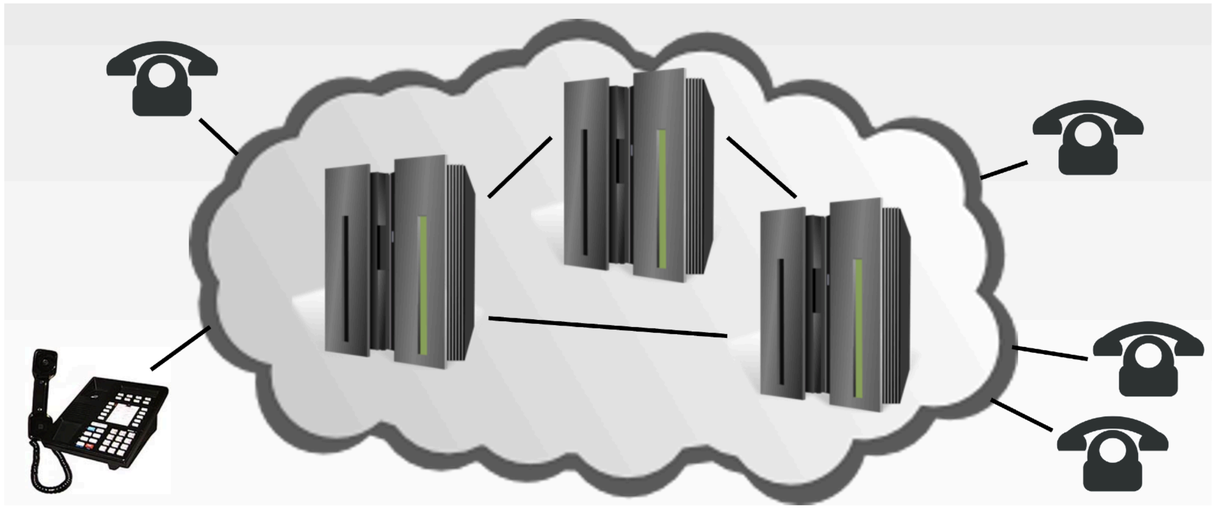
Surveillance—What is It?

- It's a way of seeing what people are doing that goes beyond what we can see or hear
- We all know about statutes, court decisions, and court orders: the Wiretap Act, the Stored Communications Act, the Pen/Trap Act, *Katz*, *Smith*, *Jones*, *Kyllo*, *Carpenter*, warrants and subpoenas, etc.
- But how does it work? How are phones tapped? Internet conversations? How are people's movements tracked? How do you look inside a house?

Wiretapping Phones

The (Traditional) Phone Network

- Customers had telephones
- Each phone was connected to a *phone switch* in a *switching office* (often called the *central office*, or *CO*)
- The switches—originally electromechanical, now computerized—are connected to each other
- A switch might handle thousands of phone lines
- The phones were dumb; the switches were smart



Tapping Phones—History

- Old-style phones had a two-wire, analog connection
- Tapping a line was as simple as climbing a telephone pole and connecting a “butt set” to that person’s wires (if you knew which they were)
- It was simple and effective, and worked for many decades
- However…

Photo from Wikimedia Commons

Loop Extenders

- The top of a telephone pole can be cold, hot, wet, etc.
- Worse yet, it's *noticeable*
- And you can't use your butt set at a phone company switching office because of the prevalence of *subscriber loop carriers*
- Solution: a *loop extender*—connect one pair of wires to the target's phone line and another to a vacant *friendly pair*
- Taps can now be done in the switching office

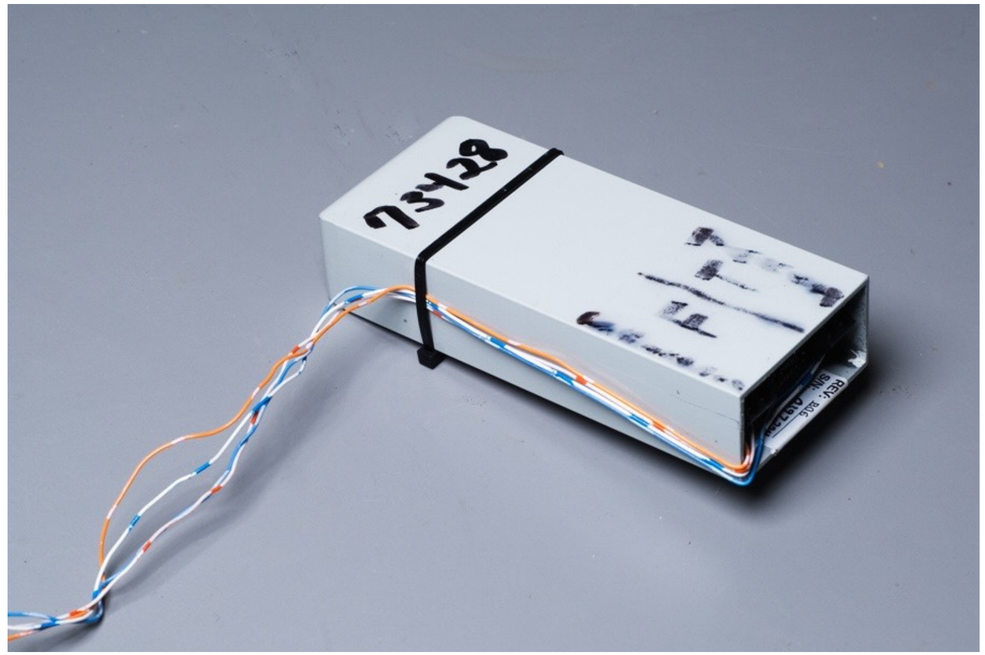


Photo courtesy Matt Blaze

This Worked, for a While…

- One brief telco visit to install the loop extender on a pole or in the SLC cabinet
  - No one is going to wonder about that
- Listening to the call could be done from the CO
- But by 1990, the FBI foresaw trouble

Photo courtesy Matt Blaze

New Phone Technologies

- New phone technologies were on the horizon
  - Cell phones existed, though they were comparatively rare
  - ISDN—*Integrated Services Digital Network*, which provided the blazingly fast speed of 56K bps—was digital
- There was starting to be an increase in modem calls
- Butt sets would no longer work—another solution was needed

The Digital Telephony Bill

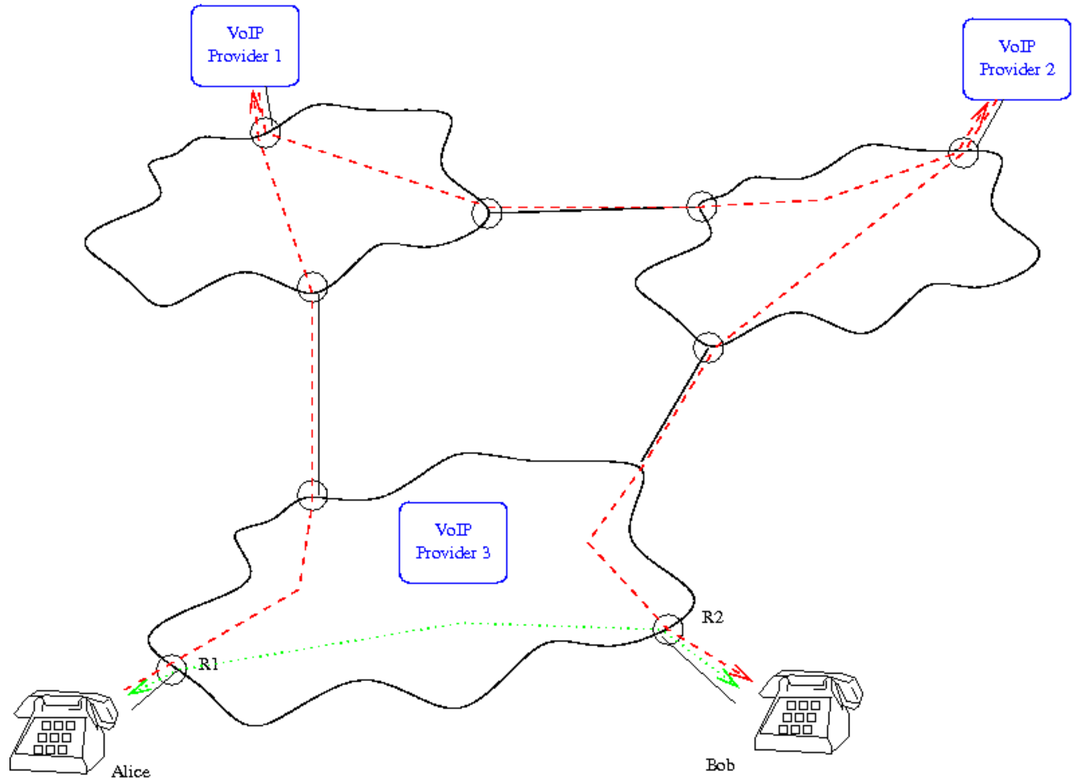
- The FBI understood that there was trouble coming
- They asked Congress to pass the *Digital Telephony Bill*, eventually enacted as *CALEA* (*Communications Assistance to Law Enforcement Act*, 47 U.S.C. §§1001-1010)
- CALEA required that phone switches have a standardized wiretap interface, regardless of the underlying technology
- Crucially—no requirement that the phone companies break encryption unless they supplied the encryption keys

CALEA Problems

- Modern phone switches are computers running very complex, highly specialized software
- Modifying this code is difficult and expensive
- CALEA authorized $500M for conversion costs—but it wasn't nearly enough
- A phone switch is a computer—what if someone hacks it and exploits the CALEA interface?
- That happened, in Greece—some blame US intelligence

Voice Over IP

- The call is set up by messages between VoIP providers
- The actual call might travel via a different path, on different networks
- The provider may not be on the path for the actual conversation, and hence *can't* tap the call
  - The provider may also be in a different jurisdiction



Cellular Calls

- Cell phones talk to *base stations* (technically: base stations connect to *mobile switching centers (MSC)*, which are linked to the conventional phone network)
- A phone announces itself to a base station, which (after authentication) tells the phone network where this particular phone is now
  - In particular, the *home register* knows what MSC should handle incoming calls to that number
  - It's possible to send queries to the home register to track someone, with or without proper legal authorization—any phone company in the world can do it…
  - If the phone moves during a call, the call is handed off to another base station or MSC
- Taps are done at the MSCs or phone switches, not by intercepting the radio signals

Where Things Stand

- The traditional landline network is effectively dead
- Almost all calls use VoIP or cellular
- This means that CALEA is the only solution—if it works at all

Tapping the Internet

Legal Authority

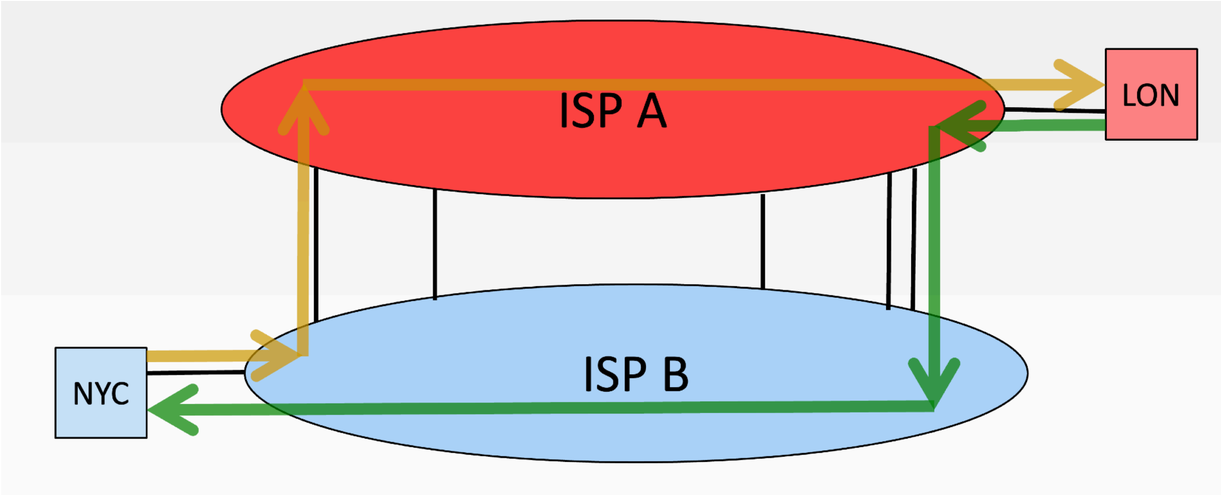
- The 1986 amendments to the Wiretap Act extended coverage to ISPs as well as phone companies
- DoJ and the FCC have decided that CALEA applies to ISPs
  - 47 U.S.C. §1001(8)(B)(ii): “The term “telecommunications carrier” means a person or entity engaged in the transmission or switching of wire or electronic communications as a common carrier for hire; and includes… a person or entity engaged in providing wire or electronic communication switching or transmission service to the extent that the Commission finds that such service is a replacement for a substantial portion of the local telephone exchange service
  - ISPs provide “electronic communications” (18 U.S.C. §2510(12)) but generally not “a replacement for a substantial portion of the local telephone exchange service” unless they also do VoIP—so why does CALEA apply? (*Am. Council on Educ. v. FCC*, 451 F.3d 226 (D.C. Cir. 2006))

Technical Challenges

- An Internet message is broken up into multiple *packets*
- The system must function correctly if packets are dropped, damaged, duplicated, or reordered
- Every packet on the net is independent; most media are *multi-access*, i.e., contain data to and/or from multiple destinations
- Conclusion: interception has to be done based on data in the packets—and these are often not knowable in advance

Internet Routing

- A web browser in London wants to query a web server in New York
- The London ISP A hands off the query to ISP B as quickly as it can, to cross the Atlantic
- The reply, via ISP B, is handed to ISP A as quickly as possible
  - This is called *hot potato* routing
- Routing is asymmetric, so taps have to be done as close to the target as possible



A Typical IP Packet

Four Sections…

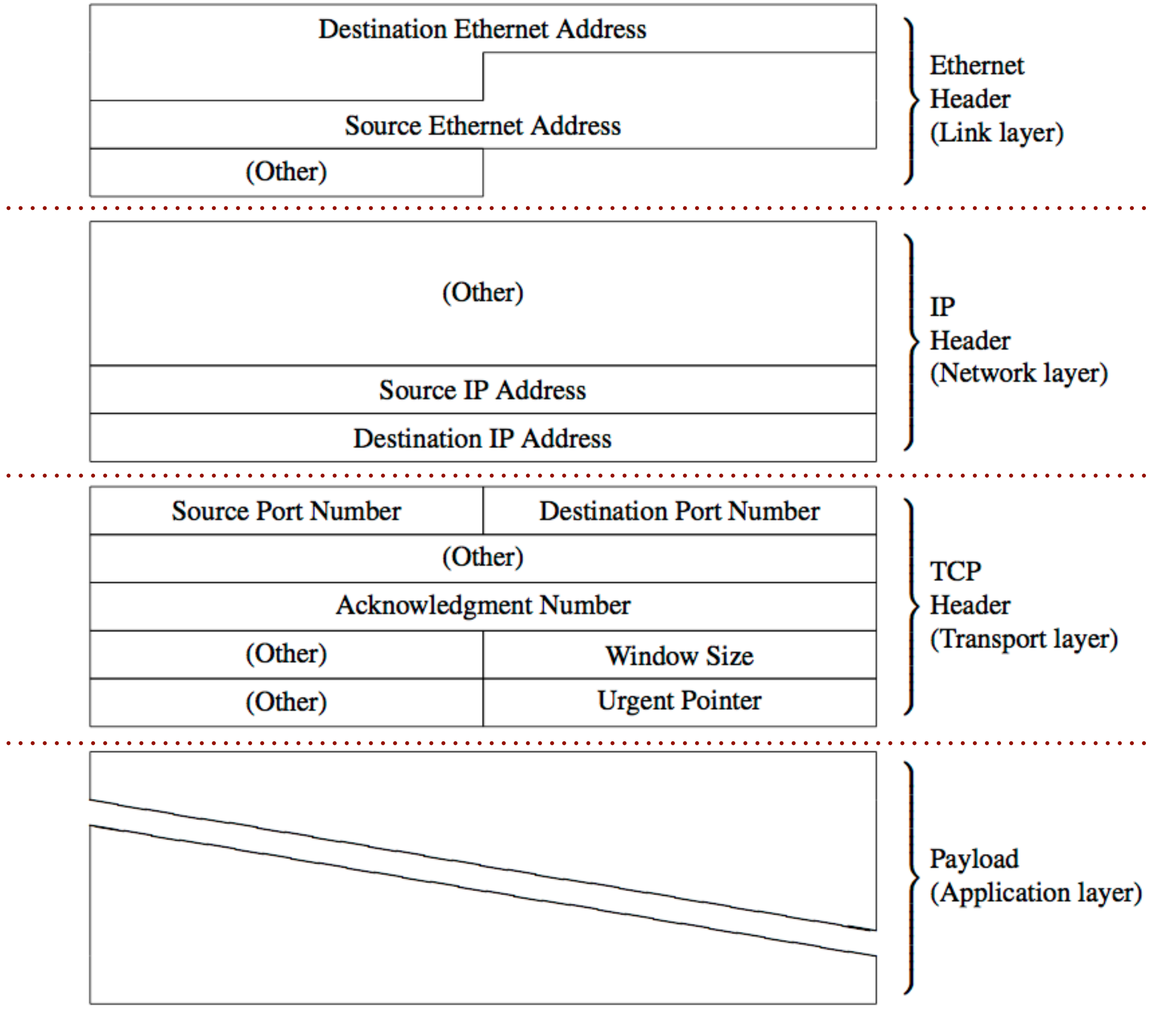
1. The *link-layer* (Ethernet, WiFi, etc.) header

2. The *Internet Protocol (IP)* header

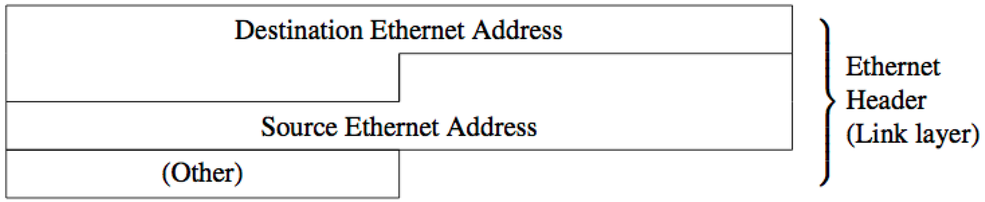
3. The *Transmission Control Protocol (TCP)* header

4. The payload (user data)

Decisions on what packets to actually intercept can be based on any of these sections.

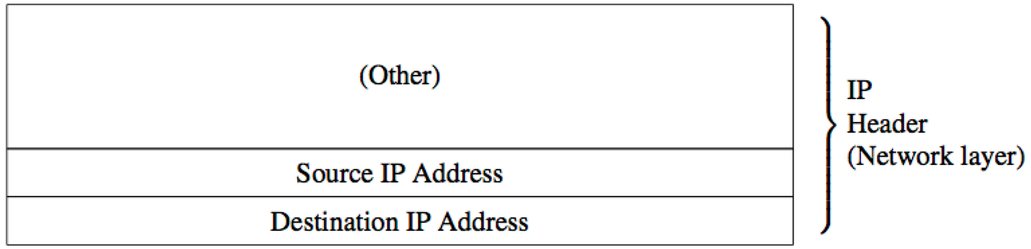


Link-Layer (Ethernet, WiFi) Addresses



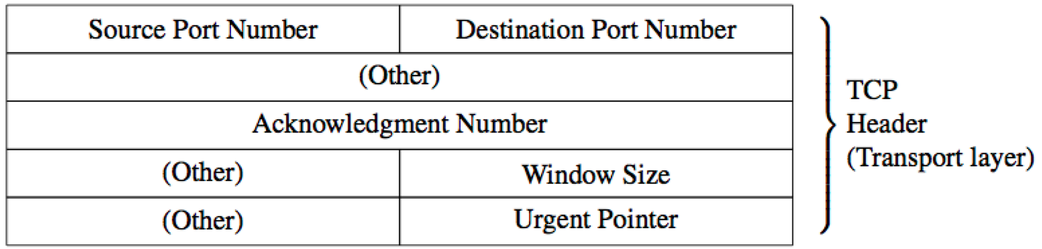
- Ethernet and WiFi addresses are manufactured into devices (more complicated for iOS)
- They do not leave the local network—you can't tap my home computer at my ISP based on its WiFi address; you can only do that if you come into my apartment
- Might be useful for tapping in office environments if the network administrators are not suspects

IP Addresses



- (Roughly) akin to street addresses
- Visible throughout the Internet—but often are changed at the border to a residence or company
  - (The reasons are complicated…)
- The easiest way to tap—*if* you know the target's IP address, which you generally don't; they're usually dynamically assigned

TCP Port Numbers



- If IP addresses are like street addresses, TCP port numbers are like rooms in a building
  - 25, 110, 143, 587 are for email; 80 is for web, etc.
  - (There are other ports for encrypted versions of the above)
- If you want to capture only someone reading mail, you'd tap ports 110 and 143

User Payloads

- Can contain anything—protocol-specific
- You might look at From/To on email traffic, or URLs on web traffic
- Utility today is hindered by near-ubiquitous encryption

How To Tap?

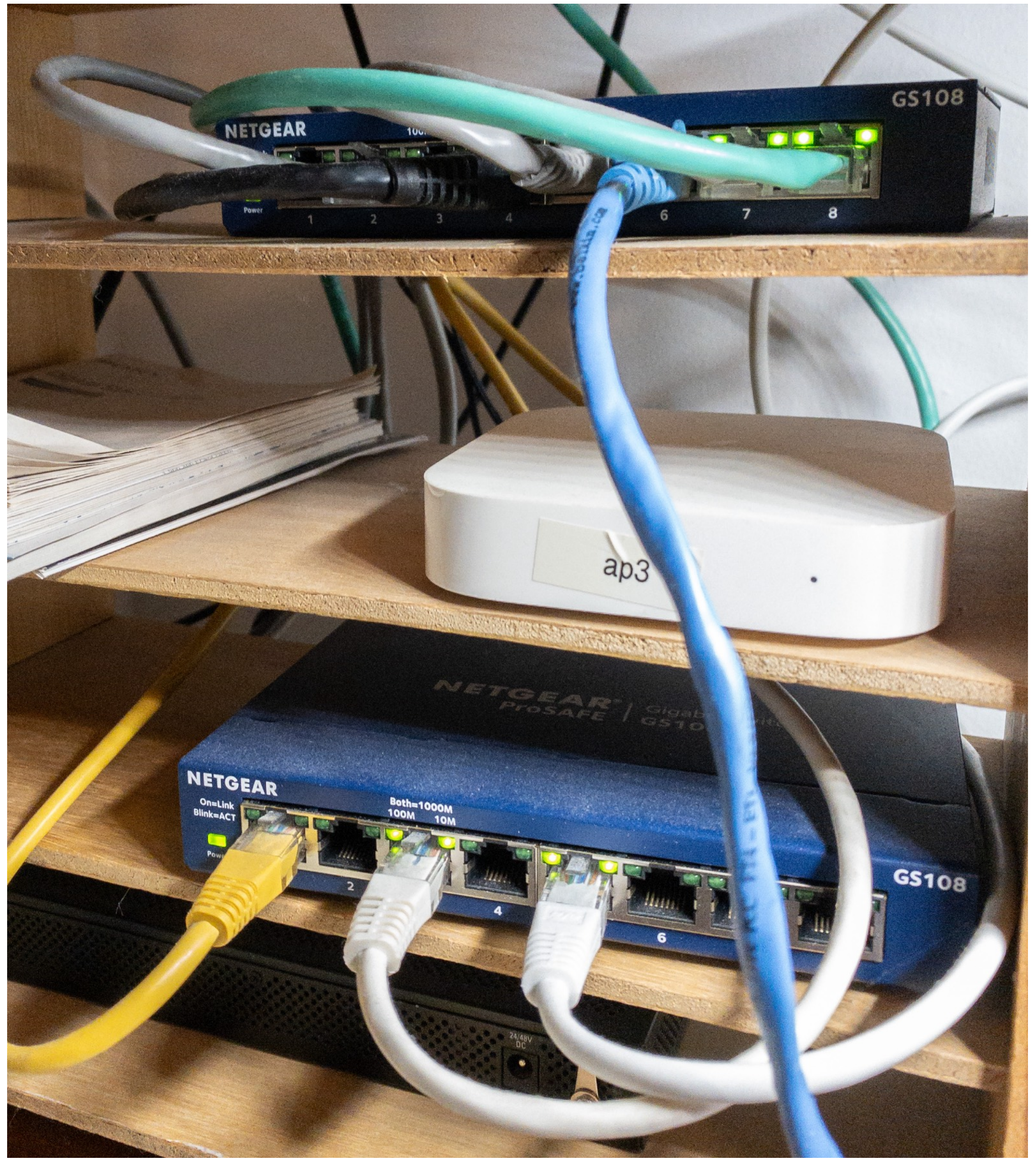
- Two issues: network access and filtering
  - (Remember that most media are “multi-access”—many devices (and people) are using the medium)
- Locations for tapping: residence (requires cooperation from another resident or surreptitious entry), business (employer, public hotspot, hotel, etc.), ISP

Residences

- Must have permission to enter—warrant or other resident—in addition to wiretap warrant
  - Sometimes, one can tap WiFi from outside, if your surveillance van isn't too obvious…
- Must have a way to connect to the residential network
  - Without cooperation, that can be hard—WiFi is generally encrypted these days, and you probably need the key
- Generally filter on Ethernet/WiFi address

Enterprise

- Easiest to tap into *switch*
  - But—switches send traffic to a port only if destined for a computer on that port
- However, unlike home switches, enterprise switches and routers generally have *mirror* or *monitoring* ports—they see all traffic sent to/from a specified other port
- No need to worry about WiFi encryption
- Cooperation can be compelled under 18 U.S.C. §2518(4)(e)



ISPs

- Must tap close to the target, not on the Internet backbone
- Easy to know the customer's IP address, but hard to do *minimization*
  - My spouse and I each have phones, tablets, and computers—whose traffic is targeted? On the Internet, all of those devices will have the same IP address.
- Again, cooperation can be compelled under 18 U.S.C. §2518(4)(e)

IP Address Assignment

- Most user computers (laptops, desktops, phones, tablets, etc.) receive their IP addresses dynamically
  - Server IP addresses are generally static—useful if tapping, say, a criminal web site
- A protocol known as *DHCP (Dynamic Host Configuration Protocol)* is used
- Computers send a *broadcast* DHCP request message
- The DHCP server (built in to home routers; a separate computer for enterprises) looks at the computer's MAC address or (sometimes) hostname to assign an address
- Previous addresses are reused if still available; if not, a new one is assigned
- IP address assignments have *lease times*—computers must renew their lease to retain their IP address

Content Versus Metadata

What Type of Tap is Authorized?

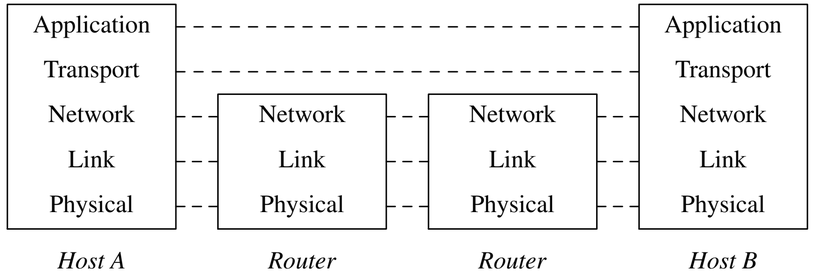
- Content taps require “super-warrants” (18 U.S.C. §2516)
- Metadata taps—“dial, routing, addressing, and signaling”—require a much simpler court order (18 U.S.C. §§3121-3127)
  - “[T]he information likely to be obtained is relevant to an ongoing criminal investigation”
- Generically known as *pen/trap* orders
  - A *pen register* records the numbers someone dials
  - A *trap-and-trace* device records the caller’s number

Pen Registers

- (Ancient) rotary dial phones produce electrical impulses for digits dialed
- Early pen registers simply made a mark on a paper tape for each impulse
- “Indeed, a law enforcement official could not even determine from the use of a pen register whether a communication existed. These devices do not hear sound. They disclose only the telephone numbers that have been dialed -- a means of establishing communication. Neither the purport of any communication between the caller and the recipient of the call, their identities, nor whether the call was even completed is disclosed by pen registers.” *United States v. New York Telephone Co.*, 434 U.S. 159 (1977)
- (Trap-and-trace is more complex, and used to work using CallerID technology)

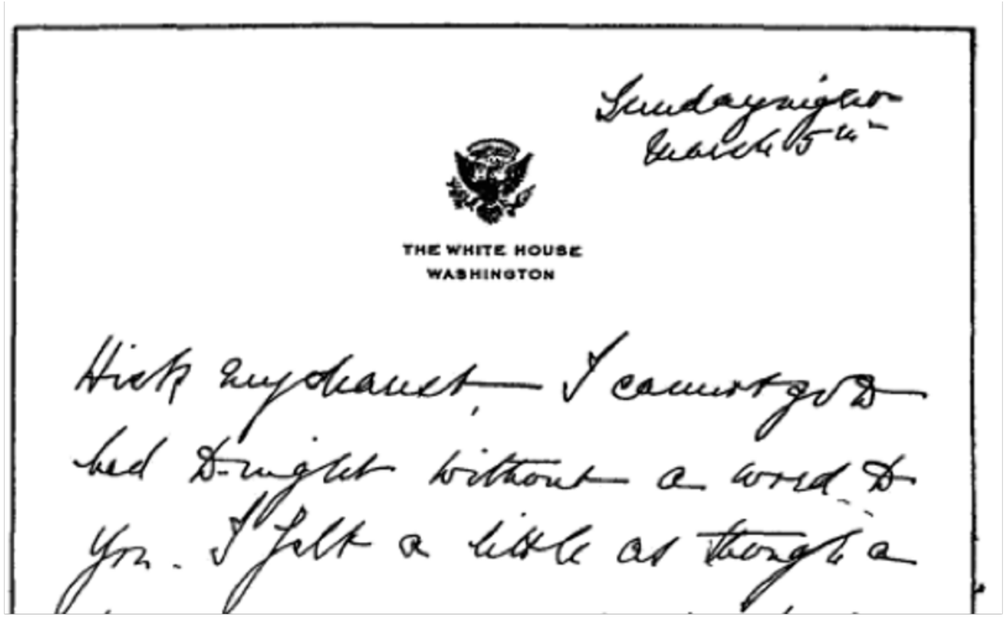
Metadata on the Internet

- What is the equivalent of a phone number on the Internet?
- The answer can be very complicated!
- IP addresses are seen by every router along the path, and hence are clearly third-party data
- Port numbers are often end-to-end only, and hence not third-party data—but it gets complicated! (They're sometimes taken by ISPs, but not voluntarily given…)
- Email addresses are sometimes even worse
- Courts (and DoJ) don't understand the subtleties



Example: Email

- Internet protocols make a distinction between a message *envelope*, message *headers*, and message *bodies*
- The message body is always end-to-end —but what about gmail's ad-scanning? Virus scanning? CSAM scanning?
- The envelope is usually third party data— but some people (like me) run their own mail servers, so it's not
- Headers are a combination of end-to-end and third party data
- It's actually more complex than that…

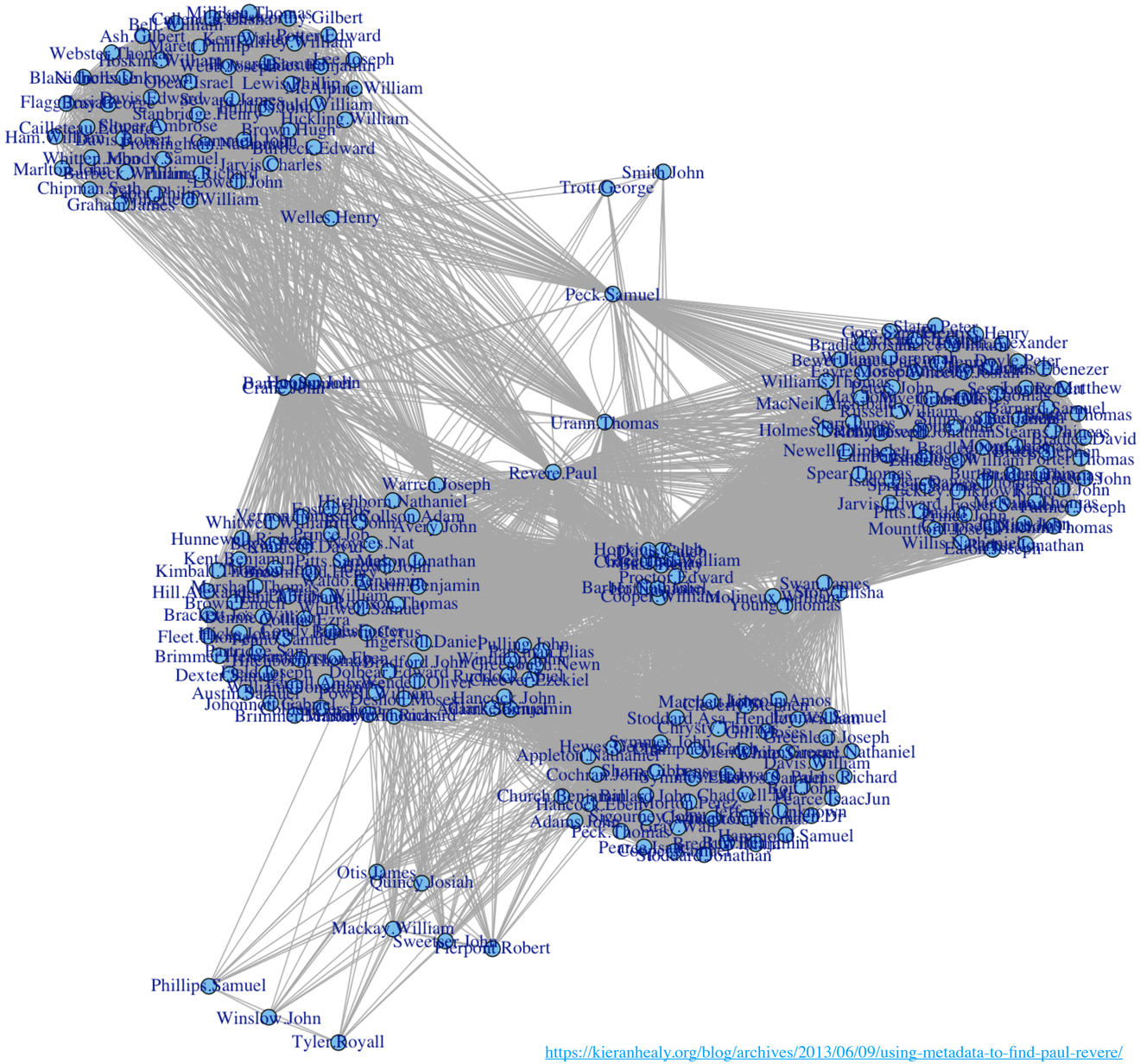


Metadata Analysis

- Metadata analysis is an extremely powerful technique
- Who talks to whom is extremely revealing
  - The military calls it traffic analysis, and has been using it for more than 100 years
  - During World War I, it was used as an adjunct to cryptanalysis, but it has independent value
- Much harder to hide metadata than to hide content
  - (For web browsing, use Tor to hide your metadata)

Finding Paul Revere Through Metadata

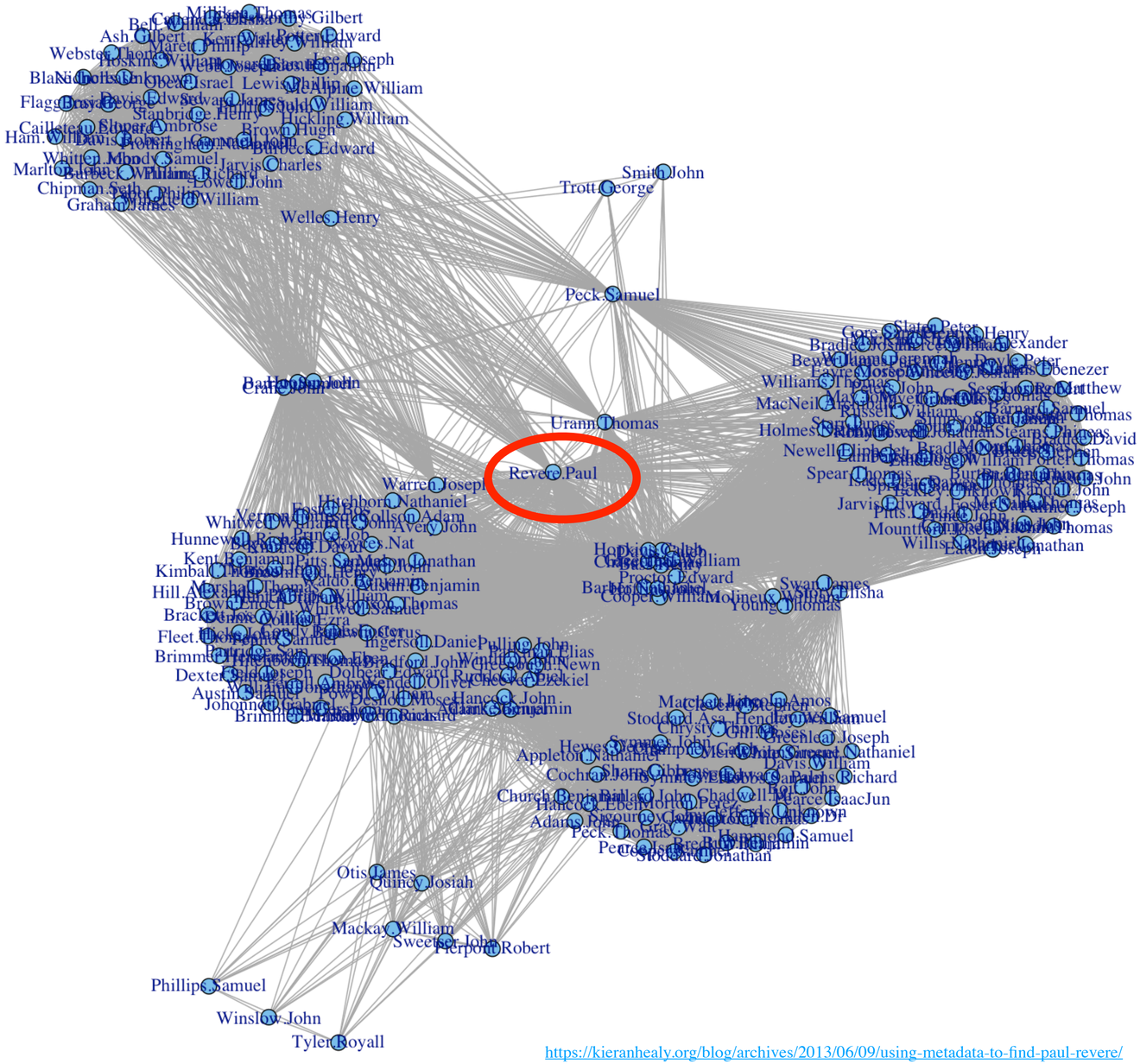
- There are good records of who belonged to which patriotic organizations in pre-Revolution Boston
- By a bit of simple analysis, it's possible to see who linked them all together: Paul Revere
- Note well: this is not based on any knowledge of what was said or done at any of these meetings



https://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/

Finding Paul Revere Through Metadata

- There are good records of who belonged to which patriotic organizations in pre-Revolution Boston
- By a bit of simple analysis, it's possible to see who linked them all together: Paul Revere
- Note well: this is not based on any knowledge of what was said or done at any of these meetings



https://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/

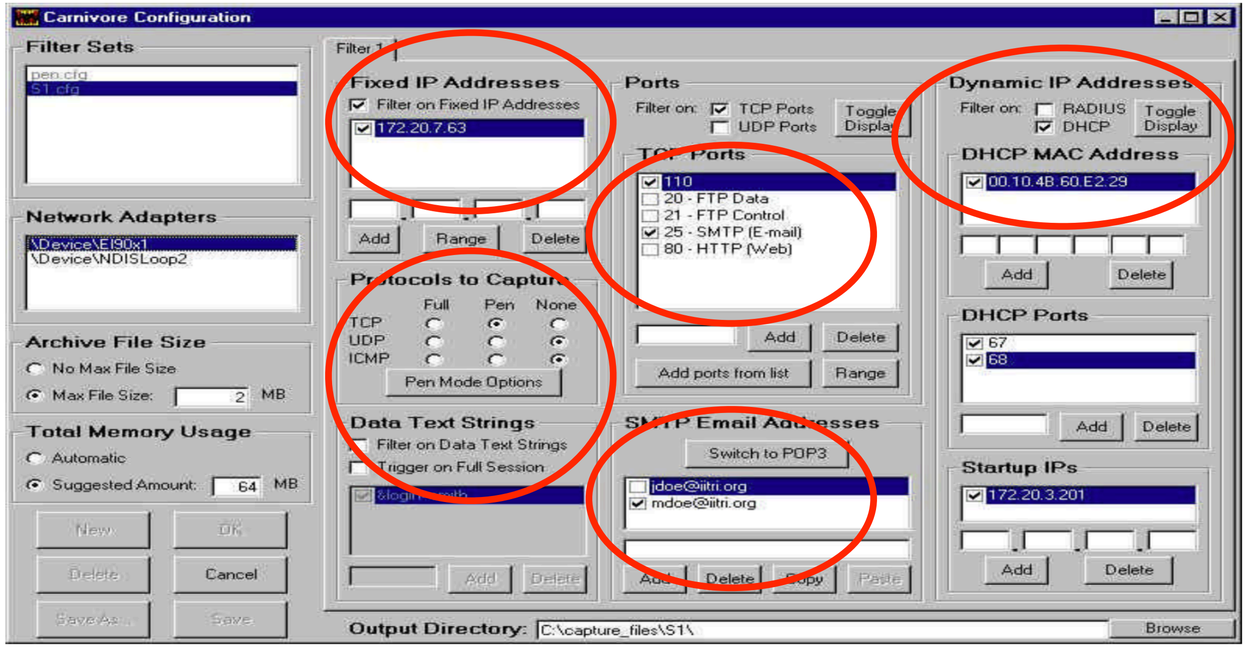
Filtering

What to Filter?

- How do you pick out the proper traffic to intercept?
- A wiretap warrant has to be particularized
- What are you filtering on?
- (Is it a Fourth Amendment search if a program looks at packets, only to discard them, if no person sees them? What if you do full-stream collection and filter it later?)

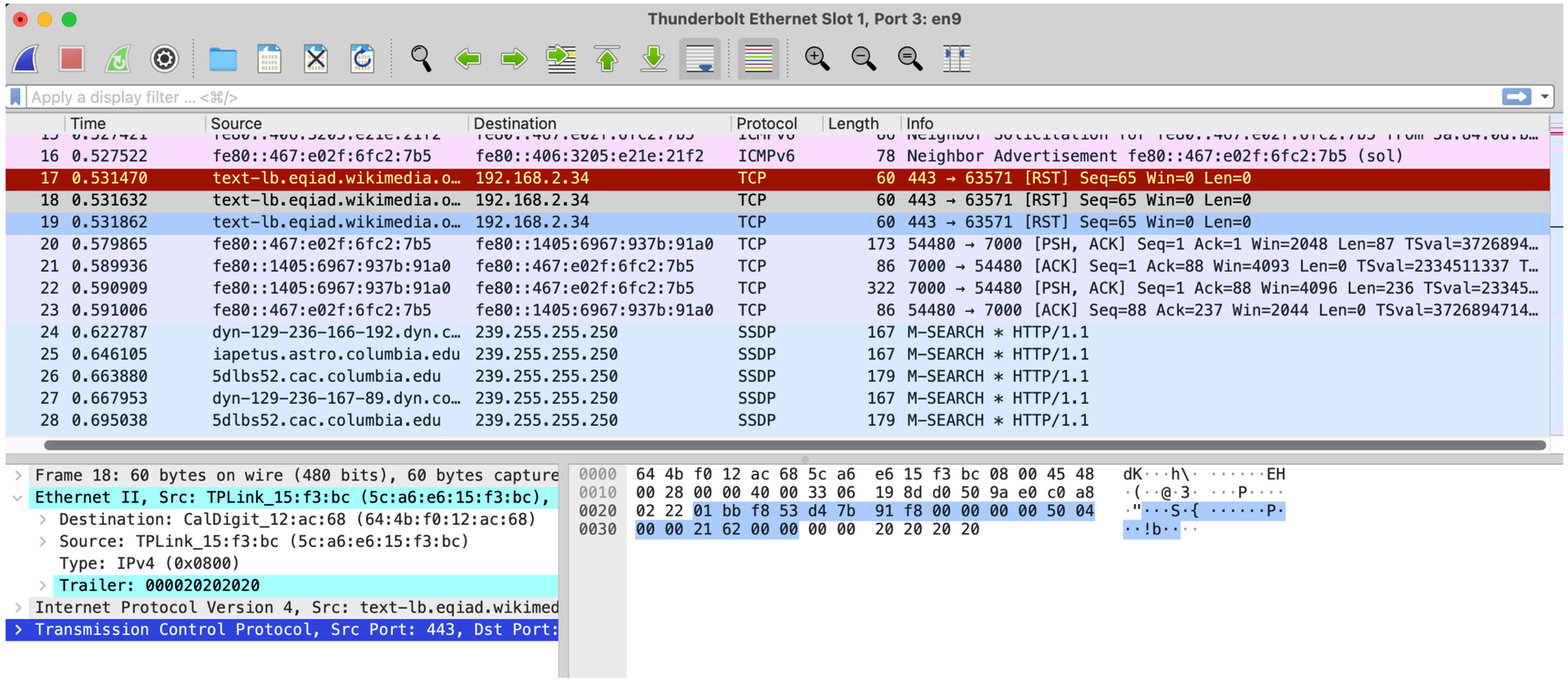
Carnivore

- About 25 years ago, the FBI had a tapping/ filtering system called *Carnivore*
  - Later renamed DCS1000
  - Now replaced by commercial software
- Note all of the filtering options
  - Fixed IP address or DHCP
  - TCP port numbers
  - Content vs. pen/trap
  - Email addresses
  - Other text strings in data



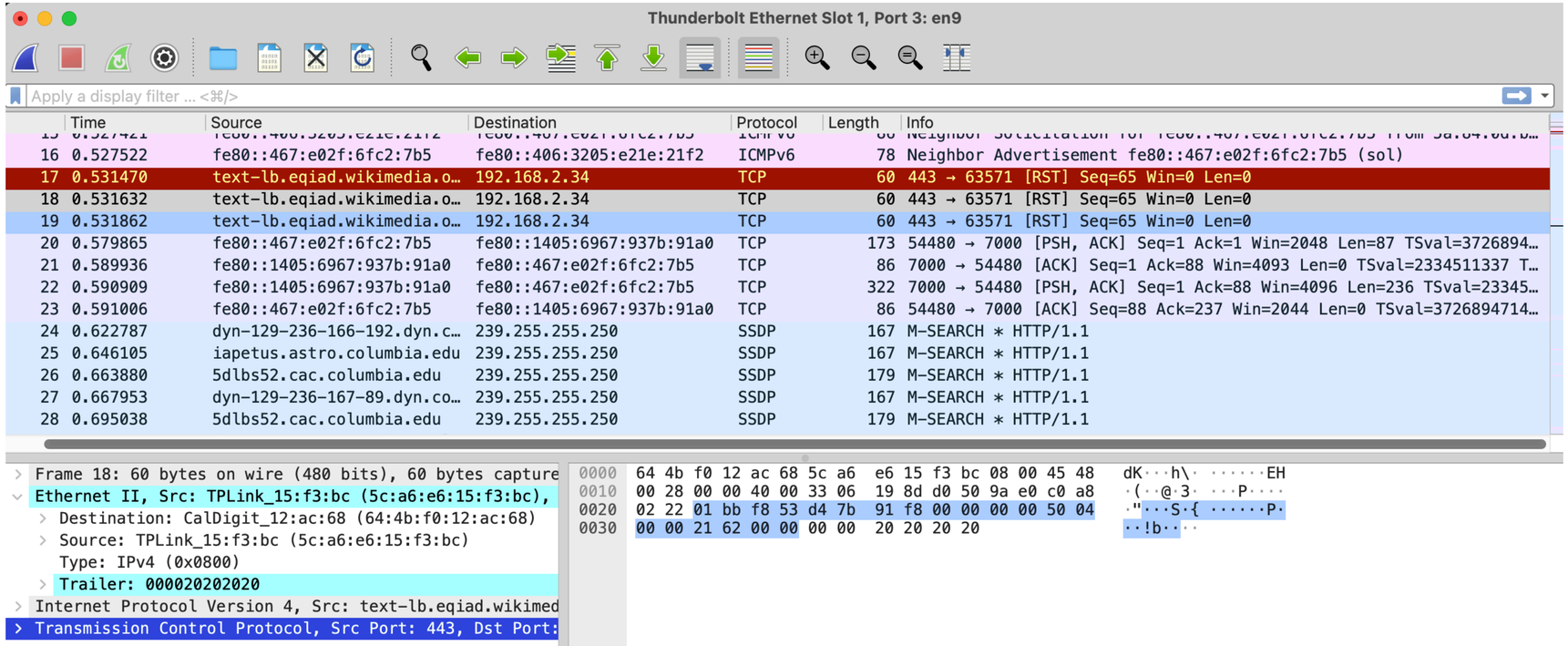
Screenshot from *Independent Technical Review of the Carnivore System—Final Report*
https://www.justice.gov/archive/jmd/carniv_final.pdf

Open Source



There is also a command line program, tcpdump, available on Linux and preinstalled on every Mac. (Why is this not barred by 18 U.S. Code § 2512?)

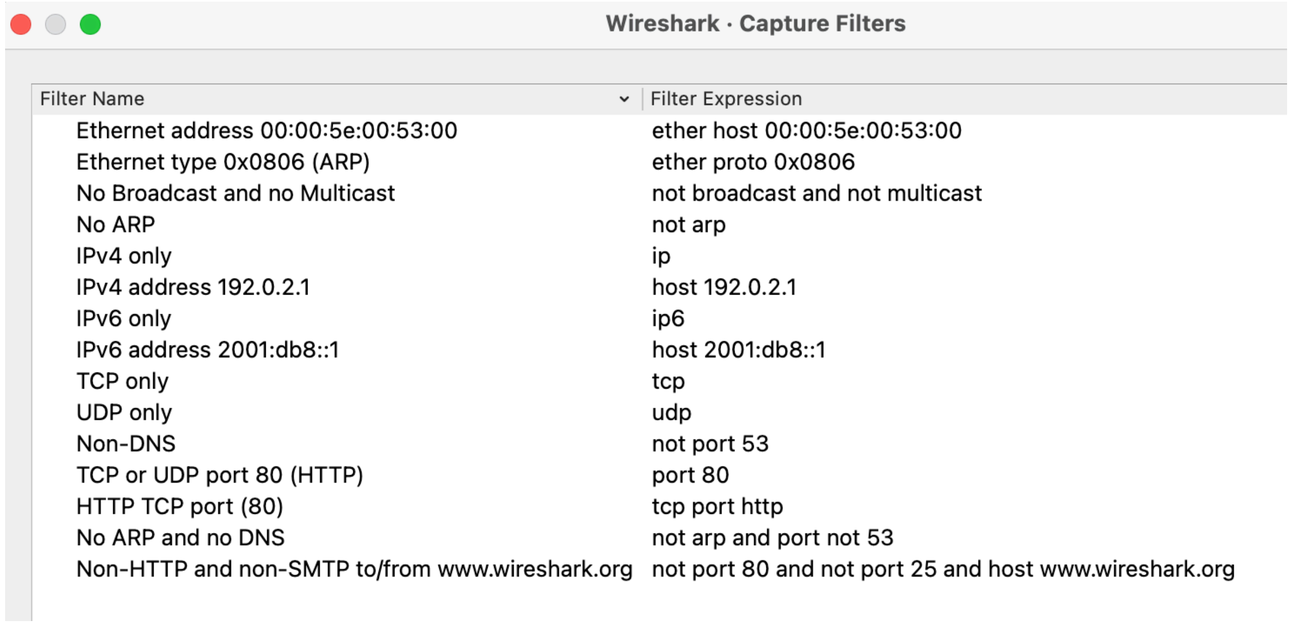
Open Source



There is also a command line program, tcpdump, available on Linux and preinstalled on every Mac. (Why is this not barred by 18 U.S. Code § 2512? The statute covers “device[s]… primarily useful for the purpose of the surreptitious interception”.)

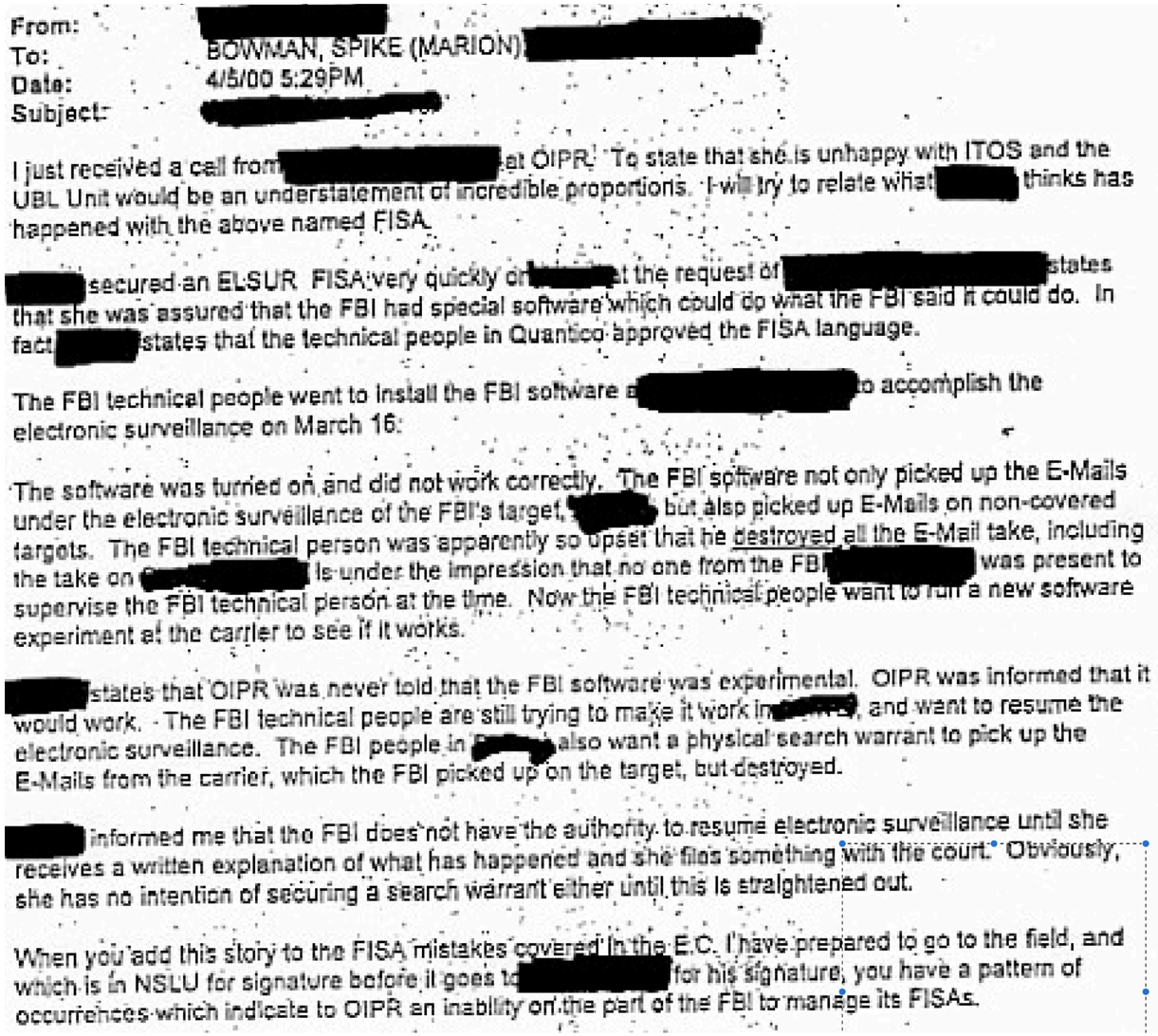
Wireshark Filtering

- All sorts of (complex) filters possible
- Not as easy to go from a MAC address to an IP address automatically by monitoring DHCP
- A standard tool used for monitoring and debugging networks and applications



Minimization

- One reason for filtering: *minimization*, making sure that taps pick up as little non-targeted conversation as possible
- Required by statute (18 U.S.C. 2518(5)) and case law (Scott v. United States, 436 U.S. 128)
- Touchstone: *reasonableness* of agents' conduct



From: [redacted]
To: BOWMAN, SPIKE (MARION) [redacted]
Date: 4/5/00 5:29PM
Subject: [redacted]

I just received a call from [redacted] at OIPR. To state that she is unhappy with ITOS and the UBL Unit would be an understatement of incredible proportions. I will try to relate what [redacted] thinks has happened with the above named FISA.

[redacted] secured an ELSUR FISA very quickly on [redacted] at the request of [redacted] states that she was assured that the FBI had special software which could do what the FBI said it could do. In fact [redacted] states that the technical people in Quantico approved the FISA language.

The FBI technical people went to install the FBI software at [redacted] to accomplish the electronic surveillance on March 16:

The software was turned on and did not work correctly. The FBI software not only picked up the E-Mails under the electronic surveillance of the FBI's target, [redacted] but also picked up E-Mails on non-covered targets. The FBI technical person was apparently so upset that he destroyed all the E-Mail take, including the take on [redacted] is under the impression that no one from the FBI [redacted] was present to supervise the FBI technical person at the time. Now the FBI technical people want to run a new software experiment at the carrier to see if it works.

[redacted] states that OIPR was never told that the FBI software was experimental. OIPR was informed that it would work. The FBI technical people are still trying to make it work in [redacted], and want to resume the electronic surveillance. The FBI people in [redacted] also want a physical search warrant to pick up the E-Mails from the carrier, which the FBI picked up on the target, but destroyed.

[redacted] informed me that the FBI does not have the authority to resume electronic surveillance until she receives a written explanation of what has happened and she files something with the court. Obviously, she has no intention of securing a search warrant either until this is straightened out.

When you add this story to the FISA mistakes covered in the E.C. I have prepared to go to the field, and which is in NSLU for signature before it goes to [redacted] for his signature, you have a pattern of occurrences which indicate to OIPR an inability on the part of the FBI to manage its FISAs.

http://www.epic.org/privacy/carnivore/fisa.html

Location Tracking

Where is Someone, and When?

- There are many ways to track someone's location
- They differ widely in prior knowledge required, accuracy, access required, and legal justification
- Some of these are untested legally

Location Tracking Mechanisms

- Phones
  - GPS
  - WiFi
  - Tower triangulation
  - Bluetooth Low Energy
- Telcos
  - Tower triangulation
  - Tower dumps
  - Cell site location information (CSLI)
- Trackers
  - Radio beacons
  - GPS devices
- IP geolocation

Phones and Location

- Phones have many ways of learning their location
  - WiFi—hear many WiFi access points; ask a server where those signals overlap
  - Bluetooth Low Energy—used for navigation within a building, due to its accuracy
  - Cell tower triangulation—how strong is the signal from several nearby cell sites?
  - GPS—listen to satellites

Phones and Location: Uses

- Many phone apps need location data
  - Mapping applications—where are you, to help you navigate?
  - Dating apps—meet someone nearby
  - Weather forecasts
  - Reminders—tell you when you enter a particular area
  - Many more
- E911—per FCC rules, phone *systems* must know the phone's location on 911 calls, to direct first responders
  - Can be implemented on the phone (the normal case, today) or via the phone network

Selling Location Data

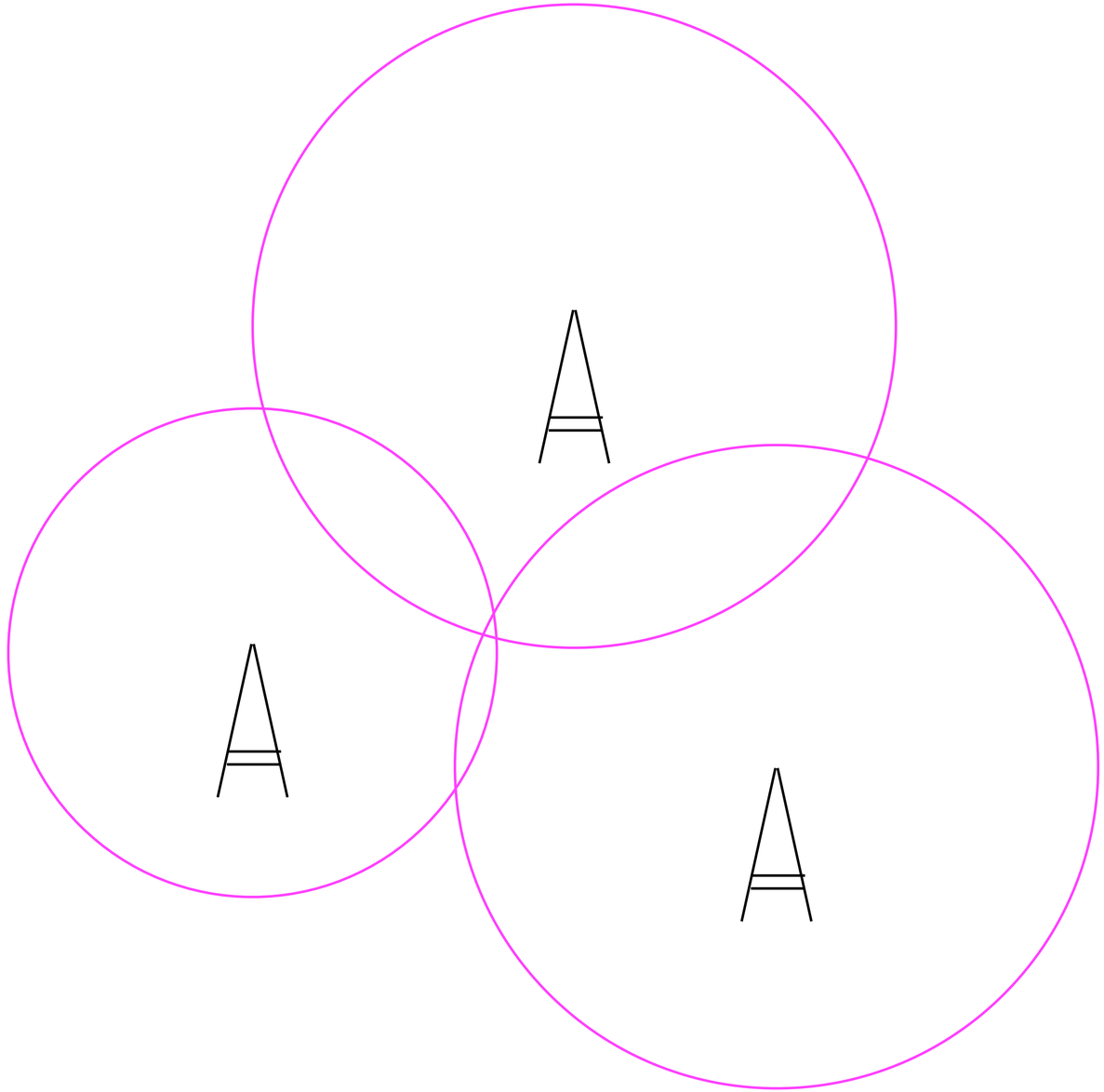
- Many app providers that collect location data sell it, including to government agencies
  - Phone companies and ISPs can't sell directly to government agencies: 18 U.S.C. §2702(a)(3)
  - App providers sell this data to data brokers, who resell it to lots of folks
- Sales to government agencies are controversial—see, e.g., S.2576/H.R.4639 — Fourth Amendment Is Not For Sale Act
- What about *Carpenter*?

Phone Companies

- Your phone connects to the cell site (“tower”) with the strongest signal
  - Stronger signals means that your phone can use less battery power to transmit
  - The phone network knows which tower you’re connected to, which *sector* (antenna, typically, out of three) is being used, and what signal strength is used
  - Phone companies keep such records for a few years, for traffic engineering
  - This is *CSLI—cell site location information*
- This means that the phone network *must* know where your phone is
  - It has to, anyway, to route calls to you
- If more than one tower hears your phone, they can triangulate

Triangulation

- Where do the circles overlap?
  - Circle radius depends on topography and signal strength used
- Today: ~1 mile radius in rural areas, much smaller in cities, and *microcells* in crowded, often indoor areas (shopping malls, hotel conference areas, etc.)



Tower Dumps

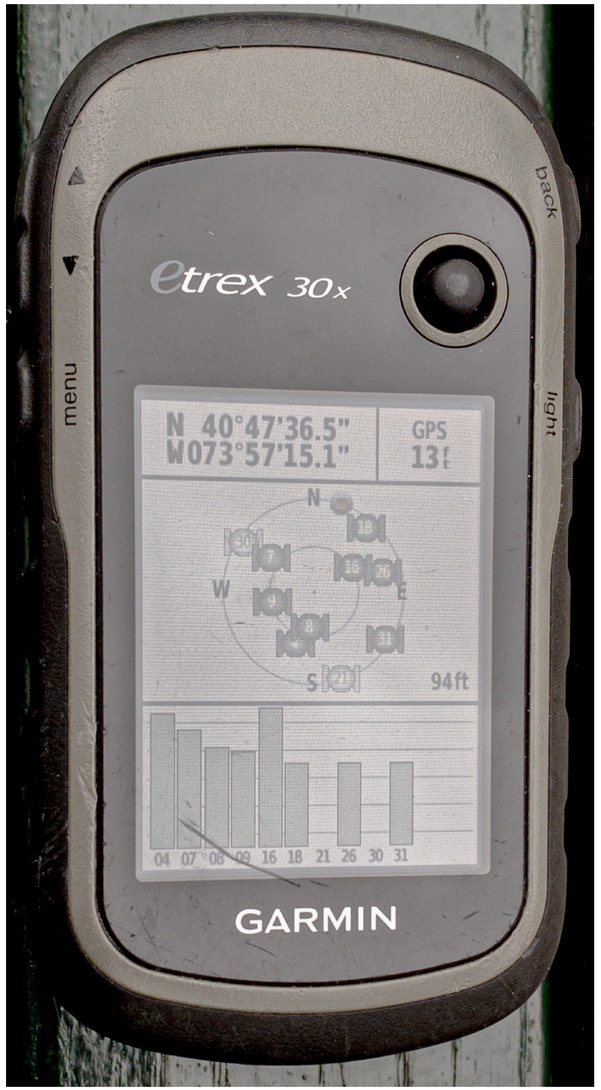
- Each cell cite knows what phones have connected to it
- It is possible to do a *tower dump*—find out what phones connected to a particular tower in a given interval
  - Often used to find suspects in a crime
- Legal status uncertain—is a geofence warrant sufficiently particularized?
  - How large is the region? How long is the time interval? How many non-suspects' phones will be found?
  - The Fourth Circuit has heard oral arguments in an appeal (22-4489) in *United States v. Chatrie*, 590 F. Supp. 3d 901 (2022); a Texas magistrate has differentiated a case there from *Chatrie* (*In re Info. That Is Stored at Premises Controlled by Google*, 2023 U.S. Dist. LEXIS 3365)—and geofence warrants played a major role in many January 6th cases
- Note: enterprise WiFi access points (and some consumer models) log the WiFi addresses of devices connecting to them

Physical Trackers

- Early tracking involved placing a radio transmitter on some vehicle or object, and locating it via *direction-finding*
  - Direction-finding goes back more than 100 years
- This requires physical access to whatever you want to track
- "A person traveling in an automobile on public thoroughfares has no reasonable expectation of privacy in his movements from one place to another." (*United States v. Knotts*, 460 U.S. 276 (1983))
- But: is a warrant necessary to attach the tracker?
  - See *United States v. Jones*, 565 U.S. 400 (2012)

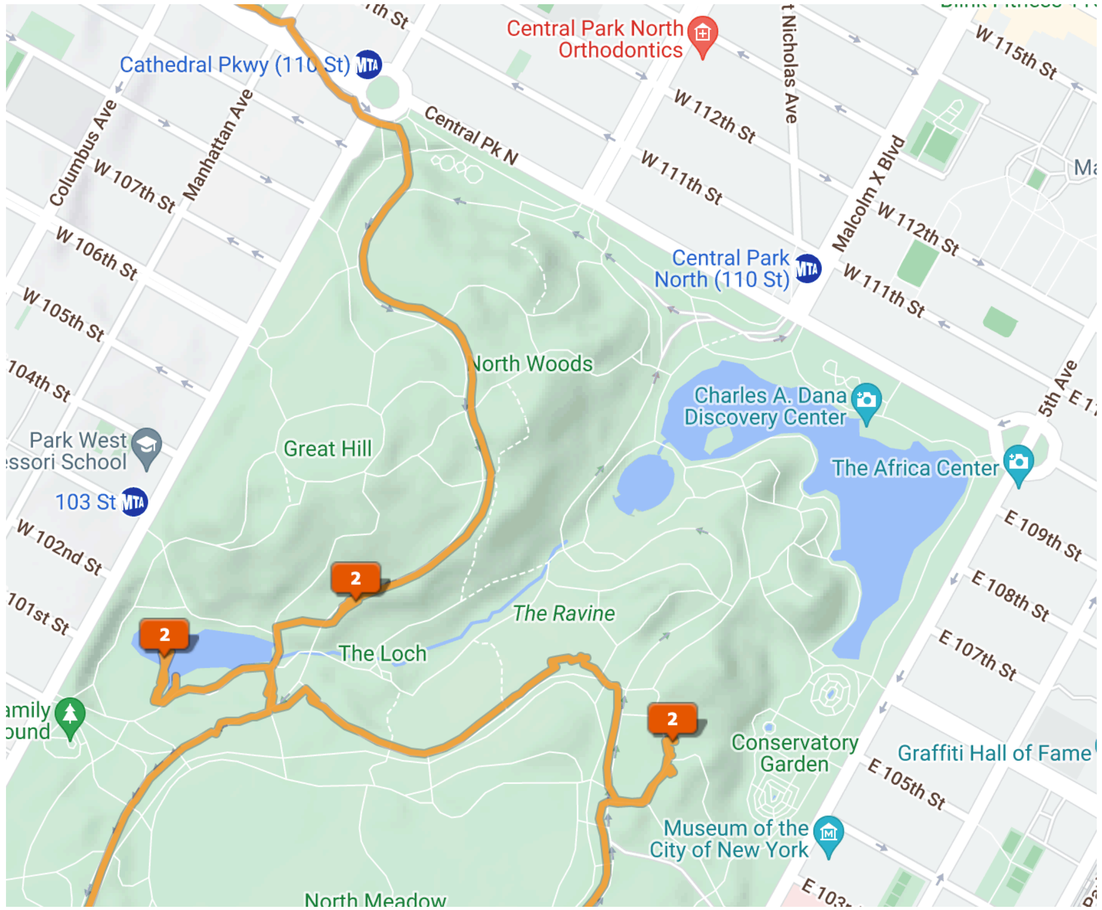
GPS Tracking

- GPS trackers listen to satellite signals
  - Note well: trackers *do not* transmit to the satellites
- The more satellites they hear, the more accurate the location will be
  - (This example was in Central Park)
- Law enforcement uses GPS trackers attached to cell phones, which relay the location



Track Loggings

- Even commercial GPS devices will log a user's track
  - (This is part of my wanderings through Central Park taking bird photos…)
- Entries are timestamped—GPS technology inherently requires extremely precise, accurate time
- Same physical access issues as in *Knotts* and *Jones*—but see the concurrences in *Jones*



GPS Errors

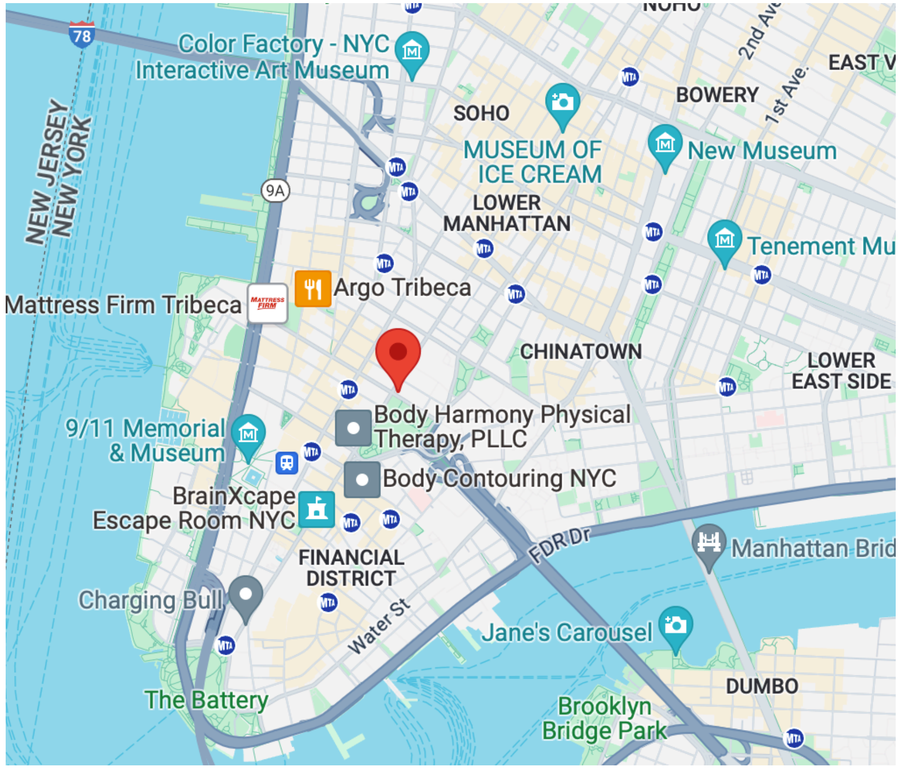
- GPS isn't perfect; there are numerous sources of error
- I did *not* venture into the Central Park Pond on this photo expedition
- The problem? Probably *multipath*—reflection of the satellite signal from the tall buildings on W 59th St.
- Other issues: GPS signals do not penetrate walls, roofs, leafy vegetation, etc., particularly well
- Canyons, natural or Manhattan, are quite challenging
- These cut down the number of satellites heard, and the geometry makes errors more likely

IP Geolocation

- Everyone talking on the Internet has an IP address
- Just like phone numbers, IP addresses are allocated hierarchically and geographically
- Any web site you contact knows approximately where you are, just based on your IP address
- (Google has better technology for this)

Neither of These is Where I Live…

But the city is correct



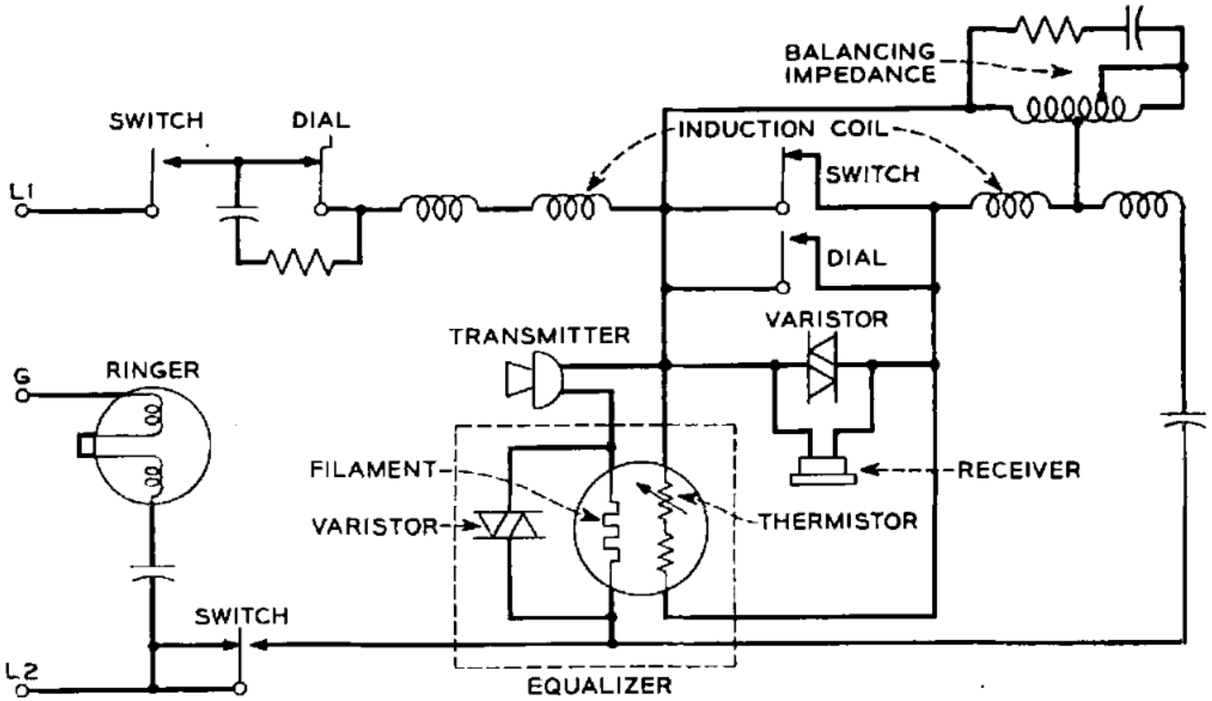
Physical Surveillance

Physical Surveillance

- Physical surveillance is still done
- Interesting questions—what is an improper enhancement to human senses?
  - Binoculars and telescopes? Probably ok; known to the Framers
  - Parabolic microphones? More doubtful, though some of the principles were known in ancient Athens
  - Alito's "tiny constables" in *Jones*?
  - Drug-sniffing dogs? Okay, per *Illinois v. Caballes*, 543 U.S. 405 (2005)
  - Infrared—an interesting case
  - Terahertz radiation—no case law (lack of standing in *Corbett v. City of New York*, 2013 U.S. Dist. LEXIS 204543)

Bugging a Room

- Can law enforcement use your devices as remote bugs?
- Not possible with old-fashioned telephones
- Not permitted with one model of car assistance feature—interfered with normal usage (*Company v. United States*, 349 F.3d 1132 (9th Cir. 2003))
- But: *U.S. v. Tomero*, 462 F. Supp. 2d 565 (S.D.N.Y. 2006), permitted turning a (dumb!) cellphone into a roving bug
- What about Alexa and kin? Your laptop?



From A. H. Inglis and W. L. Tuffnell, "An Improved Telephone Set", *Bell Sys. Tech. J.*, April 1951

CS @CU

Near Infrared

- Wavelengths a bit longer than visible red light—but invisible to people
- Used in night vision goggles—and in TV remote controls
- Some digital cameras can detect such light—but some phones filter out those signals using software

Far Infrared

- Detects heat emanations
  - Relatively new technology
  - Not in common use
- Technology at issue in *Kyllo*

Looking Through Walls

- Technology is being developed to see people through walls, using standard WiFi signals
- Obvious beneficial uses for firefighters, search and rescue teams, etc.
- Law enforcement? See *Kyllo*…

Third Party Data

Companies Have Records

- Stored records can be used to learn about a person
- Many simply require a subpoena (third party doctrine); others require a warrant (*Carpenter*)
- Many examples…

Bank Records

- Bank records may be available by simple subpoena
  - *United States v. Miller*, 425 U.S. 435 (1976)
- The same is likely true of other financial transactions
- “Follow the money”

Network Traffic—NetFlow

- Many ISPs collect *NetFlow* data—the traffic matrix
- The traffic matrix is a table showing which IP addresses sent how much data to which other IP addresses
  - Used for traffic engineering and sometimes billing
- Has a suspect visited a particular site? When? For how long? How much data was sent?
- Note: NetFlow data is based on random sampling, and hence can miss short connections

Search History

- Google, Bing, etc., keep records of folks' search histories
- How long this is retained can vary
- Always linked to a login, if any; may or may not be linked to IP address

Web Sites

- All web sites log all connections
- They have to, to catch operational errors
- But these logs are extremely revealing
  - How long are they retained?
  - Who sees them?
  - What legal process is necessary?

I heard you say

```
GET / HTTP/1.1
Host: greylock.cs.columbia.edu
User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10.15; rv:72.0) Gecko/20100101 Firefox/72.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate
DNT: 1
Connection: keep-alive
Upgrade-Insecure-Requests: 1
```

Purchase History

- Amazon and others keep detailed records of what people buy
- Has someone purchased “suspicious” items?
- Unknown if they can search for “who bought item X on date Y?”

Transportation Records

- The MTA records your Metrocard swipes and (probably) your OMNY taps
- E-Zpass records your driving
- Automatic license plate readers
- Have you bought train or bus tickets?
- Used a credit card to buy gas somewhere?
- An oil change on your car?
  - Your mechanic probably sold the car's mileage information

Data Brokers

- Data brokers aggregate information from many different sources
- They literally have thousands of data points on every adult American
- We all cast giant *data shadows*—and these paint a very complete picture of who we are and what we're doing

Questions?

December 27, 2022, Conowingo Dam, MD

CS@CU