CSE508 Network Security (PhD Section)

4/30/2015

Online Privacy and Anonymity

Michalis Polychronakis

Stony Brook University

Privacy

"The right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share information about itself with others." [RFC2828]

Anonymity

"The state of being not identifiable within a set of subjects, the anonymity set." [Pfitzmann and Köhntopp]

Very different from privacy:

An anonymous action may be public, but the actor's identity remains unknown (e.g., vote in free elections)

Real-world Privacy

Large-scale data collection examples

- Credit cards, Metrocards, Loyalty cards
- Street/public space cameras
- E-ZPass
- Named tickets
- …

Part of our everyday activities and personal information is (voluntarily or compulsorily) recorded

- Information from different sources can be **correlated**
- *Did you buy your Metrocard with your credit card?*

The same happens in the online world…

Third parties have access to…

Our email (Gmail, Yahoo, …)

Our files (Dropbox, Google Drive, …)

Our finances (e-banking, credit reporting, Mint, …)

Our communication (Skype, Facebook, …)

Our traffic (Wireless providers, ISPs, …)

Our location (3/4G, GPS, WiFi, …)

Our preferences ("Likes," Amazon, Netflix, …)

Our health (Fitbit, iWatch, …)

…

BUSINESS DAY

Millions of Anthem Customers Targeted in Cyberattack

By REED ABELSON and MATTHEW GOLDSTEIN FEB. 5, 2015

Outside the Anthem facility in Indianapolis. Anthem said it detected a data breach on Jan. 29, and that it was working with the Federal Bureau of Investigation. Aaron P. Bernstein/Getty Images

Anthem, one of the nation's largest health insurers, said late

Bits

SECURITY

Apple Says It Will Add New iCloud Security Measures After Celebrity Hack

By BRIAN X. CHEN SEPTEMBER 4, 2014 11:32 PM 21 Comments

PREVIOUS POST
Microsoft Introduces Three New Smartphones

NEXT POST
Daily Report: Apple Expected to Unveil Smartwatch and Larger iPhones

The Atlantic

The Netanyahu Disaster
By Jeffrey Goldberg

The Effects of Forgiveness
By Olga Khazan

Rural America's Silent Housing Crisis
By Gillian B. White

Introducing the Supertweet
By Ian Bogost

Armed With Facebook 'Likes' Alone, Researchers Can Tell Your Race, Gender, and Sexual Orientation

REBECCA J. ROSEN | MAR 12 2013, 2:59 PM ET

But the deeper aspects of your personality remain hard to detect.

Facebook knows you're gay before you do

3/20/13 4:29pm by Jon Green 39 Comments

Am I the only one creeped out that Facebook is now guessing, sometimes correctly, if its users are gay?

In the world of Big Data, our private lives are increasingly becoming intermingled with the shadowy, yet public, world of cyberspace.

Whenever we go online we are providing data that can be used to market to us; from Google searches to Facebook likes to eBay purchases, we are inputting data into a series of mathematical models which make *incredibly* educated guesses about the kinds of people we are.

Facebook creepily offers help to a gay guy thinking of "coming out"

Enter Matt. As BuzzFeed notes, Matt was your typical Facebook user who suddenly found an ad in his news feed for help in coming out. The weird thing was that Matt "did" need help coming out, and understandably he was more than a bit curious as to how Facebook knew.

At first, Matt wondered if Facebook had accessed his text messages, as he had confided in a close friend the previous

LATEST COMMENTS TAGS


Let's slow down this race, together: Starbuc and a bad hashtag
3/20/15 12:00pm 7 Comments


It's time to make college free
3/20/15 10:00am 19 Comments


Rick Perry's new adviser has suggested tha God isn't #ReadyForHillary. Technically, he right.
3/20/15 8:00am 14 Comments


Fatwas, gay sex tourism and the Indonesian LGBT underground
3/19/15 10:00am 6 Comments

Support AMERICAblog

Click here to donate securely via PayPal

We Recommend


Parallels between India's sexism and America's racism

Facebook

Facebook users unwittingly revealing intimate secrets, study finds

Personal information including sexuality and drug use can be correctly inferred from public 'like' updates, according to study

Most popular in US


Barcelona v Real Madrid: El Clásico - live! Jacob Steinberg


The eight best young adult books - and why grownups should read them, too


Singapore's Lee Kuan Yew dies aged 91

TECH 2/16/2012 @ 11:02AM | 2,698,356 views

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

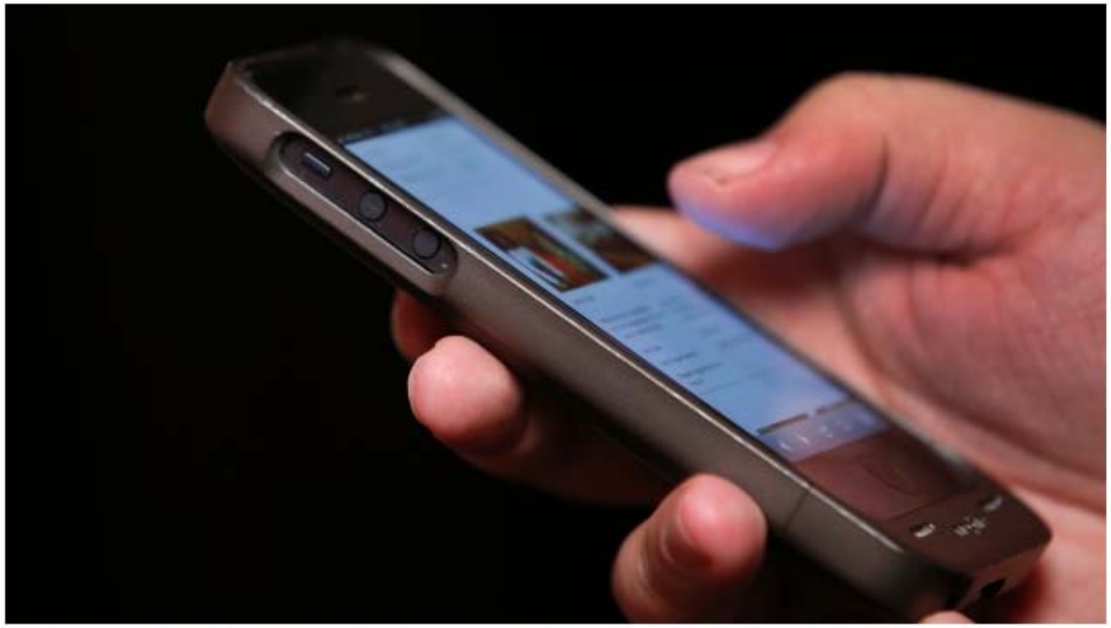
Target has got you in its aim

Charles Duhigg outlines in the New York Times how Target tries to hook parents-to-be at that crucial moment before they turn into rampant — and

Quebec resident Alain Philippon to fight charge for not giving up phone password at airport

Whether border officials can force you to provide password hasn't been tested in Canadian courts

By Jack Julian, CBC News Posted: Mar 04, 2015 9:32 PM AT | Last Updated: Mar 05, 2015 2:05 PM AT



The accused refused to divulge his smartphone password to Canada Border Services during a customs search. (Mike Segar/Reuters)

SSL-busting code that threatened Lenovo users found in a dozen more apps

"What all these applications have in common is that they make people less secure."

by Dan Goodin - Feb 22, 2015 3:45pm EST

AUTHORITIES LAUNCH MAN-IN-THE-MIDDLE ATTACK ON GOOGLE

Submitted by percy on Thu, Sep 04, 2014

WHAT HAPPENED?

From August 28, 2014 reports appeared on Weibo and Google Plus that users in China trying to access google.com and google.com.hk via CERNET, the country's education network, were receiving warning messages about invalid SSL certificates. The evidence, which we include later in this post, indicates that this was caused by a man-in-the-middle attack.

While the authorities have been blocking access to most things Google since June 4th, they have kept their hands off of CERNET, China's nationwide education and research network. However, in the lead up to the new school year, the Chinese authorities launched a man-in-the-middle (MITM) attack against Google.

We broke the news about the MITM attack on Github in January 2013. To borrow from that

Subscribe to our blog using RSS.

COMMENTS

Submitted by Marty on Mon, Sep 22, 2014
It's amazing too pay a quiick visit this site and reading
the views of all colleagues on tthe topic of this post, while I am also eager of gettihg knowledge.
Here is my page; effective weight, Marty,

Submitted by subway surfers ... on Sat, Sep 27, 2014
I'm gone to convey my little brother, that he should also pay a quick visit this web site on regular basis to obtain updated from most
recent gossip.

Submitted by Merissa on Sun, Sep 28, 2014
I think the admin of this site is genuinely working hard in support of his website, because here every stuff is

French agency caught minting SSL certificates impersonating Google

Unauthorized credentials for Google sites were accepted by many browsers.

by Dan Goodin - Dec 9 2013, 2:05pm EST

Web Browsing Tracking

Webpages are often mashups of content loaded from different sources

- Ads, images, videos, widgets, …
- IMG URLs, IFRAMEs, JavaScript, web fonts, Flash/applets, …
- Hosted on third-party servers: CDNs, cloud providers, ad networks, …

A third party involved in many different websites can track user visits across all those websites

- 2+ third parties may collude to expand their collective "view"

Need to learn two key pieces of information

- ***What webpage was visited***
- ***Who visited it***

Microsoft Announces Con

techcrunch.com/2015/04/29/microsoft-announces-continuum-turning-windows-10-phones-into-desktop

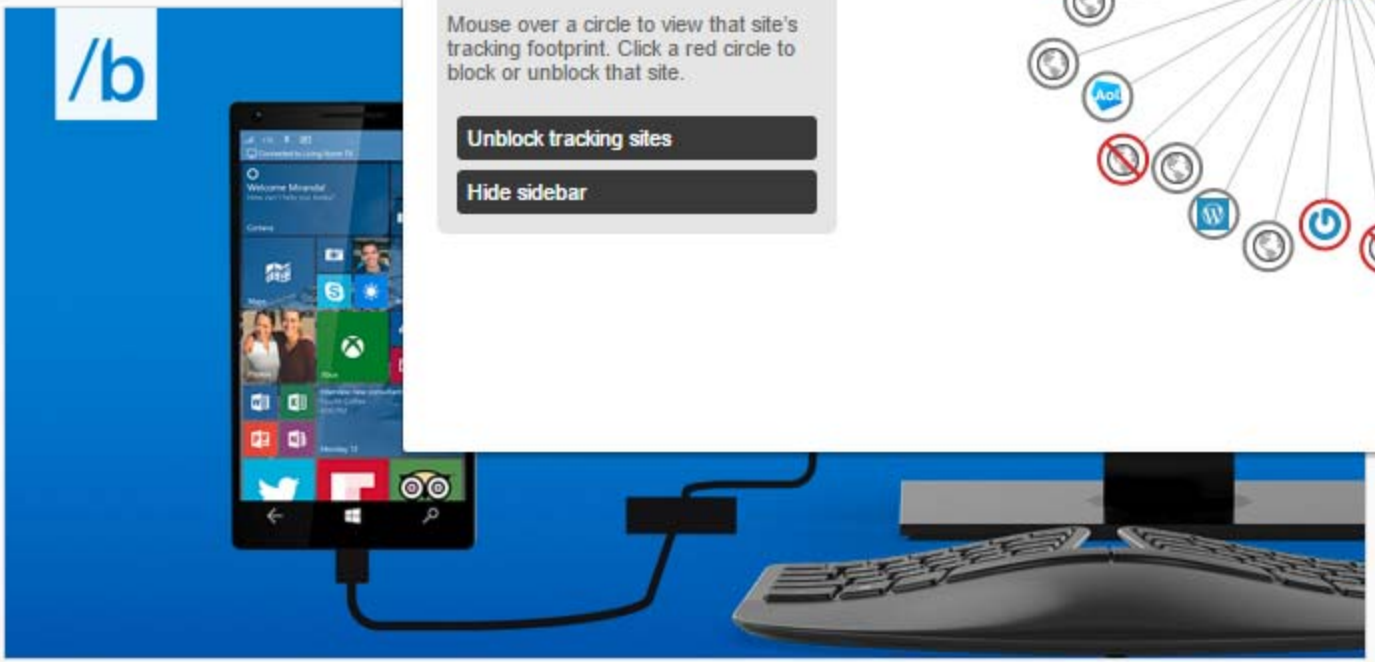
News Startups

Apps Microsoft Window

Microsoft An Phones Into

Posted 2 hours ago by Kyle

1,769 SHARES



Microsoft just demonstrated one of the intriguing possibilities from its single platform/multiple

DISCONNECT

Show list view

Browse the web normally. As you do, the graph in this popup and the counter in the toolbar will update. Each circle in the graph represents a site that's been or would've been sent some of your personal info.

Circles with a halo are sites you've visited. Circles without a halo are sites you haven't.

Red circles are known tracking sites. Gray circles aren't but may still track you.

Mouse over a circle to view that site's tracking footprint. Click a red circle to block or unblock that site.

Unblock tracking sites

Hide sidebar

What webpage was visited?

HTTP Referer [sic] header

- The URL of the webpage from which a link was followed
- Useful for statistics/analytics, bad for privacy
- Can be turned off trough browser options/extensions
- HTML5 rel="noreferrer" anchor attribute to indicate to the user agent not to send a referrer when following the link

Page-specific, session-specific, user-specific URLs

- Unique URL per page (even for the same resource) ➔ track what page was visited
- Unique URL per session/user ➔ distinguish between visits from different users

Tracking URLs are also Commonly used in promotional emails

Embedded image loading

- This is an active email address!
- Detect the time a user viewed a message
- The request reveals much more: user agent, device, location, …

Embedded links

- Learn which email addresses resulted in visits (click-through rate)

Default behavior of email clients varies

- Gmail used to block images by default, now uses image proxy servers
- Tracking through unique images still possible: senders can track the first time a message is opened (user's IP is not exposed though)

Who visited the page?

Browsing to a web page reveals a wealth of information

Source IP address

- Not very accurate (e.g., NAT, DHCP, on-the-go users) but still useful

Third-party cookies: precise user tracking

- Easy to block (configurable in most browsers, defaults vary)

"Evercookies:" exploit alternative browser state mechanisms

- Flash/Silverlight/other plugin-specific storage, ETags, HTML5 session/logal/global storage, caches, …

Browser/device fingerprinting: recognize unique system characteristics

- Browser user agent, capabilities, plugins, system fonts, screen resolution, time zone, and numerous other properties

What do web tracking techniques really track?

Distinguish between different visitors

- *Track anonymous individuals*
- Actually: track the pages visited by a particular browser running on a particular device

Better: distinguish between different *persons*

- ***Track named individuals***

The transition is easy…

Personally identifiable information (PII) is often voluntarily provided to websites:

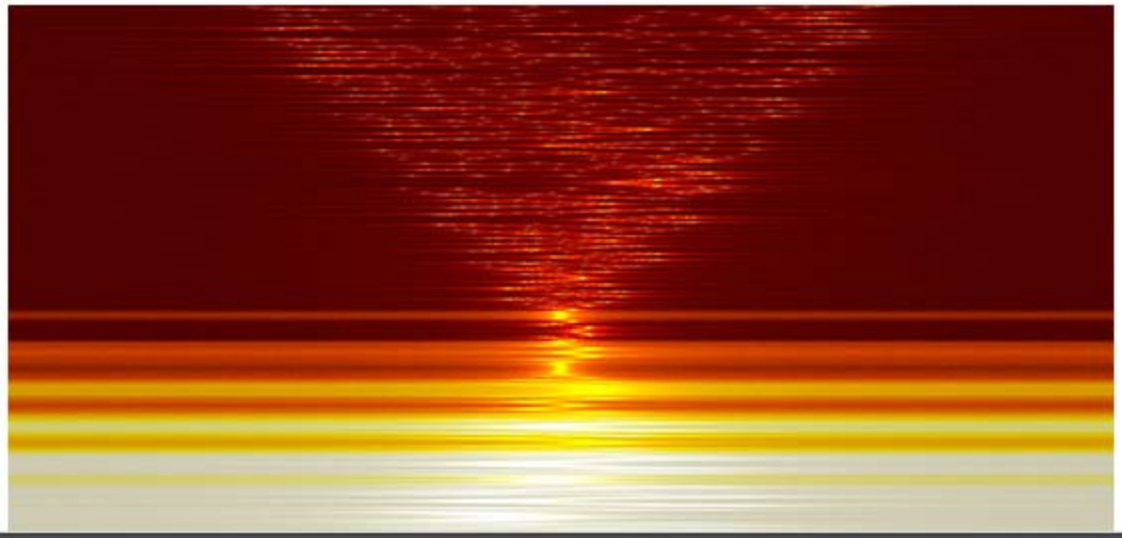
- Social networks, cloud services, web sites requiring user registration, …
- Cookies/sessions are associated with PII

Contamination: trackers may collude with services

- Previously "anonymous" cookies/fingerprints can be associated with named individuals

ROBERT MCMILLAN 10.27.14 6:30 AM

VERIZON'S 'PERMA-COOKIE' IS A PRIVACY-KILLING MACHINE



AT&T charges $29 more for gigabit fiber that doesn't watch your Web browsing

AT&T goes head to head against Google in KC on fiber and targeted ads.

by Jon Brodkin - Feb 16, 2015 12:38pm EST

AT&T

AT&T's gigabit fiber-to-the-home service has just arrived in Kansas City, and the price is the same as Google Fiber—if you let AT&T track your Web browsing history.

LATEST FEATURE STORY

FEATURE STORY (2 PAGES)

Battlefield Hardline review: an odd, cops-and-robbers facade

New twists on old formula help in multiplayer, baffle in single player.

WATCH ARS VIDEO

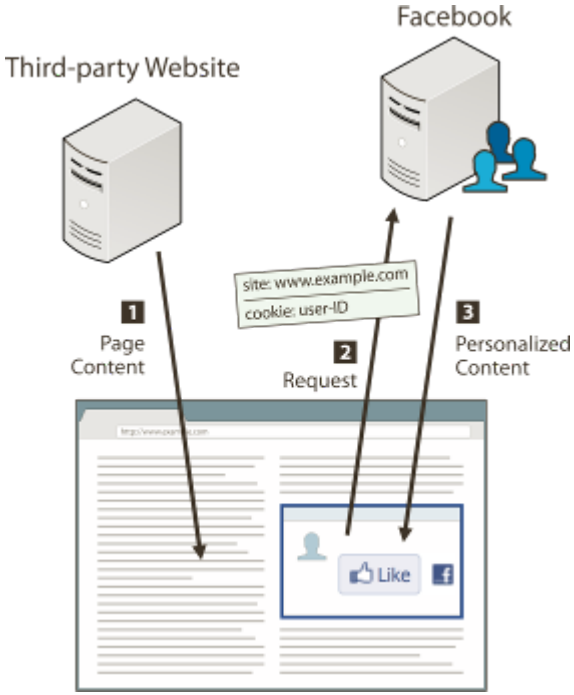
Users register on trackers!

Social plugins are prevalent

- 1.23+ billion Facebook users
- 33% of the top 10K websites have Like Buttons
- Twitter, Google+, LinkedIn, Pinterest, AddThis, …
- OS/app integration

A growing part of our browsing history can be tracked

- Not as merely anonymous visitors, but as ***named persons***
- Just visiting the page is enough (no interaction needed)
- Cross-device tracking



Existing Solutions

Log out

- Some cookies persist

Block third-party cookies

- Not always effective

Block social widgets completely

Incognito mode

All existing solutions disable content personalization

- Privacy vs. functionality dilemma

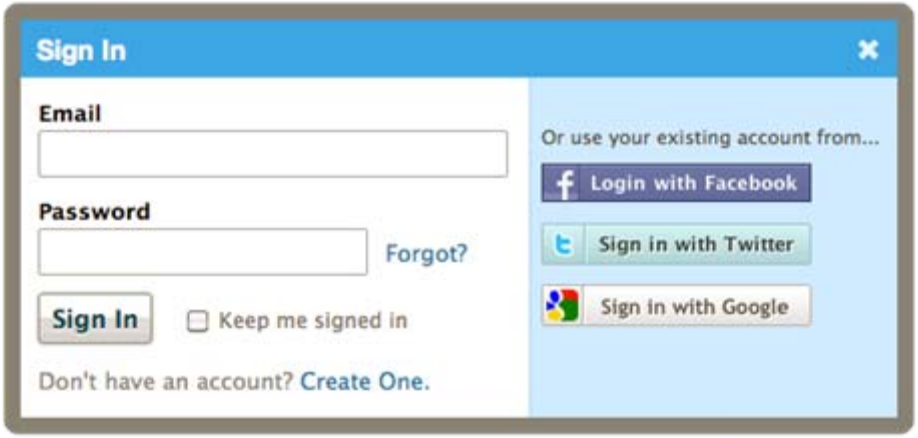
Single Sign-on/Social Login

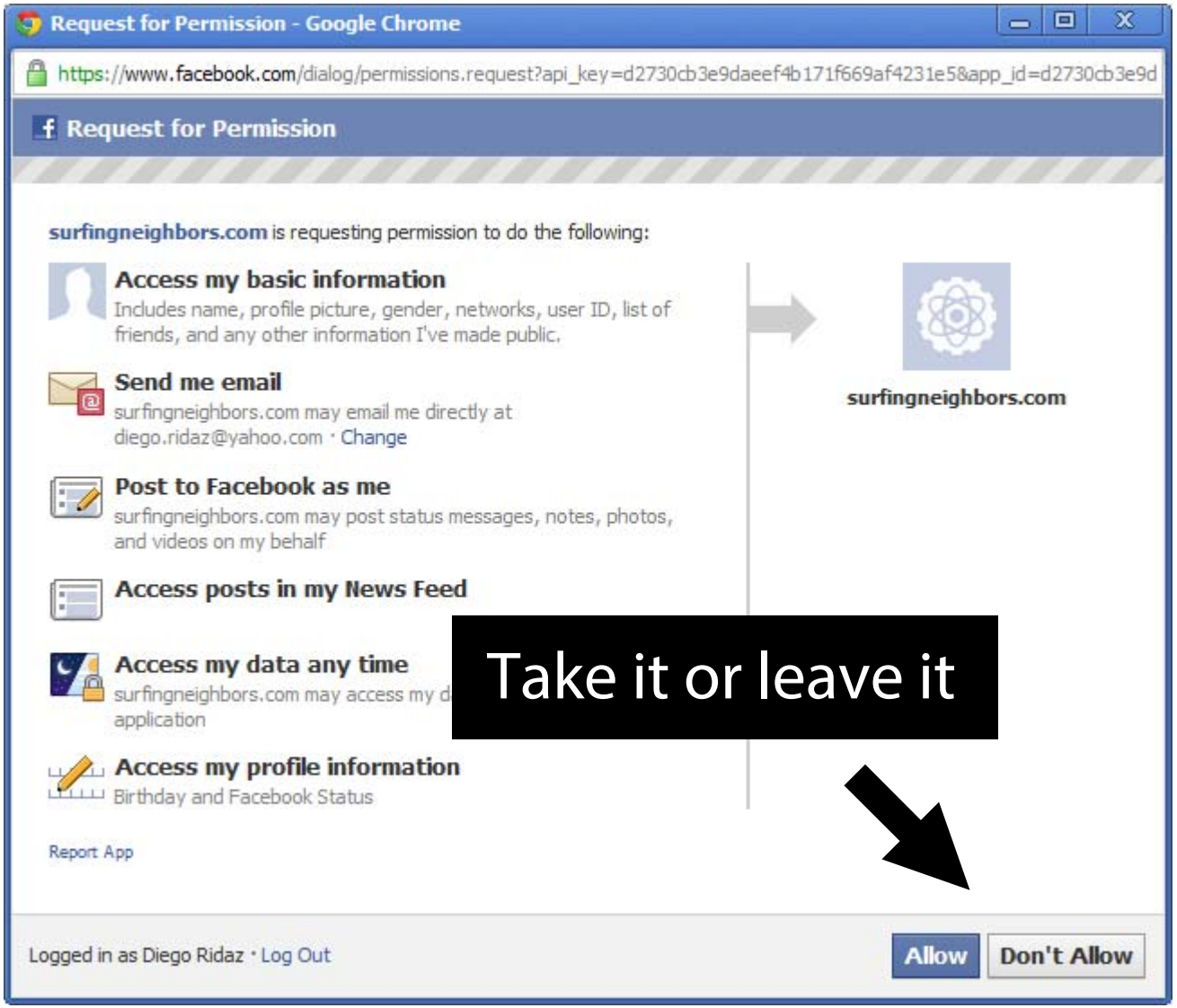
Pros

- Convenience – fewer passwords to remember
- Rich experience through social features
- Outsource user registration and management

Cons

- Same credentials for multiple sites
- User tracking
- Access to user's profile




Request for Permission - Google Chrome
https://www.facebook.com/dialog/permissions.request?api_key=d2730cb3e9daeef4b171f669af4231e5&app_id=d2730cb3e9d
Request for Permission
surfingneighbors.com is requesting permission to do the following:
Access my basic information
Includes name, profile picture, gender, networks, user ID, list of friends, and any other information I've made public.
Send me email
surfingneighbors.com may email me directly at diego.ridaz@yahoo.com · Change
Post to Facebook as me
surfingneighbors.com may post status messages, notes, photos, and videos on my behalf
Access posts in my News Feed
Access my data any time
surfingneighbors.com may access my d
application
Access my profile information
Birthday and Facebook Status
Report App
surfingneighbors.com
Take it or leave it
Logged in as Diego Ridaz · Log Out
Allow
Don't Allow

Location Tracking

IP addresses reveal approximate location information

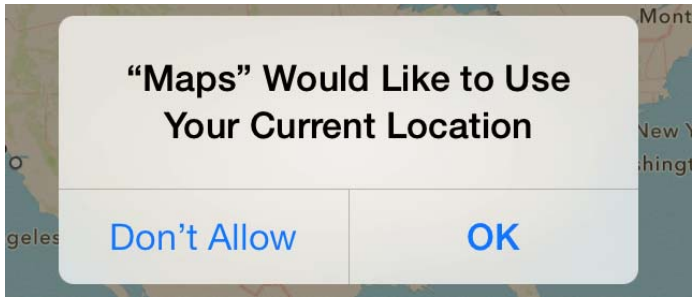
MaxMind statistics: 99.8% accurate on a country level, 90% accurate on a state level in the US, and 81% accurate for cities in the US within a 50 kilometer radius

Mobile devices allow for precise location tracking

Cell tower triangulation/trilateration

GPS

WiFi access points in known locations



Per-app permissions

Android vs. iOS:
installation vs. usage time

BUSINESS DAY

410 COMMENTS

Attention, Shoppers: Store Is Tracking Your Cell

By STEPHANIE CLIFFORD and QUENTIN HARDY JULY 14, 2013

Like dozens of other brick-and-mortar retailers, Nordstrom wanted to learn more about its customers — how many came through the doors, how many were repeat visitors — the kind of information that e-commerce sites like Amazon have in spades. So last fall the company started testing new technology that allowed it to track customers' movements by following the Wi-Fi signals from their smartphones.

Brick-and-mortar stores are looking for a chance to catch up with their online competitors by using software that allows them to watch customers as they shop, and gather data about their behavior. Video by Erica Berenstein on July 14, 2013.

But when Nordstrom posted a sign telling customers it was tracking them, shoppers were unnerved.

"We did hear some complaints," said Tara Darrow, a spokeswoman for the store. Nordstrom ended the experiment in May, she said, in part because of the comments.

Nordstrom's experiment is part of a movement by retailers to gather data about in-store shoppers' behavior and moods, using video surveillance and signals from their cellphones and apps to learn

Online Behavioral Tracking

An increasing part of our daily activities are recorded

- What we are interested in (Searches, Likes, …)
- What we read (News, magazines, blogs, …)
- What we buy (Amazon, Freshdirect, …)
- What we watch (Netflix, Hulu, …)
- What we eat (Seamless, GrubHub, …)
- Where we eat (Opentable, Foursquare, …)
- Where we go (online travel/hotel/event booking)
- What we own/owe (e-banking, credit services, Mint, …)

Mobile apps make behavioral tracking easier and more accurate

Behavioral profiles have desirable and not so desirable uses

- Recommendations, content personalization, …
- Targeted advertising, price discrimination (e.g., insurance premiums based on past behavior, higher prices for high-end device users), …

Health and Activity

Health records

- How securely are they handled and stored?

Devices track our activities and health

- Activity tracking devices
- Health monitoring devices
- Mobile phones

Many upload all data to the “cloud”…

- Who can access them?

Anonymous communication

Sender anonymity

the identity of the party who sent a message is hidden, while its receiver (and the message itself) might not be

Receiver anonymity

the identity of the receiver is hidden

Unlinkability of sender and receiver

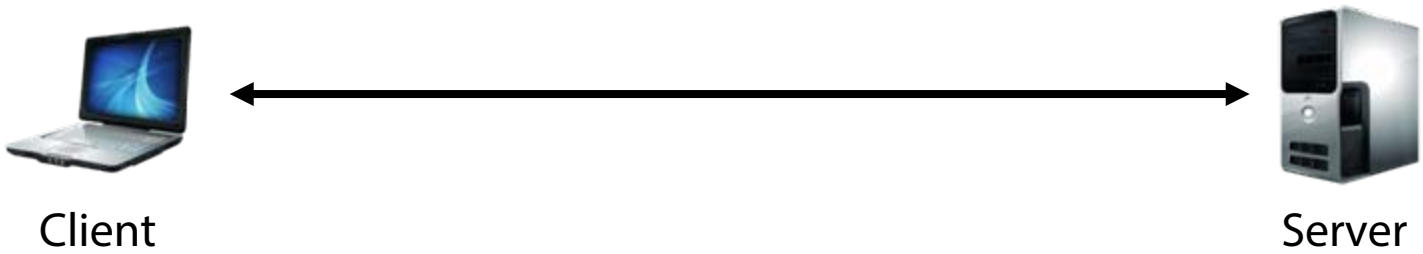
Although the sender and receiver can each be identified as participating in some communication, they cannot be identified as communicating with each other

The internet was not designed for anonymity

Packets have source and destination IP addresses

Using pseudonyms to post anonymously is not enough…

Server always sees the IP address of the client



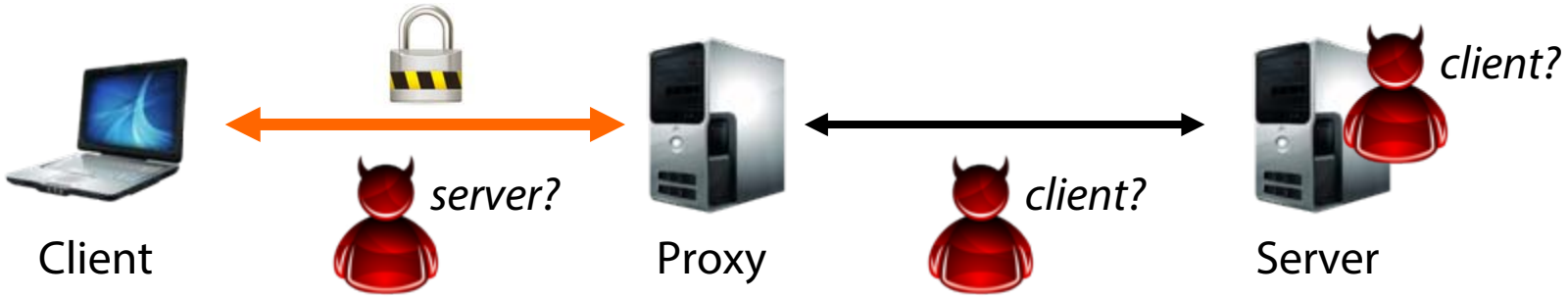
Need to hide the source IP address

Assuming no other PII is revealed (!) – OPSEC is hard

Stepping Stones: Anonymity

Remote proxies, relays, VPN services

- Server sees only the IP address of the proxy
- Since the proxy cooperates, let's also encrypt the connection to it



Sender anonymity against the server and network observers beyond the proxy

Also: receiver anonymity against local observers

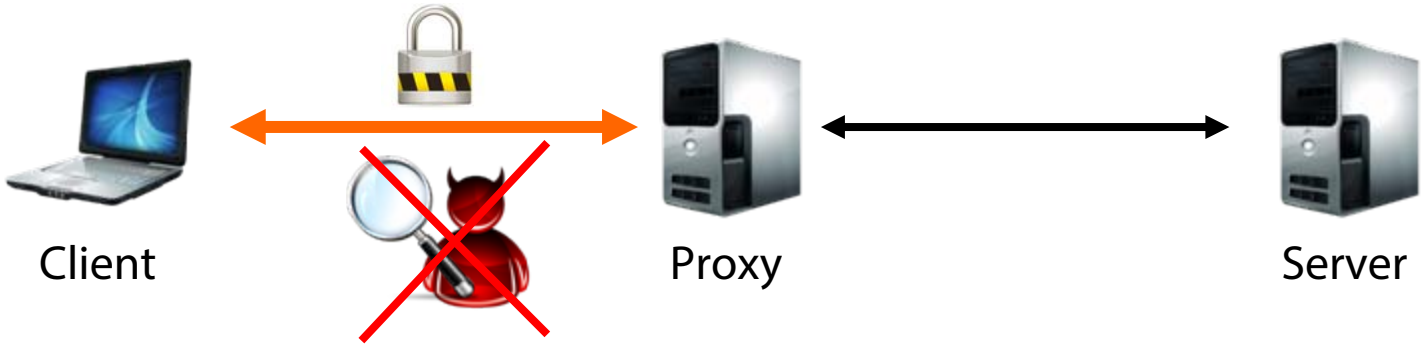
- All they can see is client <-> proxy connections
- Encryption hides the actual destination

Stepping Stones: Traffic Protection

Besides anonymity, the encrypted client <-> proxy channel offers protection against local adversaries

The definition of "local" depends on the location of the proxy

- Users in the same LAN, employer's admins, ISPs, governments, …



Protection against passive and active network adversaries (eavesdropping, MitM, MotS, …)

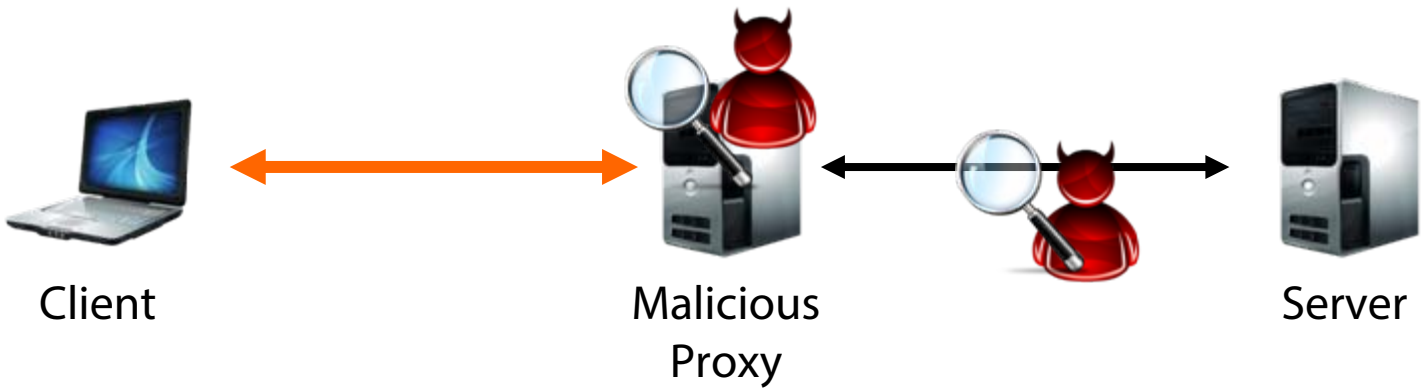
Policy and censorship circumvention

- Parental controls, company-wide port/domain/content blocking, hotel WiFi restrictions, government censorship, …

What about other adversaries?

The proxy itself may be the adversary – can see it all!

Network observers beyond the proxy can see it all!



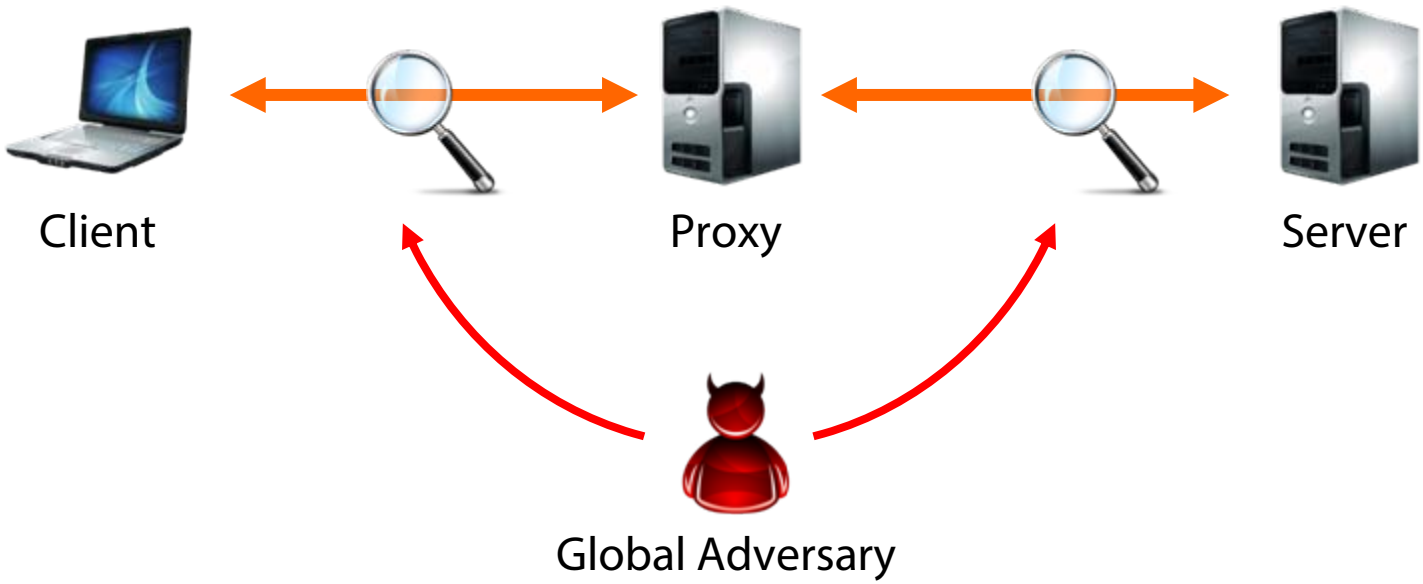
Adversaries who couldn't eavesdrop before, now can: just set up a rogue proxy and lure users

End-to-end encryption is critical!

What about other adversaries?

A “global” adversary may be able to observe both ends

Traffic analysis: communication patterns can be observed even when end-to-end encryption is used



Eavesdropping vs. Traffic Analysis

Even when communication is encrypted, the mere fact that two parties communicate reveals a lot

What can we learn from phone records?

- Who communicated with whom and when
- Activity patterns (periodic, time of day, occasional, …)
- Single purpose numbers (hotlines, agencies, doctors, …)

It's not "just metadata"…

Network traffic analysis can reveal a lot

Passive traffic analysis

Frequency and timing of packets, packet sizes, amount of transferred data, …

Active traffic analysis

Packet injection, fingerprint injection through manipulation of traffic characteristics, …

Examples:

Message timing correlation to learn who is talking to whom

Visited HTTPS web pages through structural analysis (number/size of embedded elements etc.)

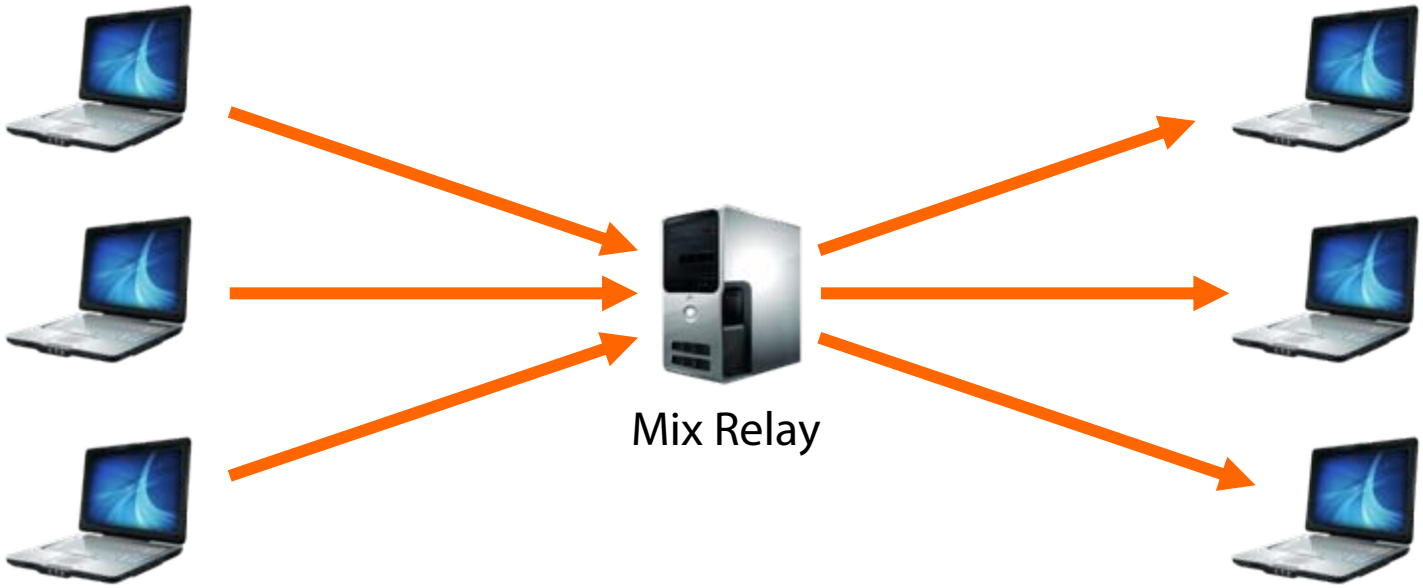
SSH keystroke timing analysis

"Traffic analysis, not cryptanalysis, is the backbone of communications intelligence."

— Susan Landau and Whitfield Diffie

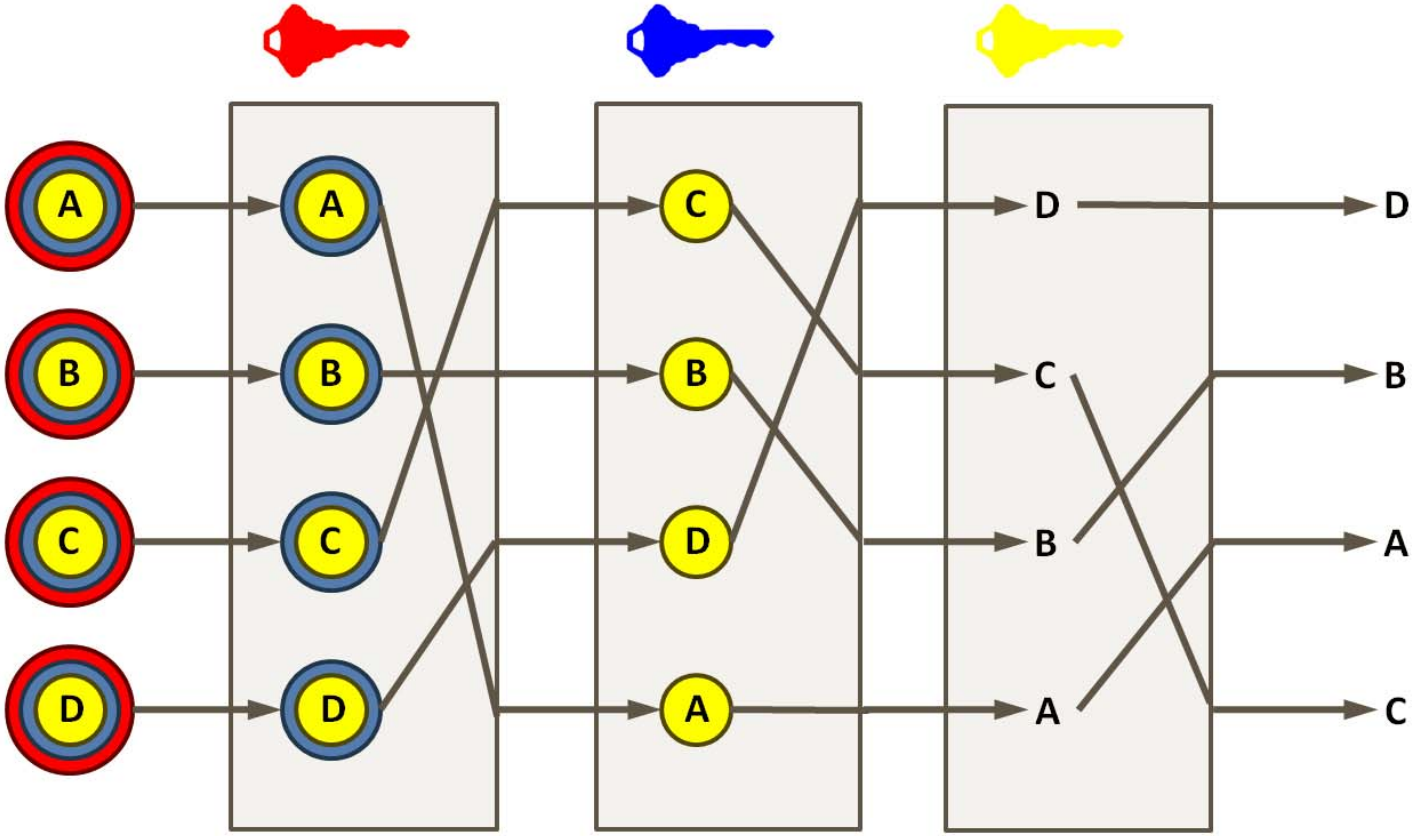
Mix Networks [Chaum 1981]

Main idea: hide own traffic among others' traffic



Originally conceived for anonymous email: Trusted remailer + public key cryptography

Additional measures are critical for thwarting traffic analysis: message padding, delayed dispatch, dummy traffic



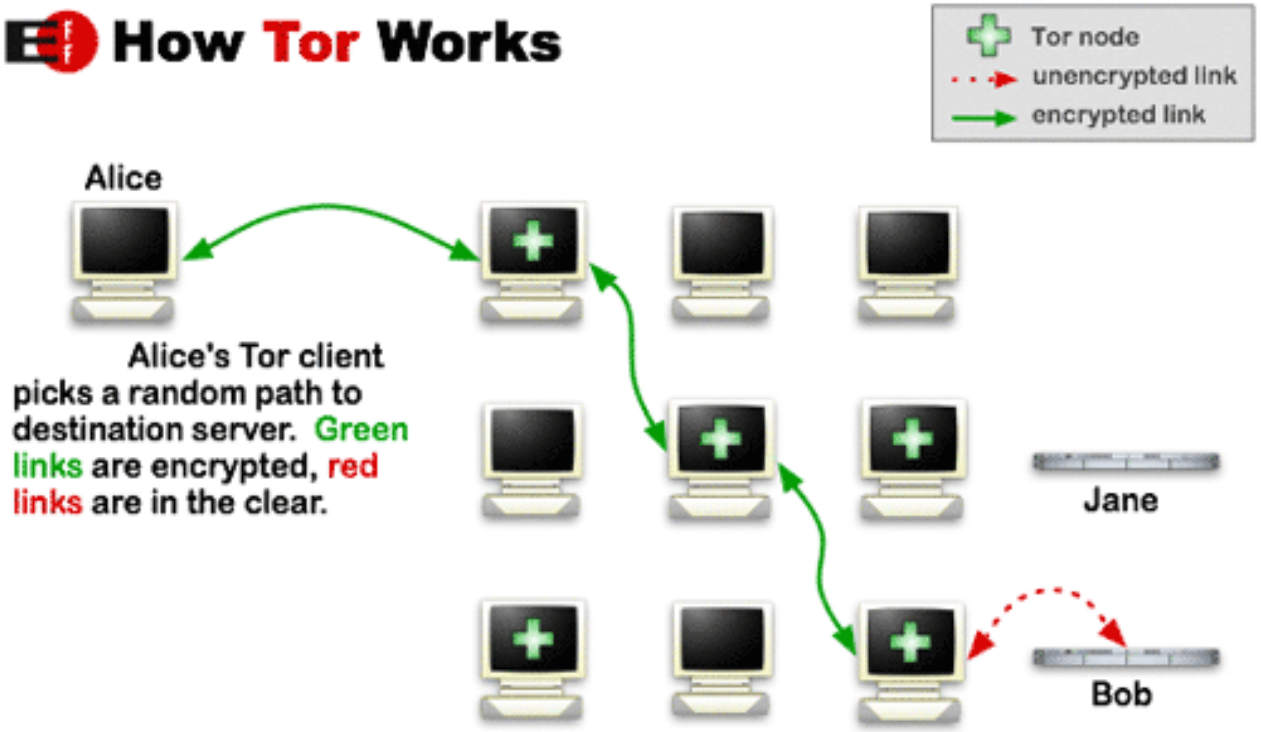
Adding multiple relays allows for anonymity even if some relays are controlled by an adversary

- Deanonymization still possible if an attacker controls *all* relays of a circuit

Tor (aka. the Onion Router)

Low-latency anonymous communication network

Layered encryption: each relay decrypts a layer of encryption to reveal only the next relay



Worldwide volunteer network of 6K+ relays

- More than 2M daily users

Three-hop circuits by default

- Entry node, middleman, exit node
- Longer circuits can be built
- Multiple connections can be multiplexed over the same Tor circuit

Directory servers point to active Tor relays

- ~10 directory servers hard-coded into the Tor client
- Monitoring for mass subscriptions by potential adversaries (sybil attack)

Applications

User-friendly Tor Browser

- Additional measures to thwart web tracking and fingerprinting

TAILS (The Amnesic Incognito Live System) Linux distribution

- Forces all outgoing connections to go through Tor

Hidden services: hide the IP of servers

- .onion pseudo top-level domain host suffix
- Not always easy: misconfigurations and leaks may reveal the real IP address of the server

SecureDrop (originally designed by Aaron Swartz)

- Platform for secure anonymous communication between journalists and sources (whistleblowers)

Censors want to block Tor

Directory servers are the easy target

- Block any access to them

Response: Tor bridges

- Tor relays that aren't listed in the main Tor directory
- Only a few at a time can be obtained on-demand (e.g., through email to bridges@bridges.torproject.org)
- Once known, adversaries may block them too…

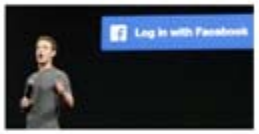
Pluggable Transports

- Censors may drop all Tor traffic through deep packet inspection
- Hide Tor traffic in plain sight by masquerading it as some other innocent-looking protocol (HTTP, Skype, Starcraft, …)

FBI Admits It Controlled Tor Servers Behind Mass Malware Attack

BY KEVIN POULSEN 09.13.13 | 4:17 PM | PERMALINK

MOST RECENT WIRED POSTS


Facebook Just Had Another Record Quarter, and It Has Apple to Thank

Comcast Renames Man 'Asshole Brown' After He Tries to Cancel Cable


A Heroin Dealer Tells the Silk Road Jury What It Was Like to Sell Drugs Online

Amazon Challenges Google and Microsoft With Its Own Email Service


These Are the Hottest New Open Source Projects Right Now

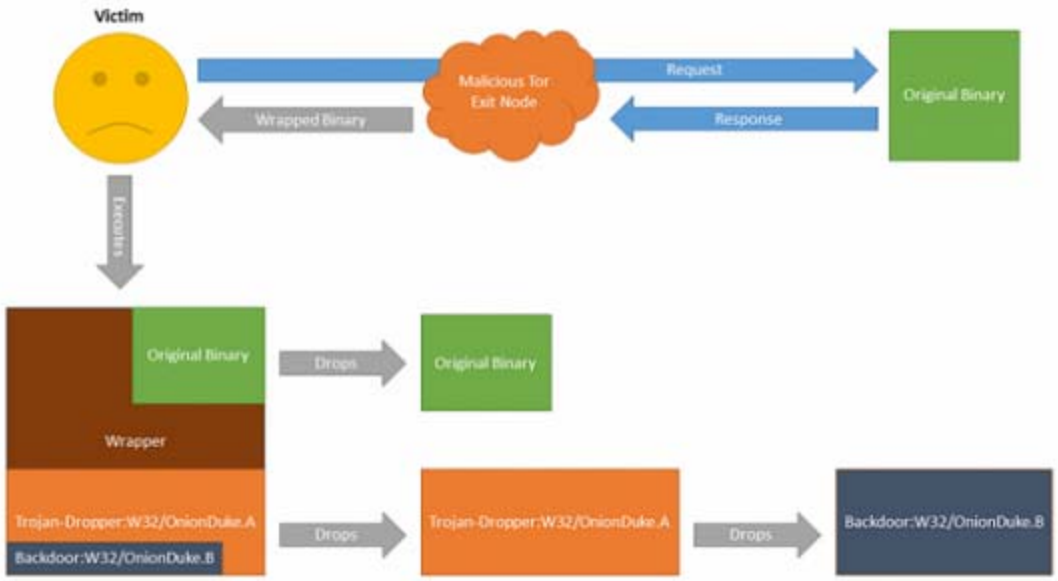

Canada Joins World Powers in

RISK ASSESSMENT / SECURITY & HACKTIVISM

For a year, gang operating rogue Tor node infected Windows executables

Attacks tied to gang that previously infected governments with highly advanced malware.

by Dan Goodin - Nov 14, 2014 10:30am EST



Enlarge / A flowchart of the infection process used by a malicious Tor exit node.

F-Secure

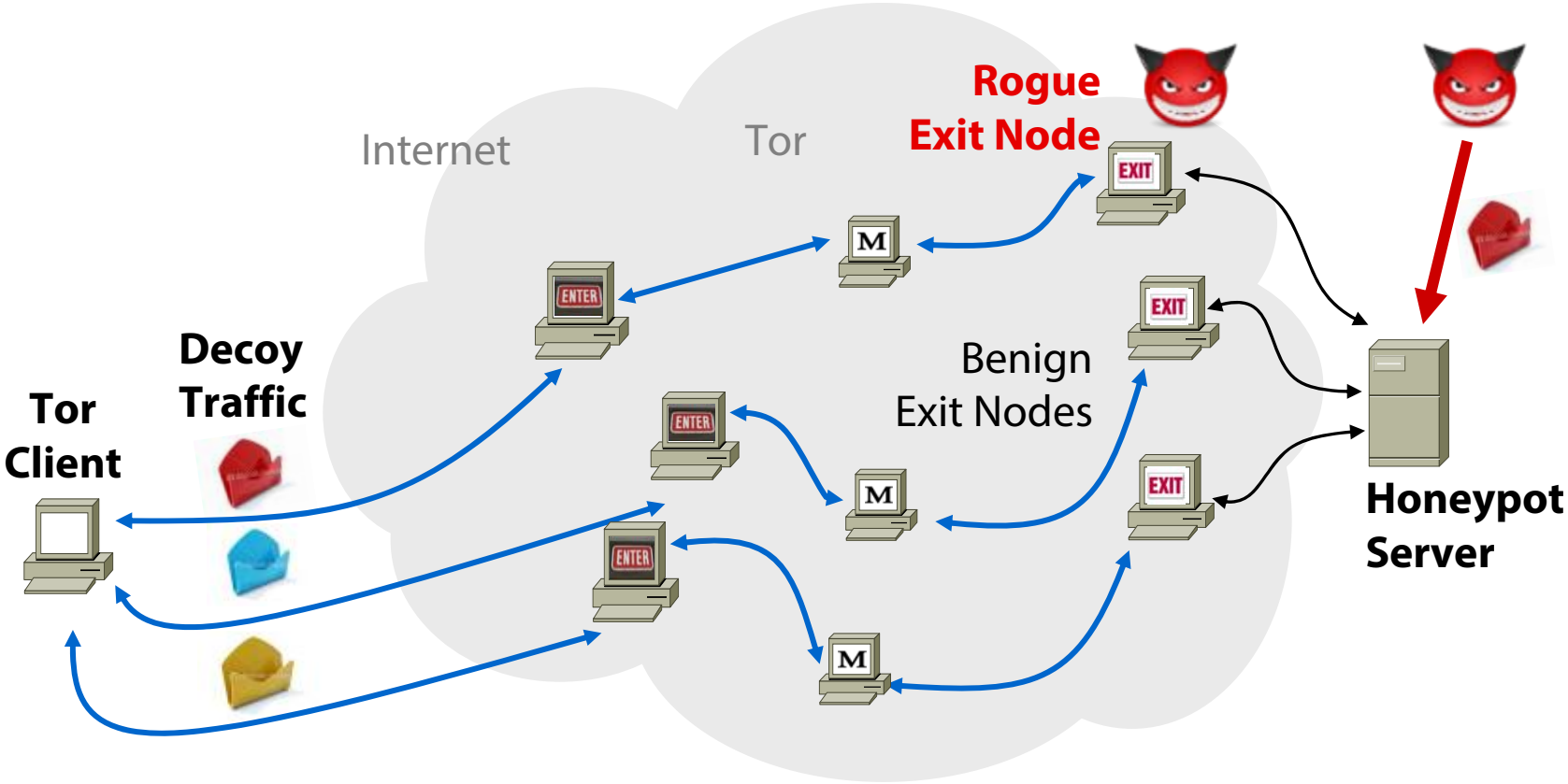
SECURITY 2/24/2015 @ 7:18AM | 13,489 views

How Hackers Abused Tor To Rob Blockchain, Steal Bitcoin, Target Private Email And Get Away With It

Across October and November of last year, some unlucky users of the world's most popular Bitcoin wallet, Blockchain.info, and one of the better-known exchanges, LocalBitcoins, had their usernames and passwords silently pilfered. They were robbed of significant sums, probably tens of thousands of dollars worth of the virtual currency, possibly more. Security-focused email services, Riseup and Safe-mail were also targeted by the same crew. And according to the man who witnessed the attacks go off last year, Digital Assurance director Greg Jones, it looks like buyers and sellers of dark markets were the targets.

The attackers used a tried-and-tested method to begin with, setting up a number of malicious exit relays on Tor. Legitimate exit relays act as the final jump from the anonymising Tor network, which loops users through a number of randomly-chosen servers across the world to protect their identity, onto the clear web. But any nefarious type who runs a malicious relay can use an encryption removal technique known as SSL stripping, where connections are

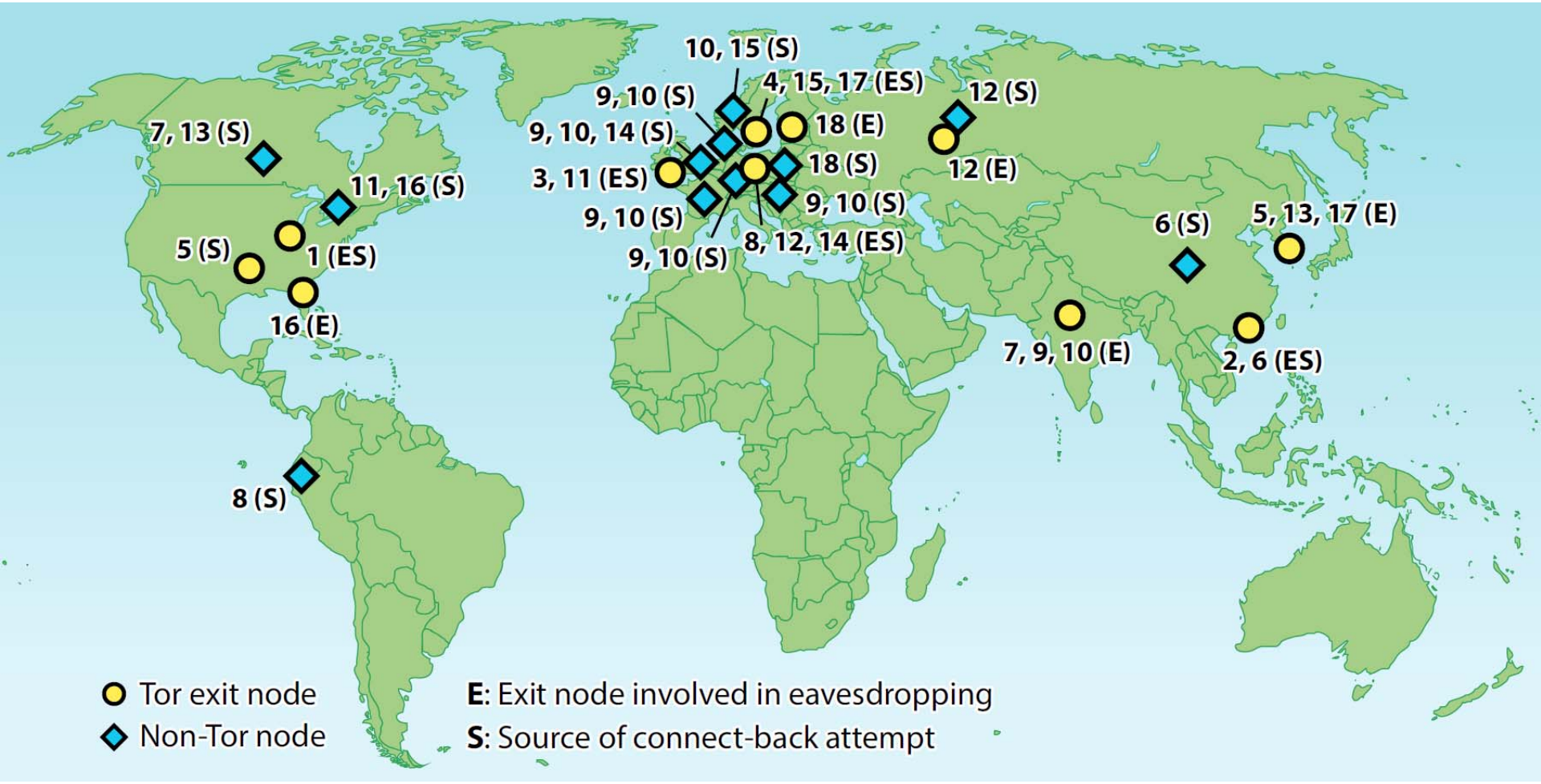
Detecting Traffic Snooping in Tor using Decoys



Expose unique decoy username+password through each exit node

Wait for unsolicited connections to the honeypot server presenting any of the exposed bait credentials

Detected Rogue Exit Nodes



30-month period: detected **18 cases** of traffic eavesdropping that involved **14 different Tor exit nodes**

What can we do?

Technical solutions exist

- Encryption
- Self-hosted services
- Anonymous communication
- …

But they are not enough

- Privacy vs. usability tradeoff
- Wrong assumptions
- Implementation flaws

Many users are not even aware of privacy issues, let alone solutions

Protect the right of individuals to control what information related to them may be collected

- *With technical means, not promises…*