Secure Systems Engineering: Introductory Comments

Bryan Ford
Ecole Polytechnique Federale de Lausanne (EPFL)

http://bford.info/

FOSAD – Bertinoro, Italy – August 29, 2016

About Me

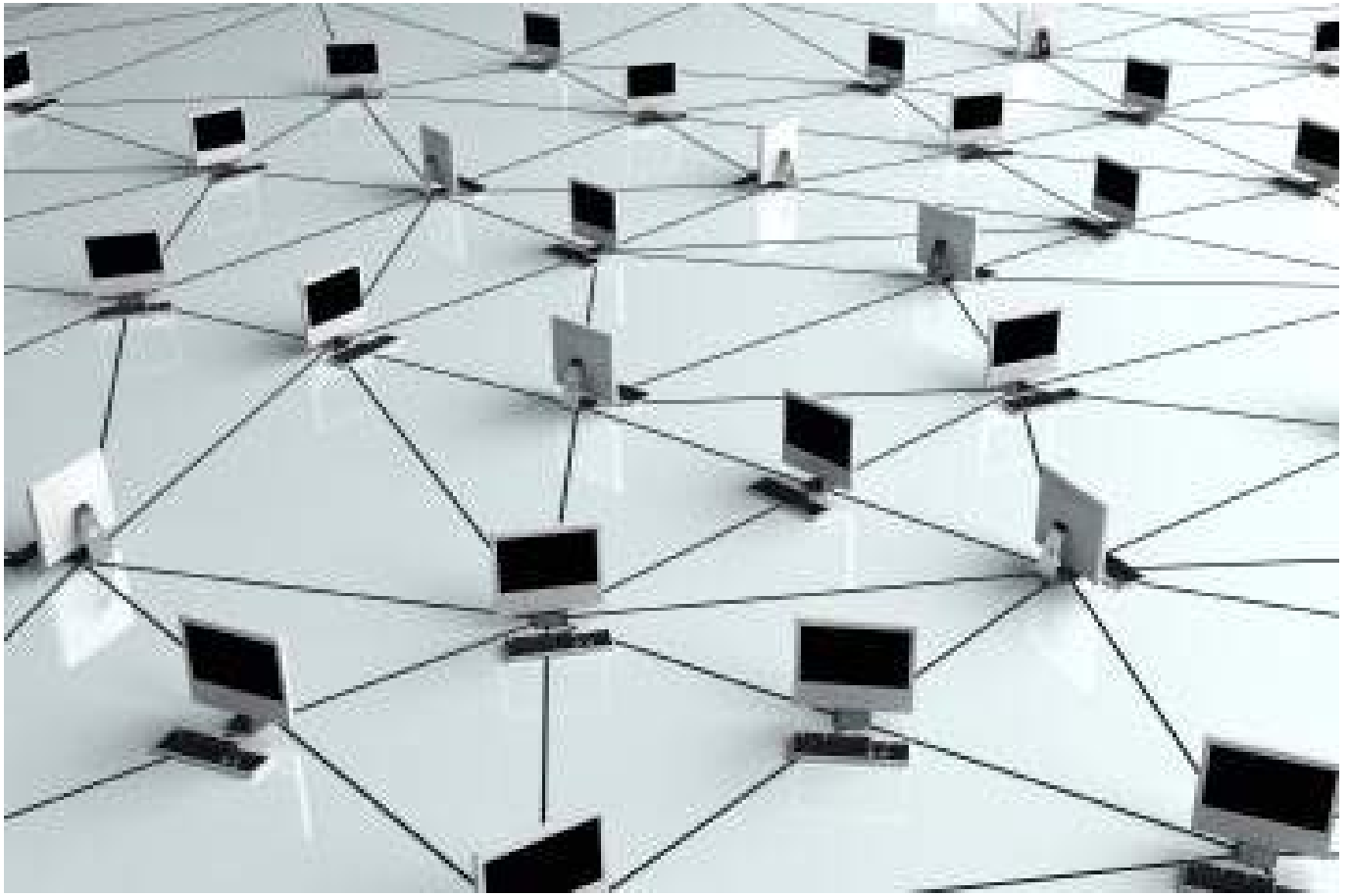
- Originally from California, but lived at times in Burkina Faso, Rwanda, Utah, Germany, …
- Started research during BS, University of Utah
- PhD at MIT CSAIL – Kaashoek, PDOS group
- Postdoc at MPI-SWS, Saarbrücken, Germany
- Prof at Yale University (New Haven) for 6 years
  - Old group website: http://dedis.cs.yale.edu
- Prof at EPFL (Lausanne) for 1 year
  - New lab website: coming (refer to old for now)

Which “Systems” do we care about?

Computer systems?

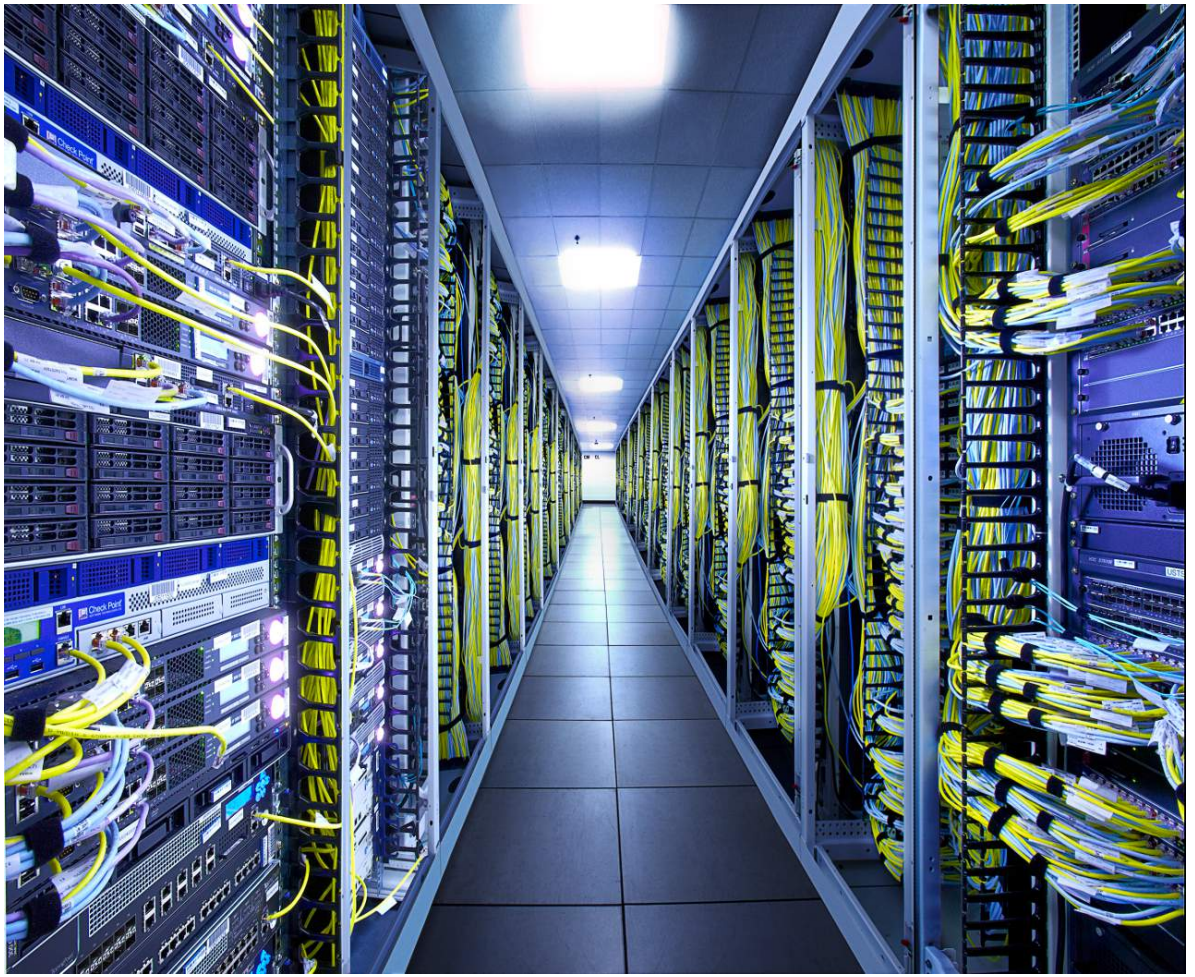
Which “Systems” do we care about?

Distributed systems?



Which “Systems” do we care about?

Cloud or “big data” systems?



Which “Systems” do we care about?

Or *human systems* – people, groups, institutions that depend on computation and communication?

Don’t forget the people!

Avoid the (natural) geek tendency
and think *only* about pure technical challenges

Goals of Systems-Security Work

Push the limits of what's known to be buildable

- More secure, faster, scalable, efficient, …
  - Often inevitably a bit incremental, unfortunately
- Define new approaches to old problems
  - Different algorithmic foundations, abstractions
- Explore new functionality, security properties
  - What do users need, but don't know it yet?

Connections, opportunities w/ formal methods

- We want to build "provably secure" systems
- Systems suggest interesting formal problems

System-building Considerations

The system builder must consider and balance:

- Simplicity, comprehensibility, maintainability
- Functionality, often including “feature creep”
- Security, privacy against realistic adversaries
- Performance, both average- and worst-case
  - Denial-of-service = adversary-induced worst-case
- Scalability, often to unanticipated demands
- Backward compatibility with legacy code/data
- Formal analyzability: can anything be proven?

Choosing Topics

You need:

- Strong **interest** in a practical goal or problem
- An **idea** about how it could be done [better]
- A realistic design and implementation **plan**
- A lot of **time** and **dedication**
- Optional but ideal: follow-through **deployment**
  - Is the product mainly the paper,
    or to produce something people can use?

Why [do I] build secure systems?

For the **learning** or **accomplishment**?

To make stuff **go faster**?

In hopes of creating **something useful**?

In order to **empower people**?

Can technology empower people **to be themselves?**

"On the Internet, nobody knows you're a dog."

Can technology empower people **in free expression and discourse?**

Queers Anonymous: Lesbians, Gay Men, Free Speech, and Cyberspace

Edward Stein[*]

Anonymity is a shield from the tyranny of the majority. It thus exemplifies the purpose behind the Bill of Rights and of the First Amendment in particular: to protect unpopular individuals from retaliation—and their ideas from suppression—at the hand of an intolerant society.[1]

Pseudonymity allows people who are experimenting with different sorts of interests to do so without social repercussions. People can temporarily obscure their real life and play with a different conception of what their life might be.[2]

Can technology empower people **in political self-organization?**

Can technology empower people **in economic self-determination?**

And does technology empower

“The Little Guy™”?

or

the loudest **trolls**,
the best-funded **criminals** and **astroturfers**,
the most sophisticated **surveillance agencies**?

Systems Security vs Real People

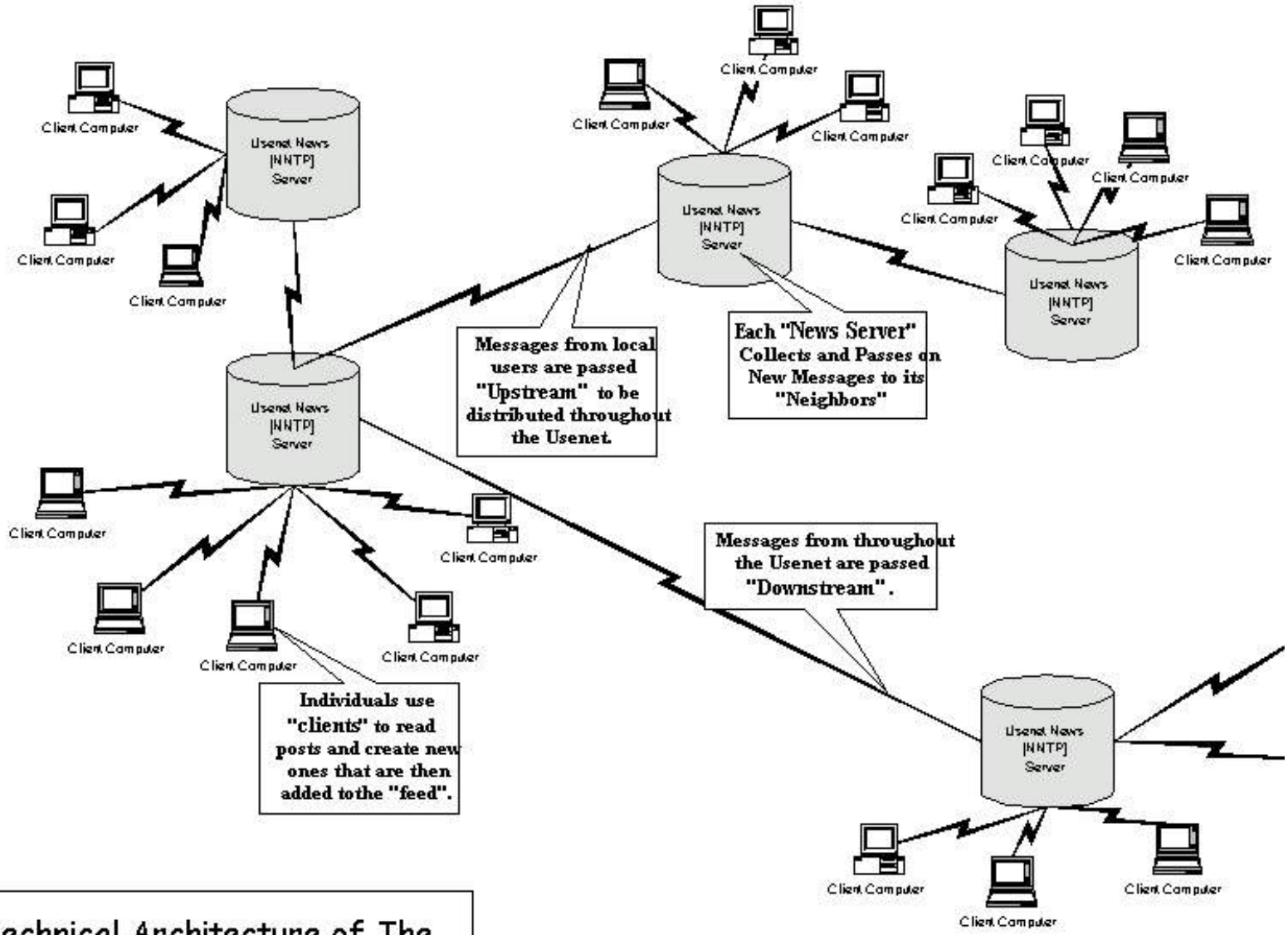
You can’t think about security of a system without thinking about:

- How legitimate users need to use the system
- How attackers might realistically misuse it

Unusable security/privacy features won’t get used, users instead choose usable insecure systems

- Usability is critical, but *really* hard
- “Nothing is foolproof because fools are so ingenious” - corollary to Murphy’s Law

Example: USENET



Technical Architecture of The Usenet Network

I Remember USENET

by Brad Templeton
12/21/2001

The word came to me from several
archive of USENET postings to go

USENET is the world's largest onlir
were discussed. The place where p

Google's new USENET archive bro
Web and 15 years before the dot-c
yet to come.

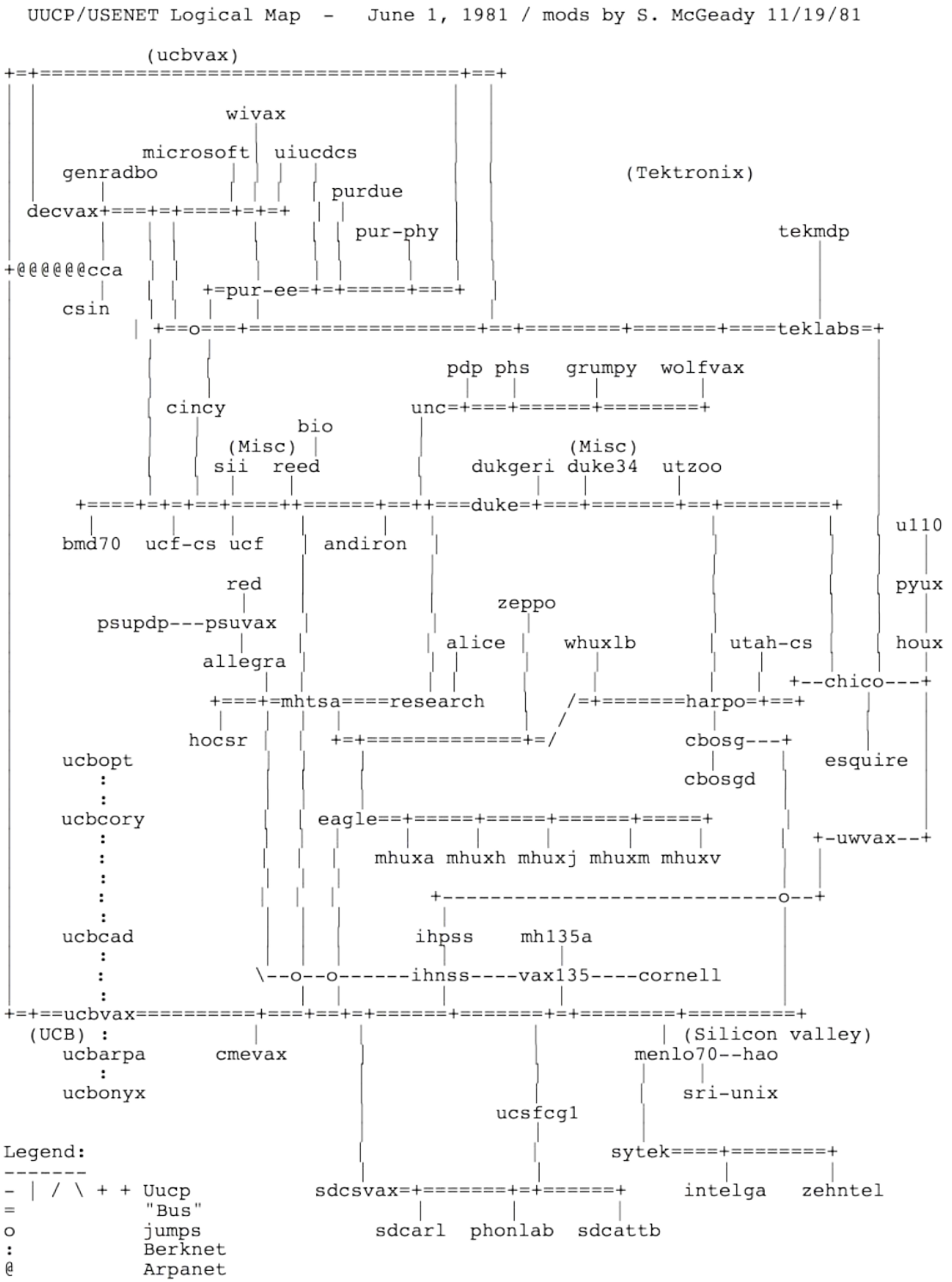
And So It Begins

USENET was created in 1979 whe
a general tool that could support co
run the same software and exchang
would read by running a local progr

A decade before there was a Web and 15 years before the dot-com bubble, we dreamed of the "WorldNet" yet to come.

UUCP/ USENET Logical Map

June 1, 1981



USENET's "Too-Strong" Security

Gossip: designed to "route around" faults or censorship – of *any* kind

- Redundant across paths: if one path censors, you'll get censored message via another

The problem: *spam* was just as uncensorable as content that people actually wanted

- Cheap for spammer, real costs distributed → classic tragedy of the commons

First USENET Spam

sci.stat.edu ›

Global Alert For All: Jesus is Coming Soon

7 posts by 7 authors G+1

Clarence L. Thomas IV 1/19/94

The earthquake in Los Angeles, California, the flood in Europe, the seemingly unstoppable war in the former Yugoslavia, the devastating fires in Australia, the flood in the Midwest of the United States of America, the devastating fires near Los Angeles, California, the rapid and appalling increase in violence in cities, towns, villages all over the world, the famines, the diseases, the rapid
decline of the family unit, and the destructive earthquake in India (in 1993) are signs that this world's history is coming to a climax. The human race has trampled on God's Constitution, as given in Exodus 20:1-17 (King James Version Bible), and Jesus is coming to set things right. These rapidly accelerating signs are an indication that Jesus is coming soon (Matthew 24).

First Commercial USENET Spam

alt.pub.coffeehouse.amethyst ›

Green Card Lottery- Final One?

9 posts by 8 authors G+1

Laurence Canter 4/12/94

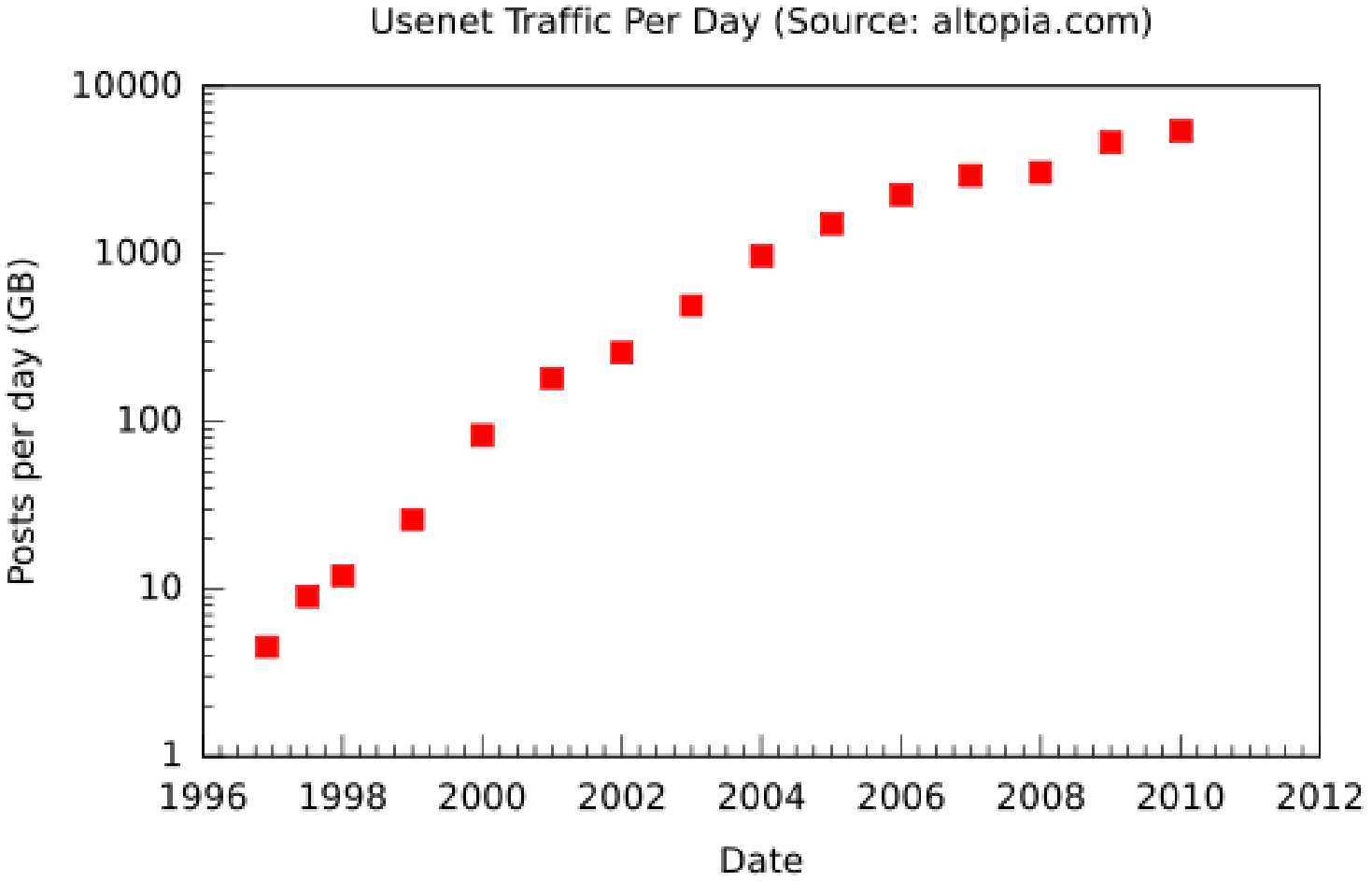
Green Card Lottery 1994 May Be The Last One!
THE DEADLINE HAS BEEN ANNOUNCED.

The Green Card Lottery is a completely legal program giving away a certain annual allotment of Green Cards to persons born in certain countries. The lottery program was scheduled to continue on a permanent basis. However, recently, Senator Alan J Simpson introduced a bill into the U. S. Congress which could end any future lotteries. THE 1994 LOTTERY IS SCHEDULED TO TAKE PLACE SOON, BUT IT MAY BE THE VERY LAST ONE.

PERSONS BORN IN MOST COUNTRIES QUALIFY, MANY FOR FIRST TIME.

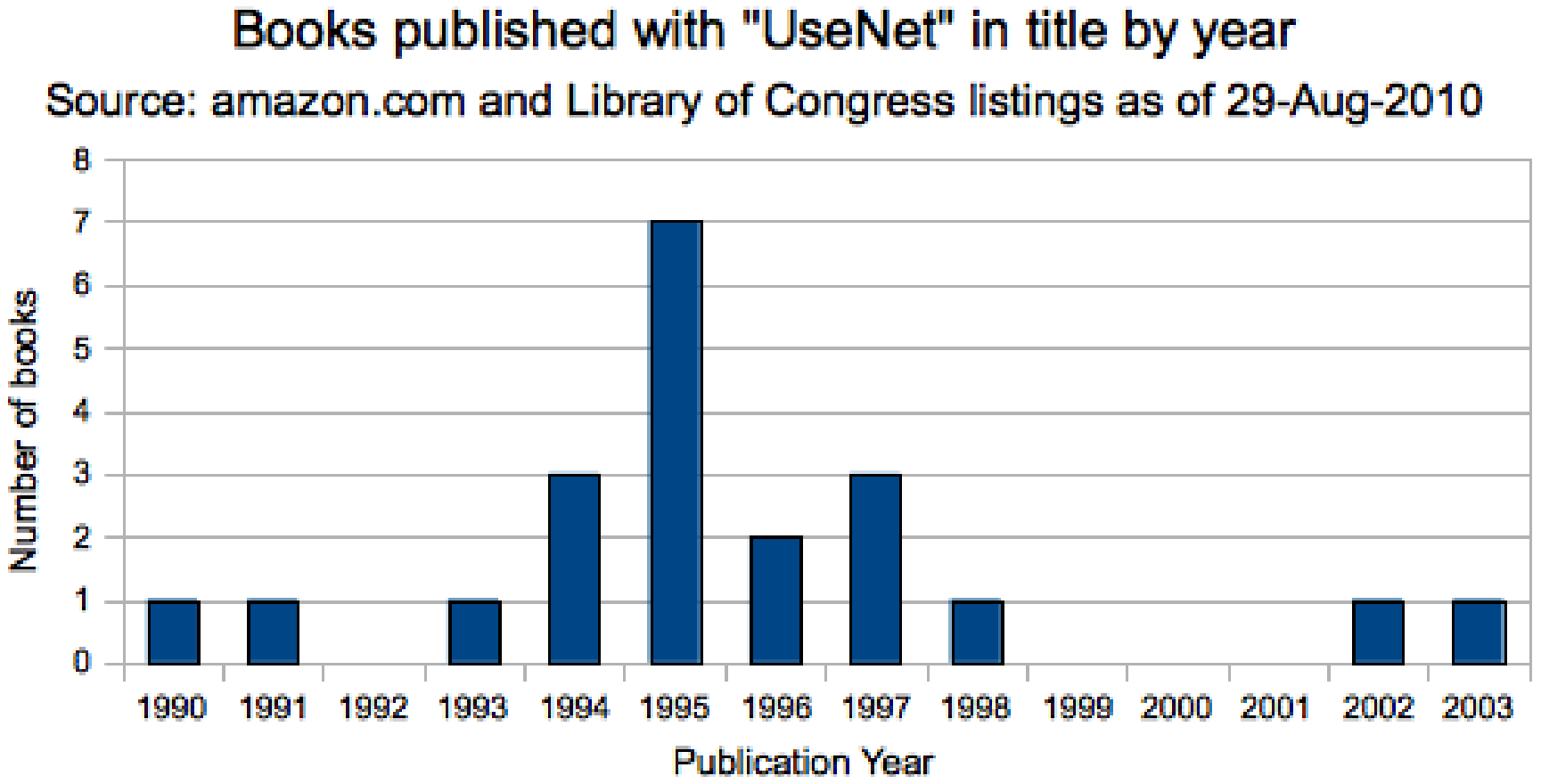
USENET Traffic: “Still Going Strong”

But comprised mostly of spam and “binaries”



USENET's Effective Heat Death

When spam took over, most *real* users fled...



Key Recurring Lessons

Once you make a system
"provably secure" in one way,
real users will find a way
to turn that security into insecurity

Building secure systems
is about constantly rethinking
what security actually means
in terms of what users need

A Brief History of Secure/Private Communication Tools

- USENET, E-mail: minimal security, no crypto
- SSH, SSL/TLS: end-to-end encryption
- Anonymous relaying: e.g., Mixminion, Tor
- Unstructured P2P: Napster, Freenet, Gnutella
- Structured P2P: e.g., Distributed Hash Tables
  - Chord, Pastry, Kademlia, Whanau, …
- Reputation: EigenTrust, Credence, dSybil
- Provable anonymity: Herbivore, Dissent, …

Rest of Course: Lessons from Experimental Decentralized Systems

Samples of my group's work at Yale and EPFL

- **Dissent:** principled, provable anonymity?
- **Buddies:** uh oh, intersection attacks!
- **AnonRep:** can we avoid pseudonymity?
- **Cothorities:** scalable collective authorities
- **RandHound:** random coins, random groups
- **ByzCoin:** large-scale consensus, coinage

Hiding in a Panopticon: Lessons and Challenges in "Provable Anonymity"

Bryan Ford
working with David Isaac Wolinsky, Joan Feigenbaum, Henry Corrigan-Gibbs, Ewa Syta, John Maheswaran, Daniel Jackowitz, and Ramakrishna Gummadi **– Yale**

Vitaly Shmatikov, Amir Houmansadr, Chad Brubaker **– UT Austin**

Aaron Johnson **– US Naval Research Lab**

FOSAD – Bertinoro, Italy – August 29, 2016


“Nobody knows
you're a dog?”


Dogs of the World
Terrier Group
Hound Group
Sporting Group
Toy Group
Non-Sporting Group
Herding Group
Actually, they know
exactly what kind you are

Who your friends are...


A dog party!
What you're doing

dogazon

What you *and* your friends like to buy

Gift suggestion ...

Rawhide Bone Dog Treat Size: 24" by Pet Time

~~$18.29~~ **$16.73** Prime

Order in the next **27 hours** and get it by **Monday, Feb 24.**

Only 19 left in stock - order soon.

More Buying Choices

$5.65 new (19 offers)

★★★★½ (55)

Pet Supplies: See all 25,595 items

... based on Rover's Dogbook likes

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

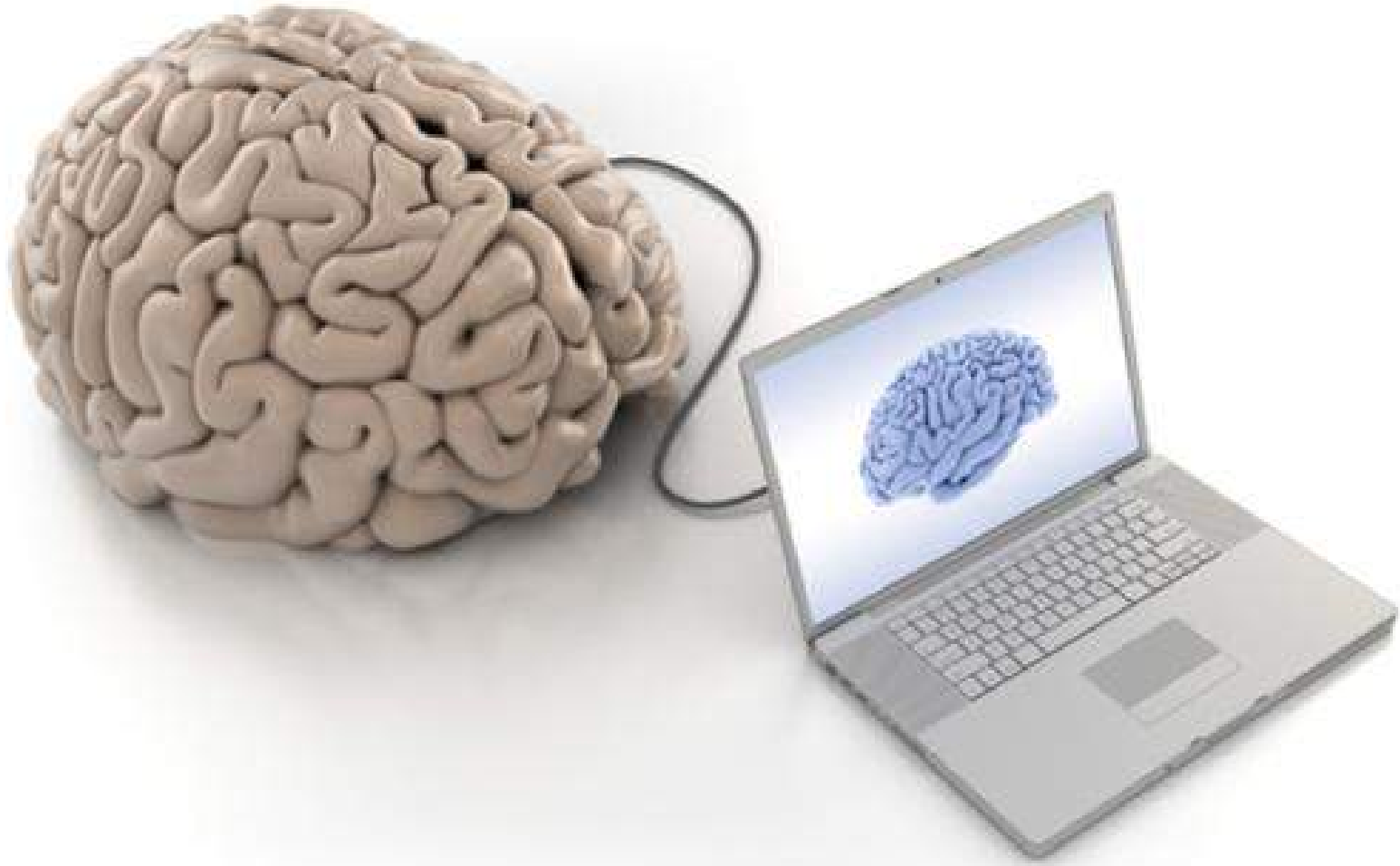
+ Comment Now + Follow Comments

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.

Why should I care about privacy if I have nothing to hide?

Reason 1: Freedom of Thought

- We invented computers to help us think.



Reason 1: Freedom of Thought

- We invented computers to help us think.
- Ubiquity brings dependence

Reason 1: Freedom of Thought

- We invented computers to help us think.
- Ubiquity brings dependence
- Whoever can read your private data can read your thoughts

Reason 2: Personal Security

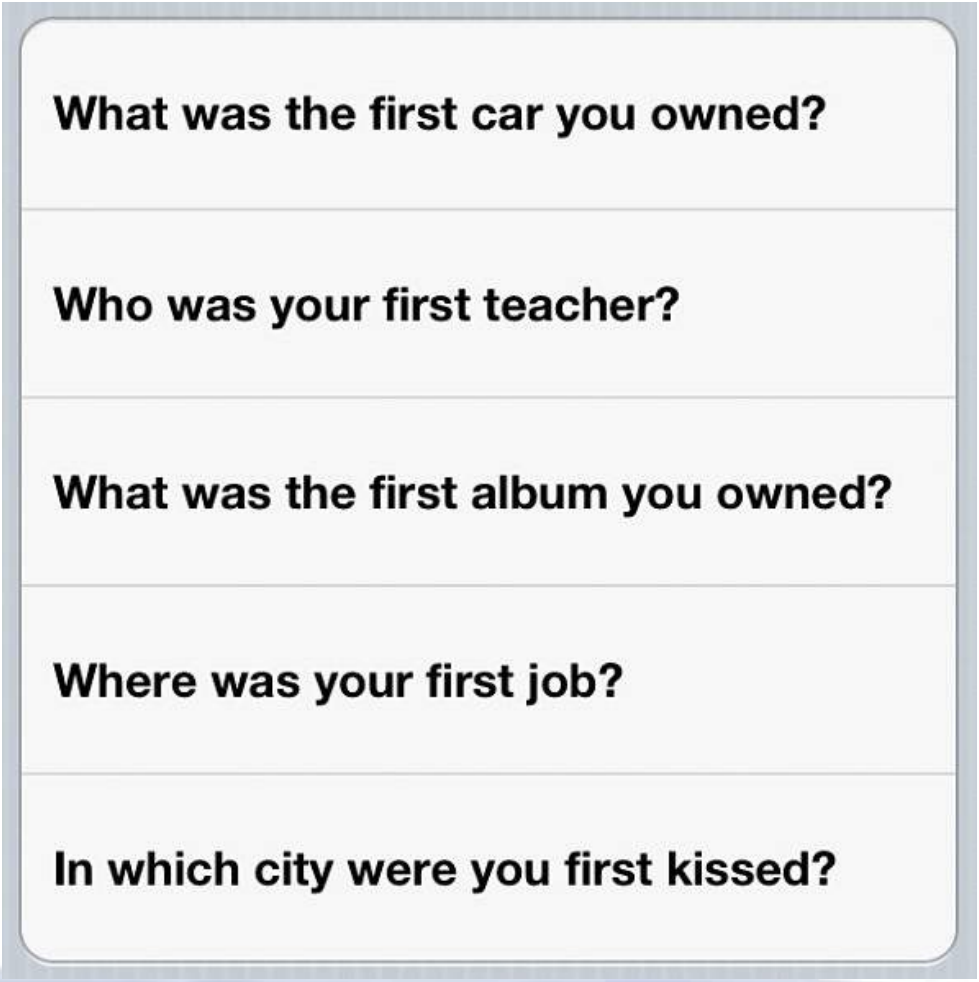
You think *this* is your password?

Reason 2: Personal Security

No, that's just a temporary access token.

This is your password.

Your life is your password.



Reason 2: Personal Security

Whoever can data-mine your life has your password

Who Wants to Track You Online?

- Advertisers (if you ever spend money)
- Vendors (if you ever buy things)
- Thieves (if you have any money)
- Stalkers (if you're a domestic abuse victim)
- Competitors (if you're a business)
- Extremists (if you're minority/gay/pro-choice...)
- The Police (if you're “of interest” w/in 3 hops)
- The Mob (if you're the police)

What tracking protection do we need?

Some people really need anonymity...

What tracking protection do we need?

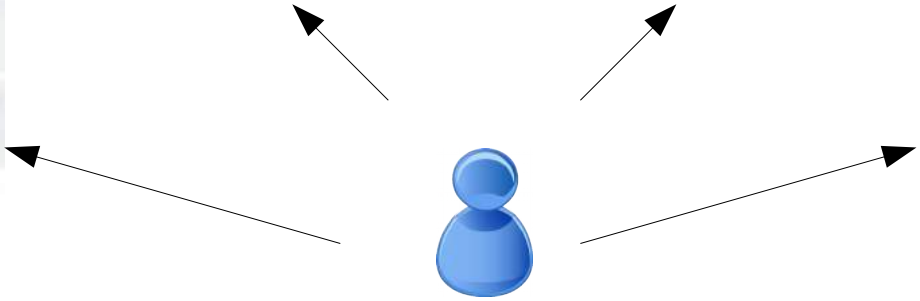
Many people just tend to wear multiple hats

Family Hat

Hobby Hat

Professional Hat

Party Hat



The Real You

Talk Outline

- ✔ Why Anonymity?
- **Current State of the Art**
- Grand Challenges in Anonymity
  - Global traffic analysis
  - Active interference attacks
  - Intersection attacks
  - De-anonymizing exploits
  - Accountability provisions
- Status and Ongoing Work

What protection can we get now?

Many weak defense options

- Disable cookies, browser history, Flash, Java
- “Do-Not-Track” HTTP option
- “Hide” behind NATs, firewalls, corporate VPNs
- Commercial proxy/VPN providers

Current state-of-the-art

- Onion routing systems – e.g., Tor

Do Not Track

Universal Web Tracking Opt Out

```
GET /something/here HTTP/1.1
Host: example.com
DNT: 1
```

Do Not Track

Universal Web Tracking Opt Out

Please don't track me,
pretty please???

Of course we'll
respect your privacy –
promise!

International New York Times

Business Day
Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS

SLIPSTREAM

Do No

By NATASHA

Published: Oc

THE cam

last mont

ZDNet

White Papers | Hot Topics | Downloads | Reviews | Newsletters

Topic: Privacy *Compare* *Follow via:*

Why

miser

Summary: *A*

the big data-co

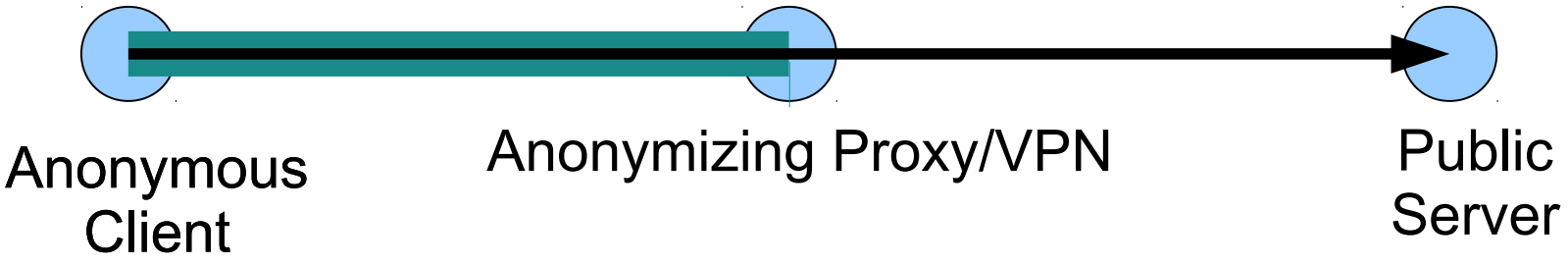
nothing at all.

By

Commercial VPN services

Popular for circumventing the Great Firewall

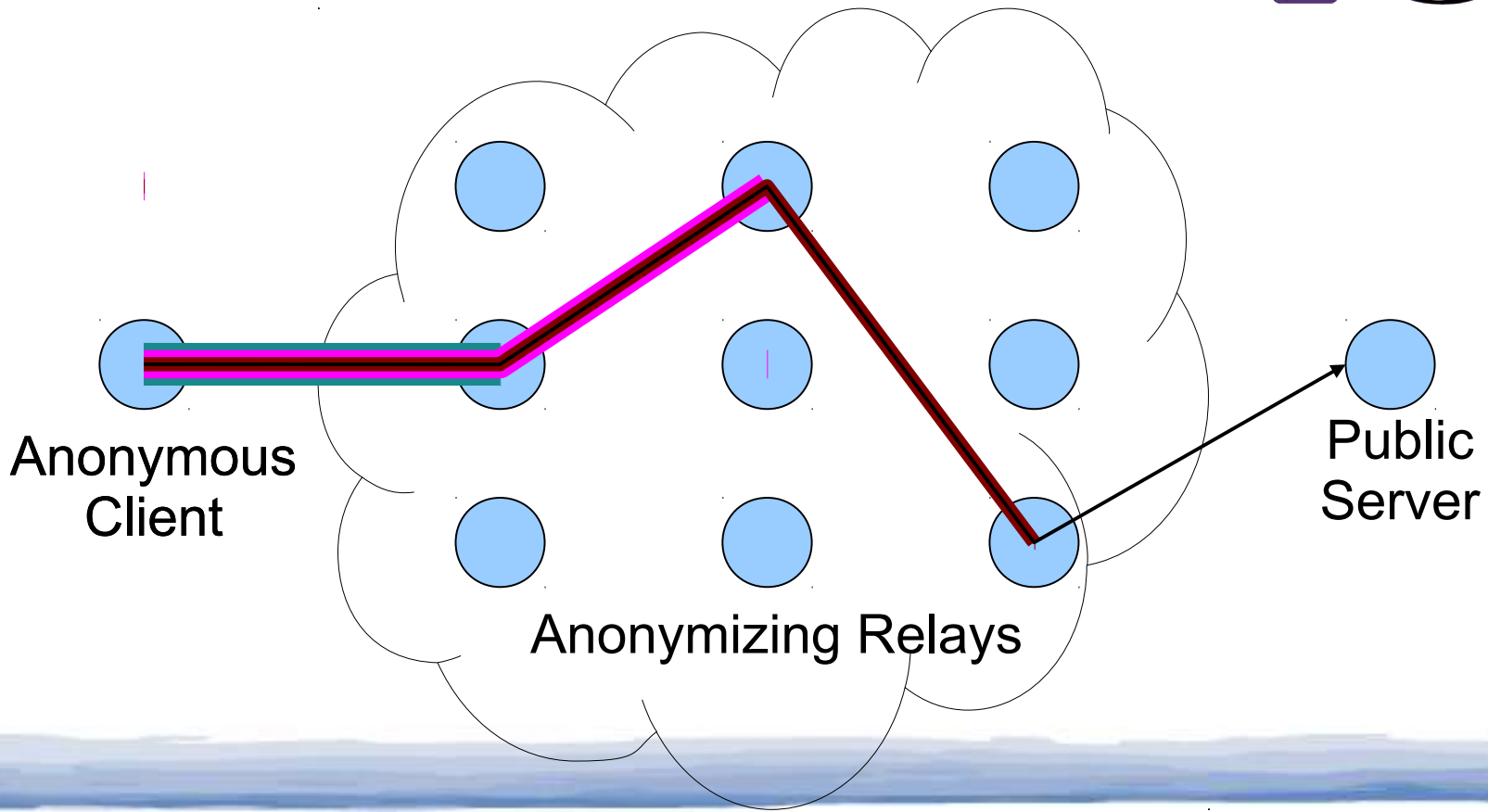
- You build encrypted tunnel with VPN server
- VPN server forwards traffic to destination
- Looks like it's coming from VPN server
- Hope the server operator protects your privacy



The current state-of-the-art

Onion routing tools such as **Tor**

- **https://www.torproject.org**



The Dissent Project

Clean-slate anonymous communications design

- Offer *quantifiable* and *measurable* anonymity
- Build on primitives offering *provable security*
- Don't just *patch* specific vulnerabilities, but *rearchitect* to address whole *attack classes*

http://dedis.cs.yale.edu/dissent/

[CCS'10, OSDI'12, CCS'13, USENIX Sec'13, ...]

Why rethink online anonymity?

NSA said Tor is the “King of Anonymity” – maybe onion routing is good enough?

Sampled Traffic

Internet-Exchange-Le

A Practical Congestion Attack on Tor Using Long Paths

Nathan S. Ev

Colorado Researc

for Security and

University of D

Email: nevans6

DSSS-Based Flow Marking Technique for Invisible Traceback *

Denial of Service or Denial of Security?

Traffic Correlati

htt

Aaron Johnson[1] Chris Wacek[2] Rob Ja

[1]U.S. Naval Research Laboratory, Washington

{aaron.m.johnson, rob.g.jansen, paul.syverson}@nrl

Low-Resource Routing Attacks Against Tor

Limits of Anonymity in Open Environments

STATISTICAL DISCLOSURE ATTACKS

Traffic Confirmation in Open Environments

Browser-Based Attacks on Tor

Timothy G

- E

vulnerable to **five**

  - Global traffic a
  - Active attack
  - Denial-of-se
  - Intersection
  - Software exploits

- Question is *when & h*

Attacking Tor: how the NSA targets users' online anonymity

Secret servers and a privileged position on the internet's backbone used to identify users and attack target computers

Bruce Schneier

theguardian.com, Friday 4 October 2013 10.50 EDT

Jump to comments (238)

Some De-anonymization Incidents

Tor is being broken – or *circumvented* – regularly

The Boston Globe

Harvard undergrad arrested in bomb hoax

By Eric Moskowitz | GLOBE STAFF DECEMBER 18, 2013

A Harvard student trying to get out of a final exam admitted to the FBI that he sent a bomb threat that forced the university to evacuate multiple buildings and rattled the campus, federal officials said Tuesday.

Inside the Tor exploit

Summary: *Some of the people who were most concerned about Internet privacy, and were using the Tor anonymous Internet service to protect it, may have been the most exposed.*

By Steven J. Vaughan-Nichols for Networking | August 5, 2013 -- 21:56 GMT (14:56 PDT)

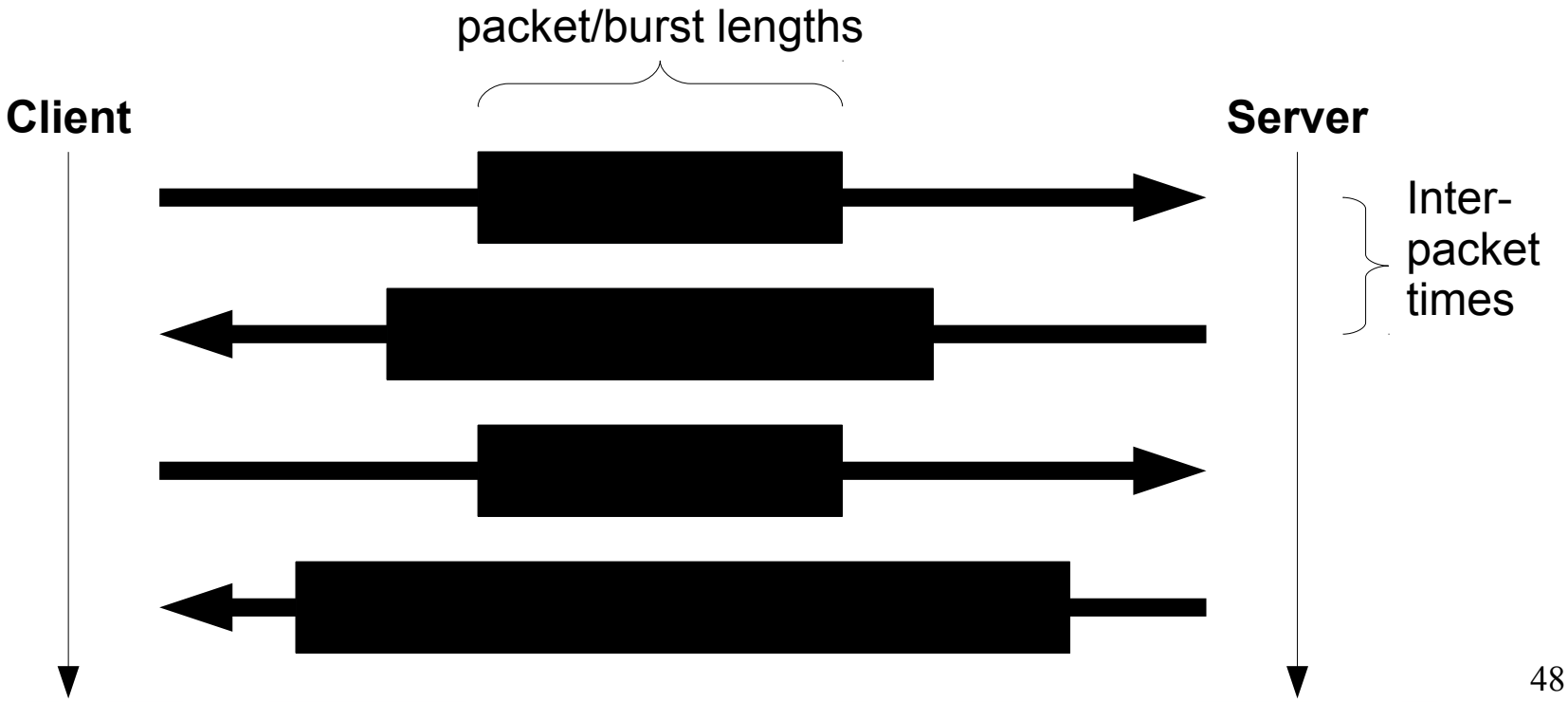
Follow @sjvn

Talk Outline

- ✔ Why Anonymity?
- ✔ Current State of the Art
- **Grand Challenges in Anonymity**
  - Global traffic analysis
  - Active interference attacks
  - Intersection attacks
  - De-anonymizing exploits
  - Accountability provisions
- Status and Ongoing Work

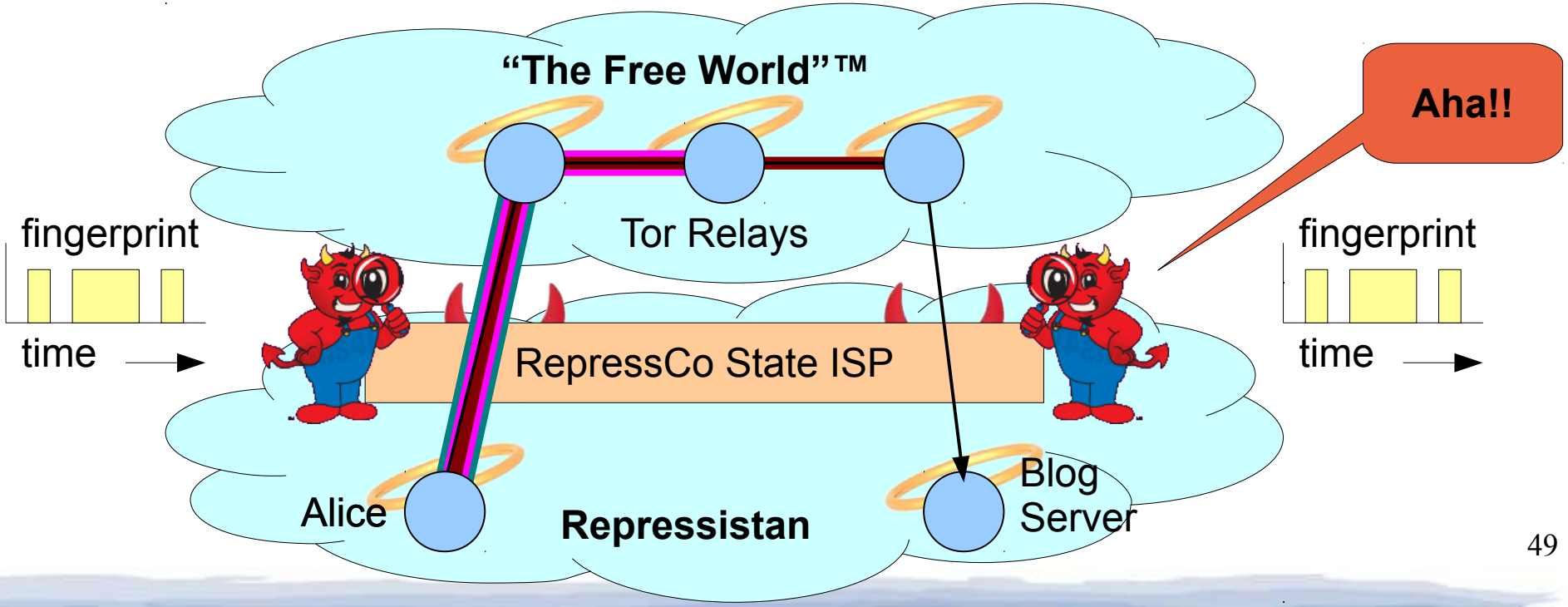
The Traffic Analysis Problem

- Most communication has a *traffic pattern*
  - Lengths and timings of packets in each direction
  - Pattern can be *fingerprinted* without seeing content



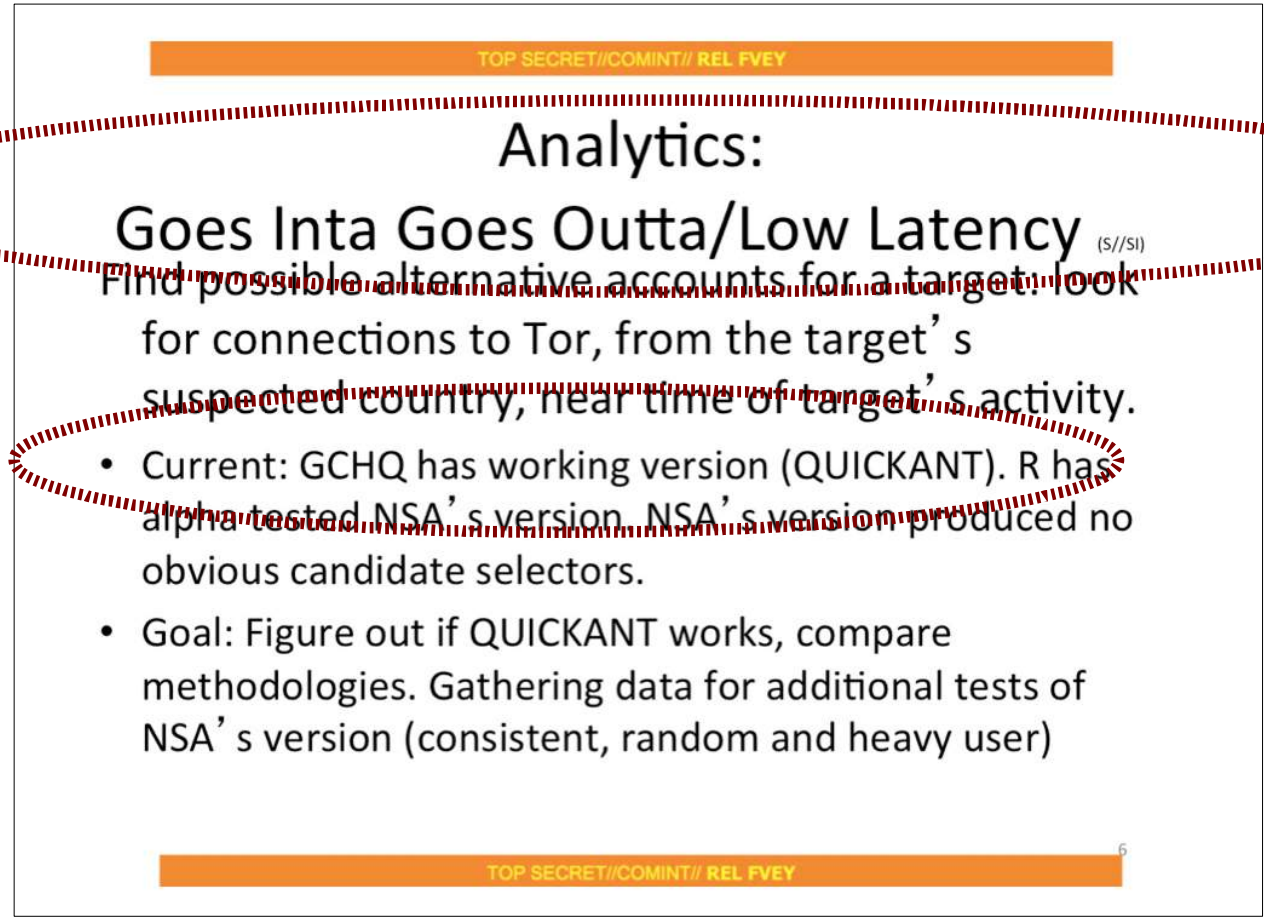
Tor Traffic Analysis Scenario

- Alice in Repressistan uses Tor to post on blog server hosted in Repressistan
- State ISP controls *both* entry and exit hops
- Fingerprint & correlate traffic to **deanonymize**



Do Attackers Actually *Do* This?

Not sure, but some are *working hard on it...*

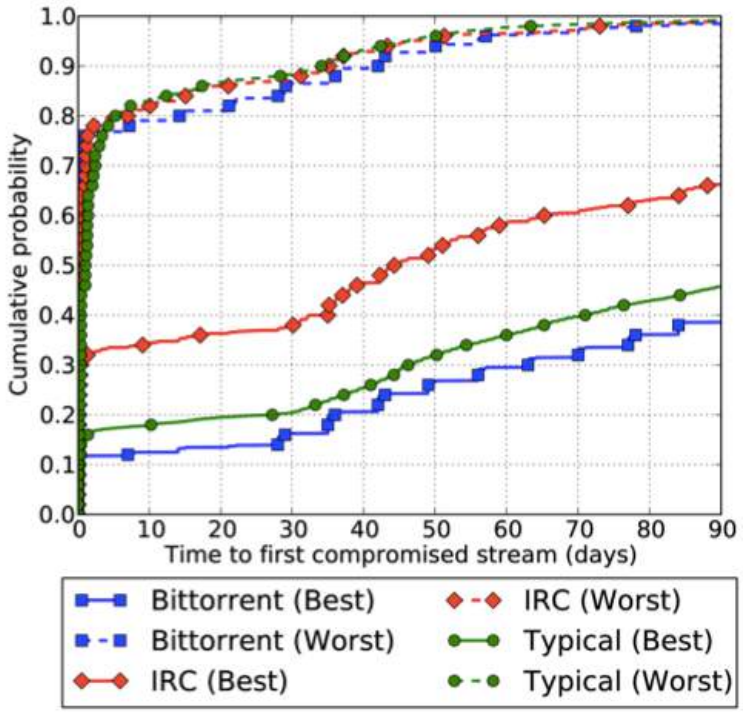


("Tor Stinks" slide deck, Guardian 10/4/2013)

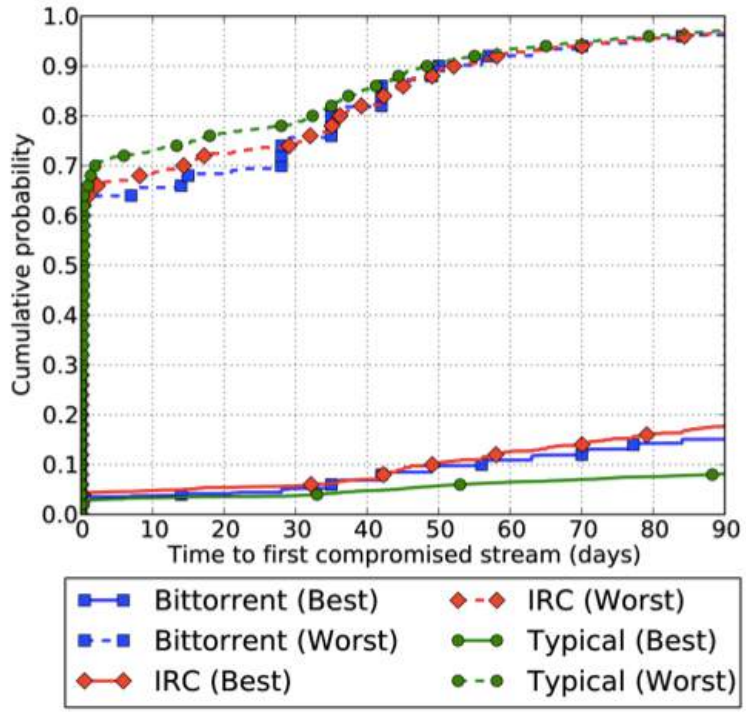
Can De-Anonymize "Real" Users?

Yes, if attacker can monitor an Internet AS or IXP

- "Users Get Routed", Johnson et al. CCS 13



(a) Time to first stream compromised by AS adversary.



(b) Time to first stream compromised by IXP adversary.

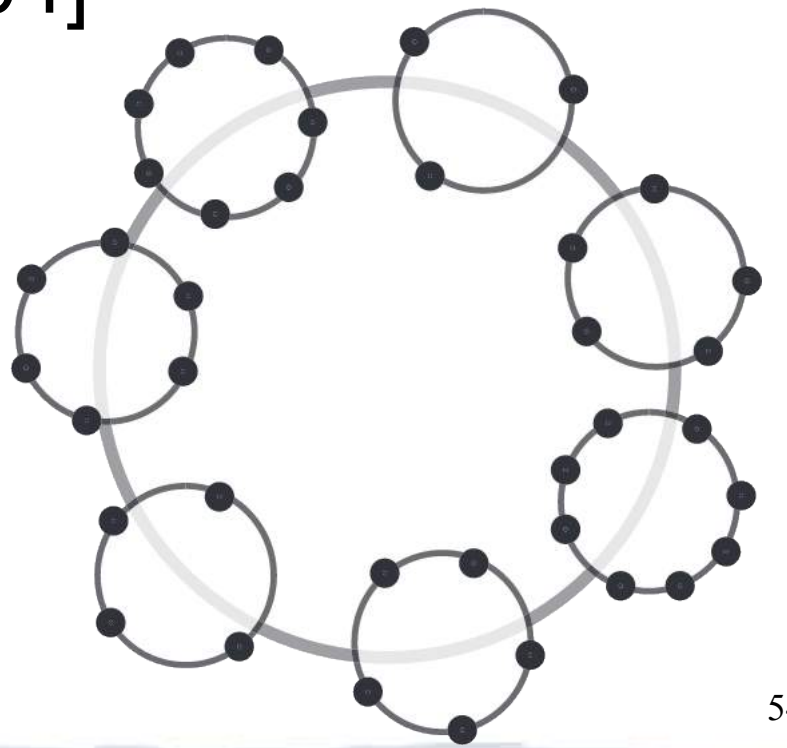
Precedent: Herbivore

First attempt to build practical anonymity system providing *provable, quantifiable* anonymity

- "Eluding Carnivores: File Sharing with Strong Anonymity" [Sirer et al, 2004]

Anonymity foundation

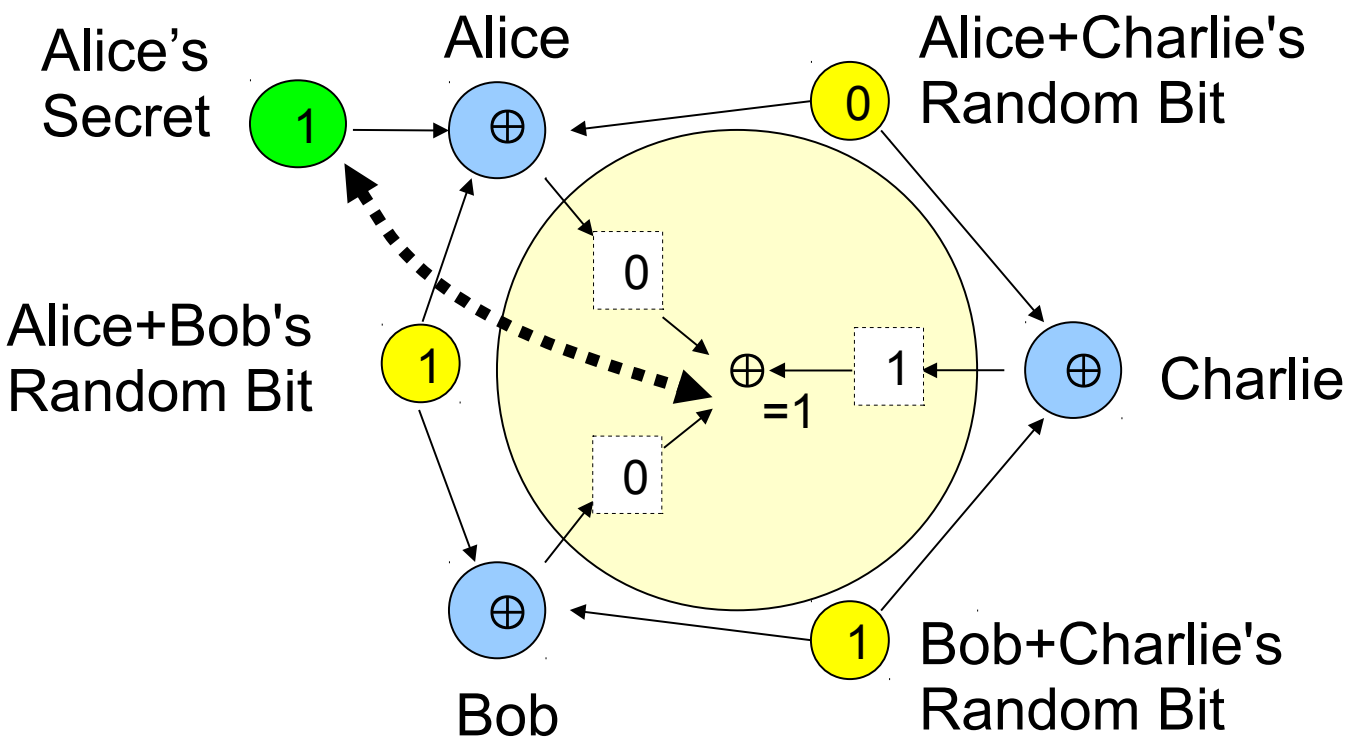
- Not *relaying* but *dining cryptographers* (DC-nets)
- Break network into small *k*-anonymity sets



Can We Resist Traffic Analysis?

Dining Cryptographers or DC-nets [Chaum '88]

- Key property: provable anonymity within a group



Key difference

Why is DC-nets resistant to traffic analysis?

- Nodes act *collectively* rather than *individually*
- Send same amounts of cryptographically-indistinguishable data in each round

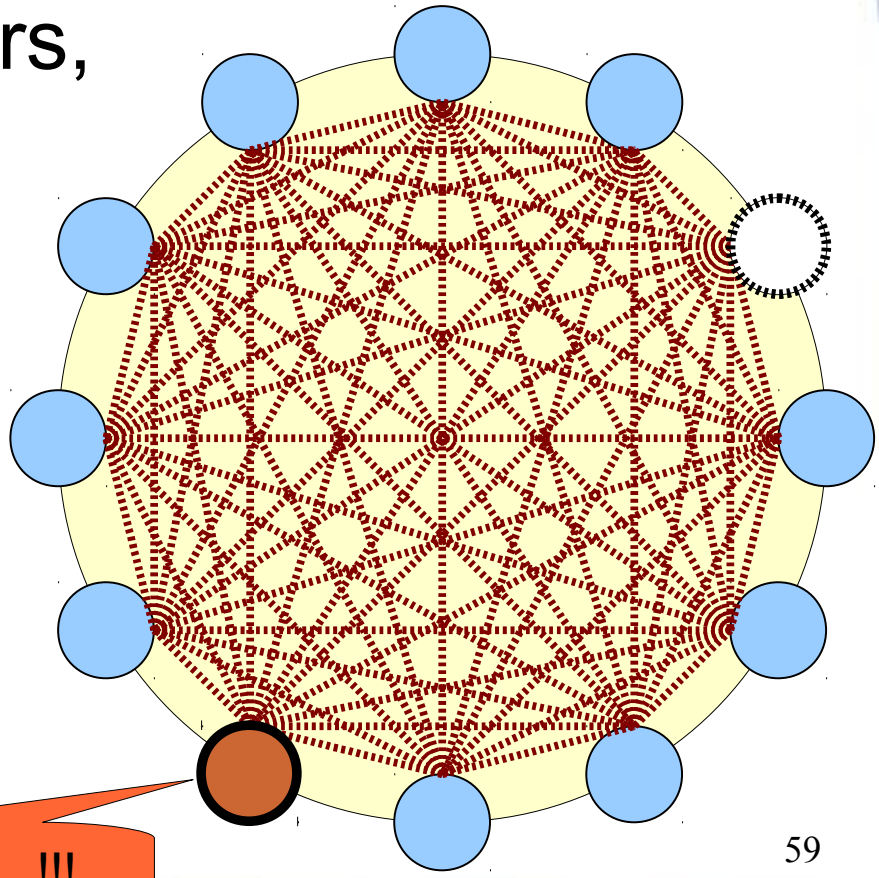
Versus onion routing: nodes makes *individual* decisions what path to take, and when to send

- Individual decisions yield traffic patterns that can be fingerprinted to identify individuals.

But: is DC-nets practical? Before, no.

Why DC-nets Doesn't Scale

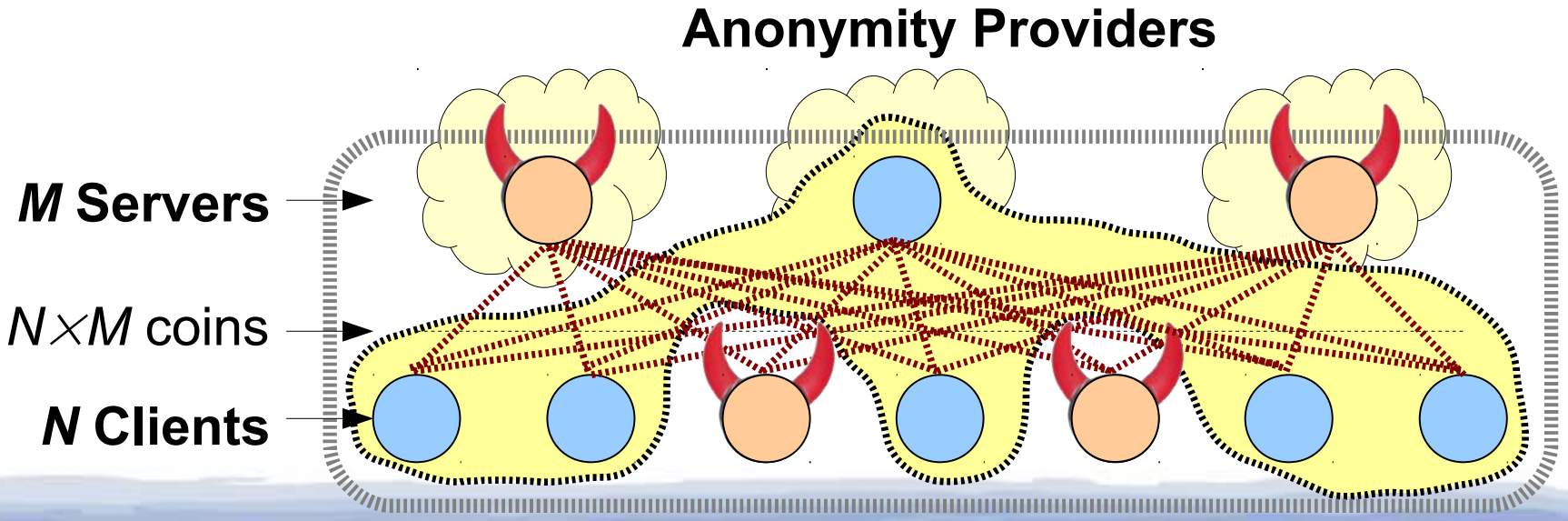
- **Computation cost:** $N \times N$ shared coin matrix
- **Network churn:**
  if *any* participant disappears,
  all nodes must start over
- **Disruption:**
  any single “bad apple”
  can jam communication



“Dissent in Numbers” [OSDI 12]

Scalable DC-nets using client/multi-server model

- Clients share coins *only* with servers
- As long as *at least one* honest server *exists*, yields ideal anonymity among *all honest clients*

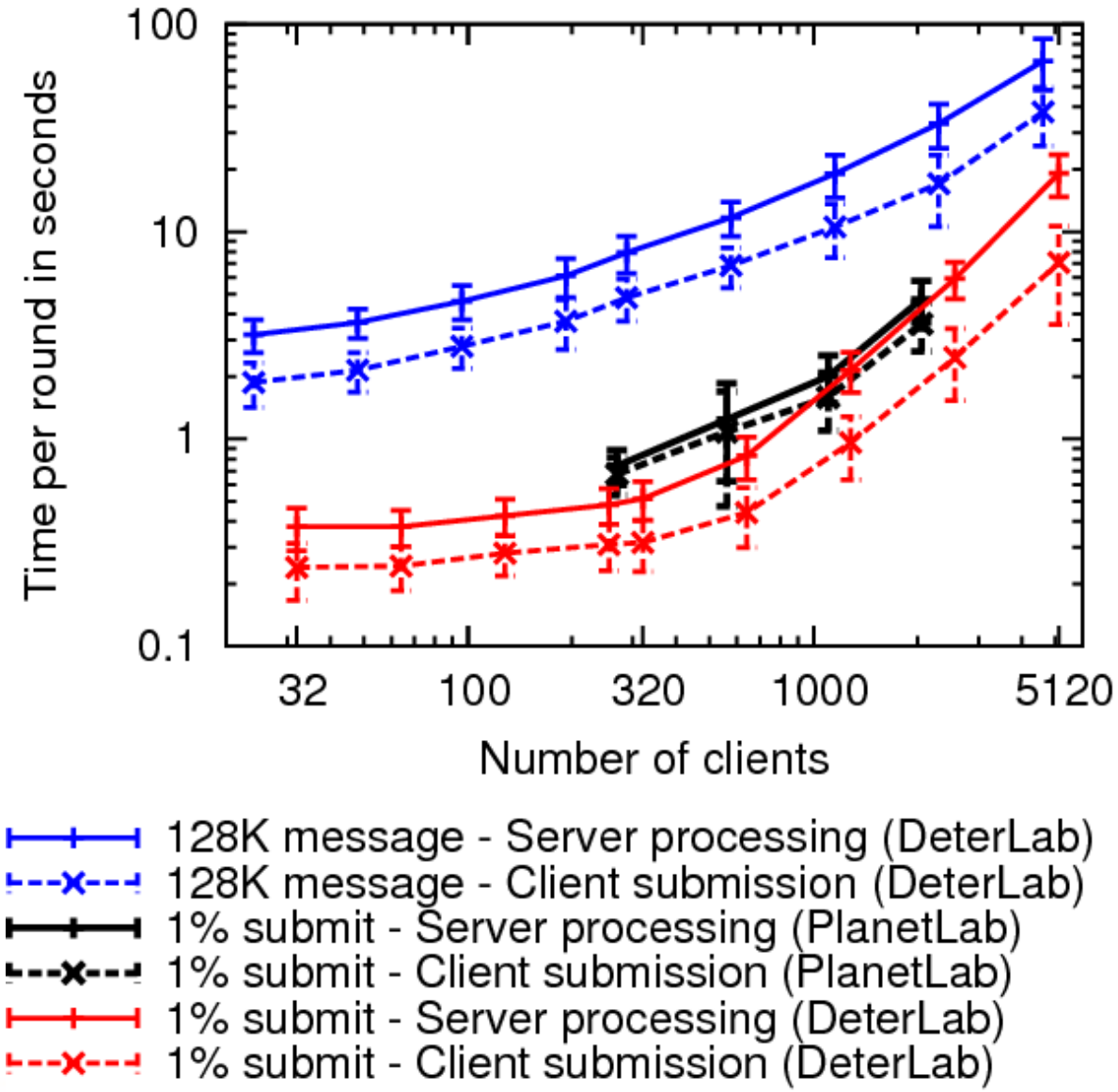


Scaling to Thousands of Clients

100× larger anonymity sets

- (Herbivore, Dissent v1: ~40 clients)

<1 sec latency w/ 1000 clients



“Provable” versus “Proven”

“Dissent 1.0” has a full, extensive, rigorous proof

- “Security Analysis of Accountable Anonymity in Dissent” [Syta et al, TISSEC 2014]
- But this was “first-cut” version that doesn’t scale

Dissent in Numbers, follow-ons:

- Have “security arguments”, *should* be provable
- But rigorous formalization, analysis still open

Talk Outline

- ✔ Why Anonymity?
- ✔ Current State of the Art
- **Grand Challenges in Anonymity**
  - ✔ Global traffic analysis
  - **Active interference attacks**
  - Intersection attacks
  - De-anonymizing exploits
  - Accountability provisions
- Status and Ongoing Work

Tor hides you in a tangle of wires...

...or a plate of spaghetti

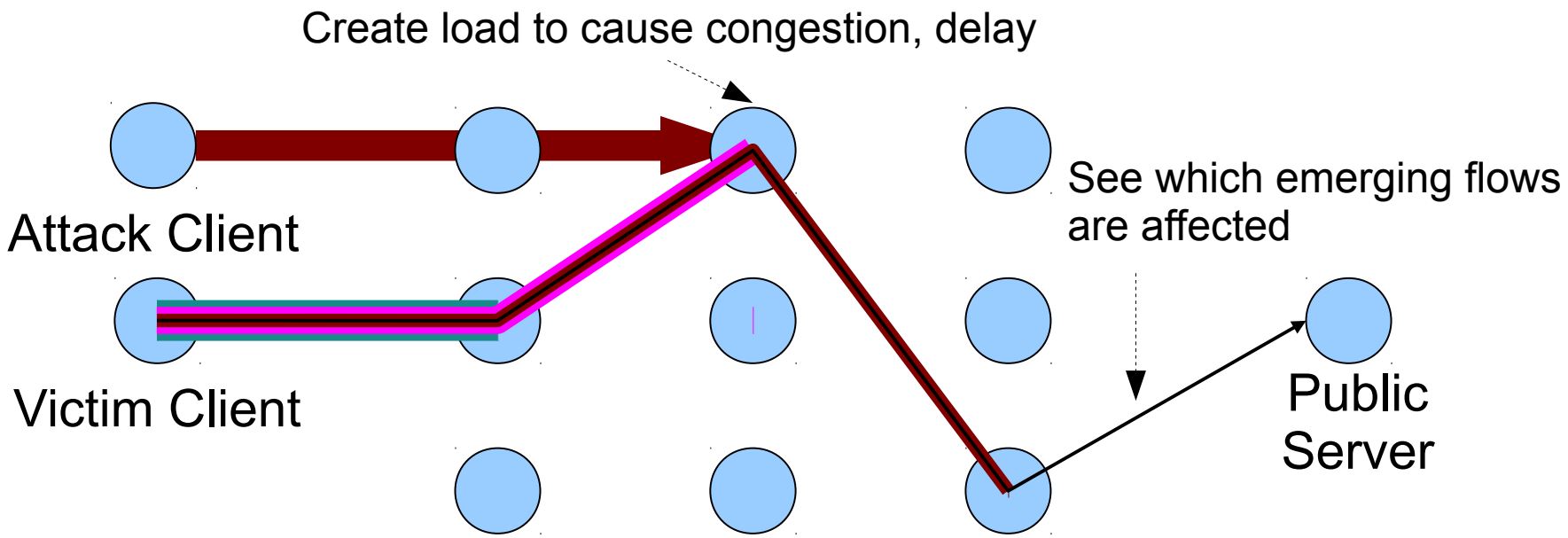
But tug on either end of a strand...

...and you'll find the other

Active Attacks: Tugging on Spaghetti

Attacker perturbs flow performance at either end traceable side-channel “markers”

- Congestion attacks: [Murdoch 05, Evans 09]
- Related “relay early” attack confirmed in July 14

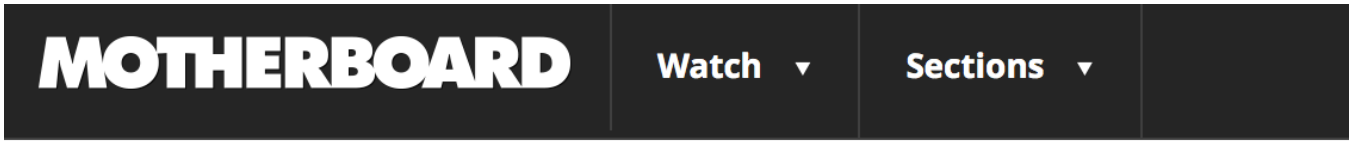


Active Attacks in the Wild

CMU researchers found, exploited side-channel

- Entry relay “marks” flow, colluding exit detects

FBI gets wind, BlackHat talk vanishes, …



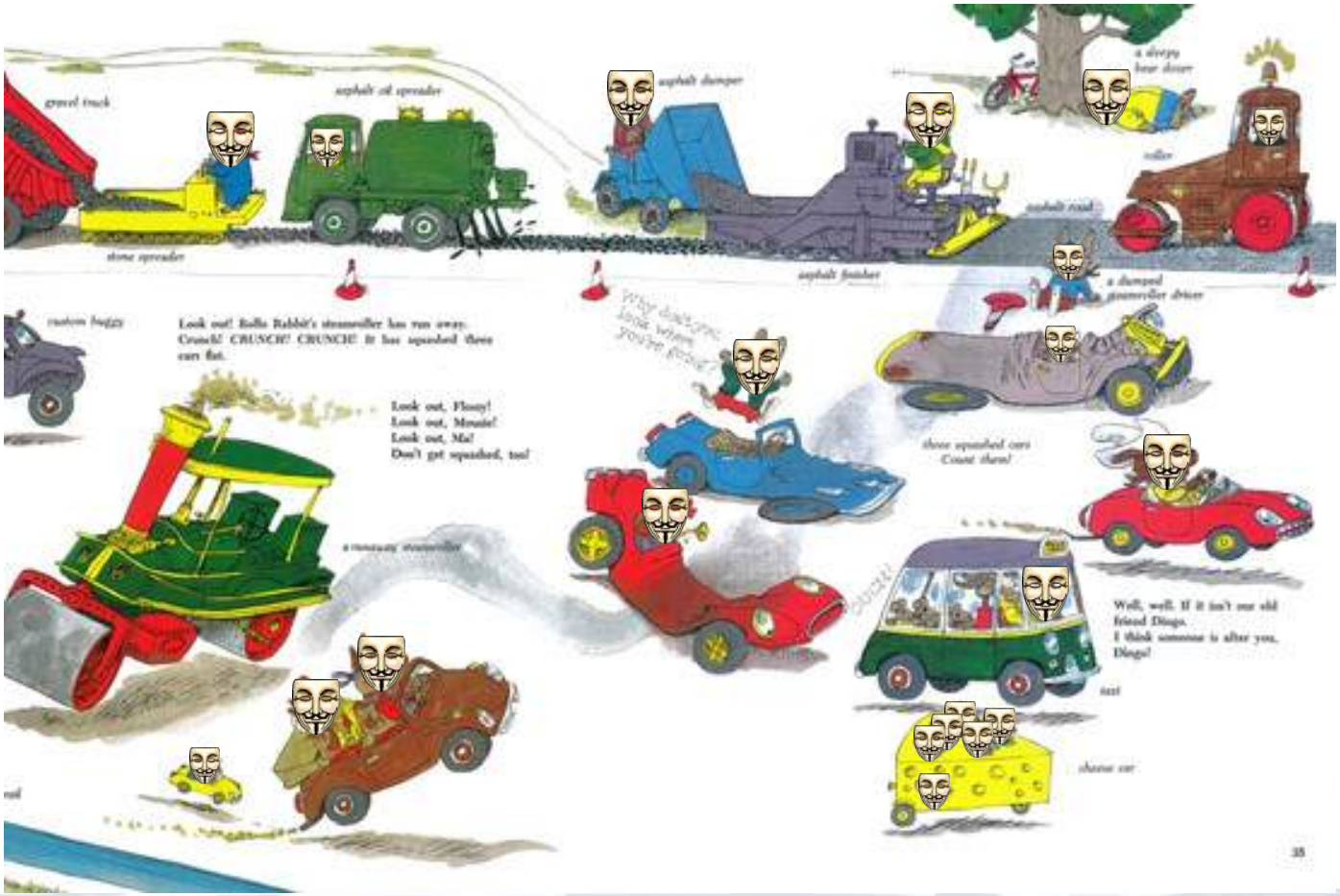
Confirmed: Carnegie Mellon University Attacked Tor, Was Subpoenaed By Feds

Written by **JOSEPH COX**

February 24, 2016 // 09:05 AM EST

Anonymity Depends On **Similarity**

People doing many **different** types of activities are fingerprintable, trackable by flow behavior…



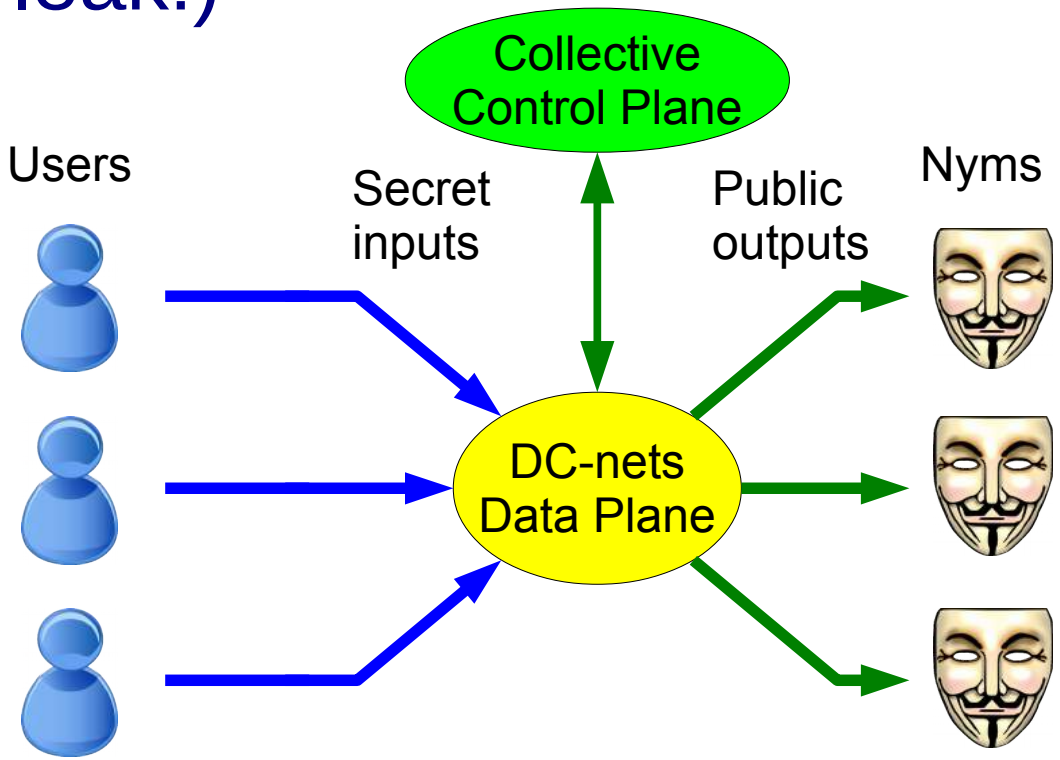
Anonymity Depends On **Similarity**

…but not so distinguishable when involved in a single **common, collective** activity

Collective Anonymity in Dissent via **Collective Control Plane (CCP)**

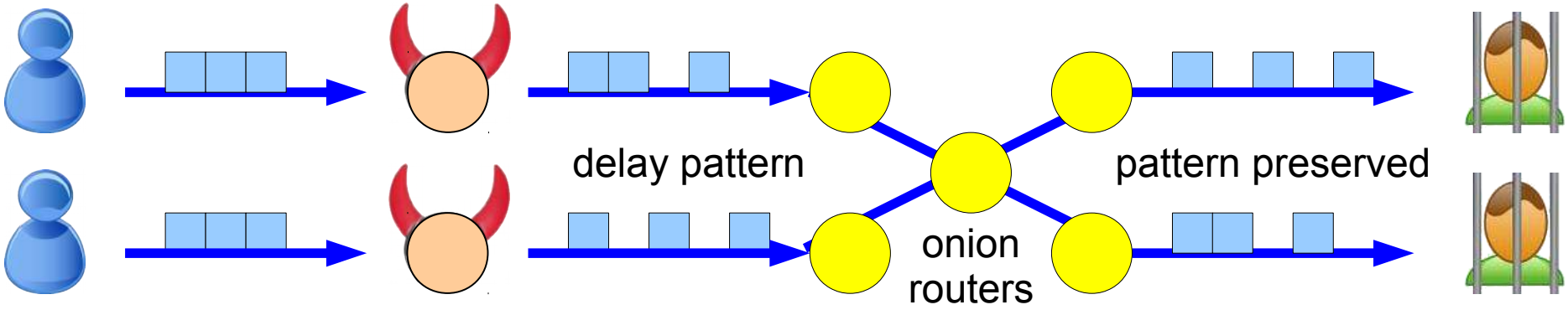
Policy Oracle controls when/how much to send

- But *does not know* who owns which nyms (can't leak!)

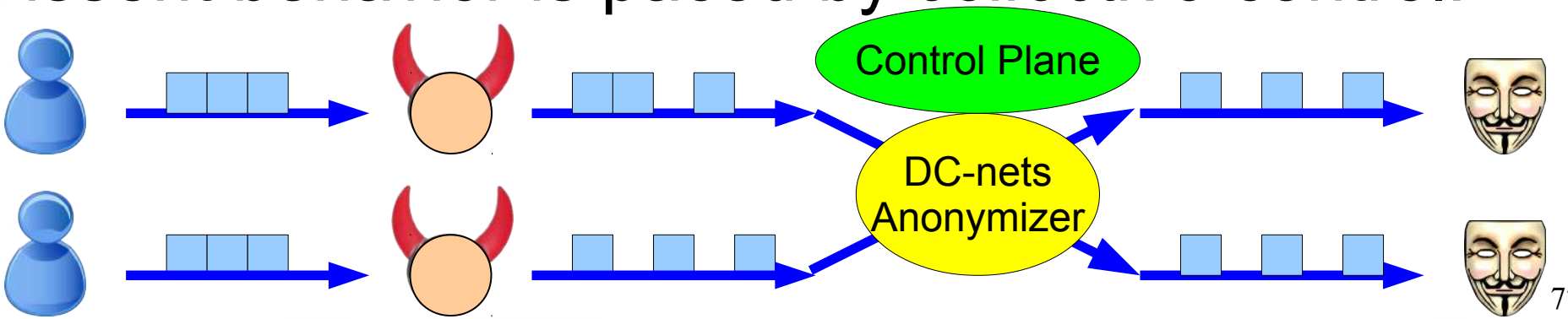


How CCP Counters Active Attacks

Onion routing preserves *individual* flow properties:



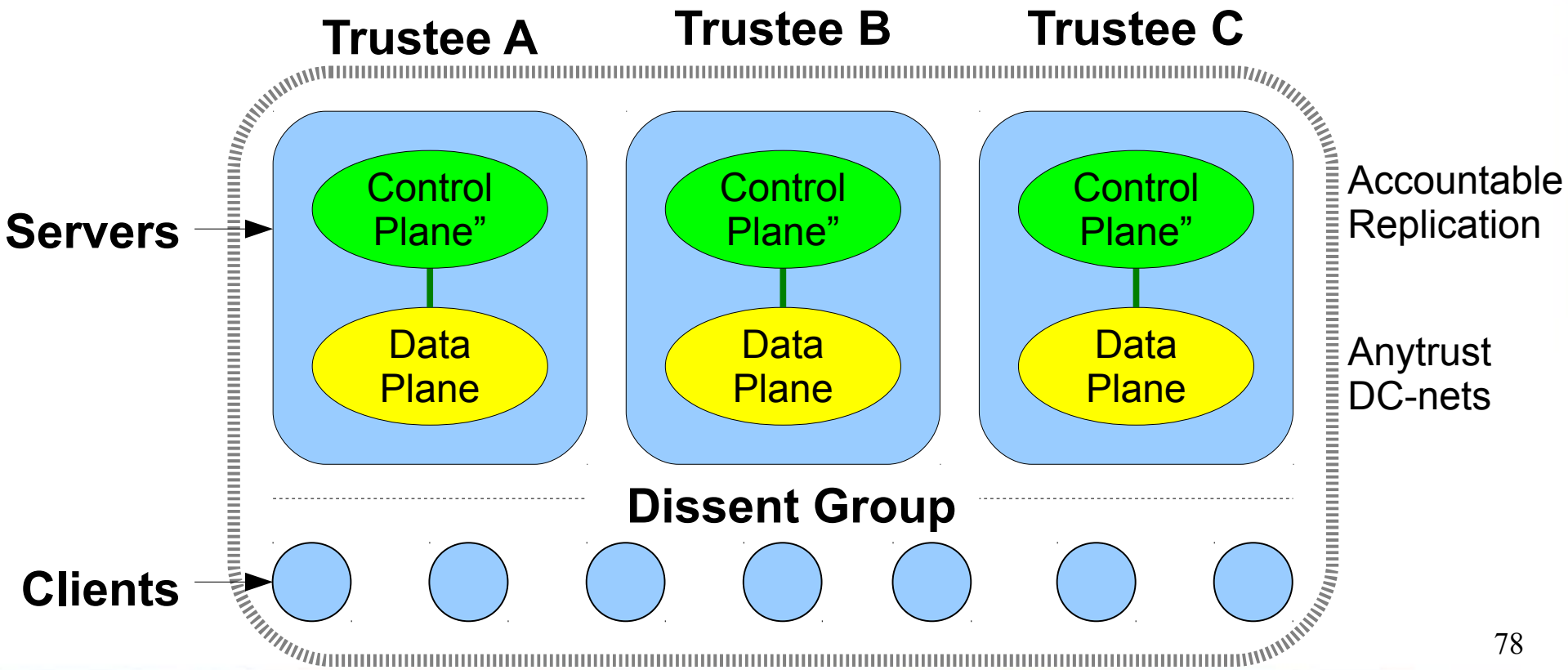
Dissent behavior is paced by *collective* control:



Implementing the CCP

Accountable replication of control plane logic

- Each server implements copy, all must agree

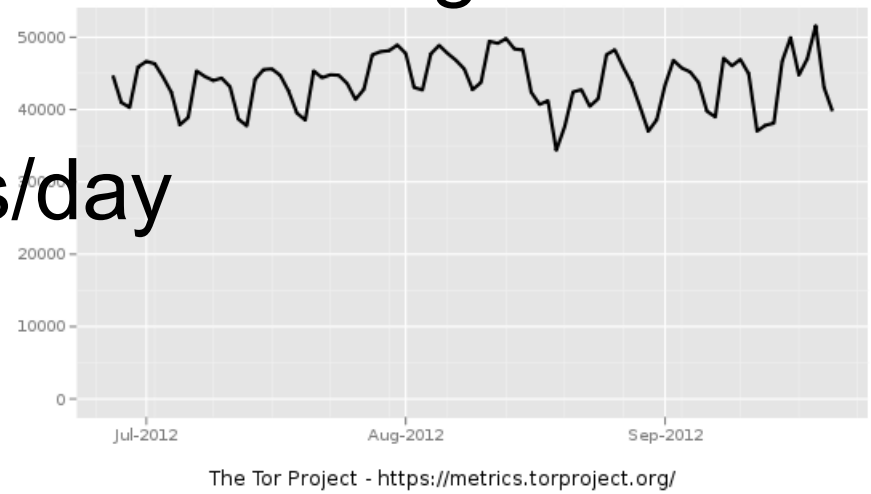


Talk Outline

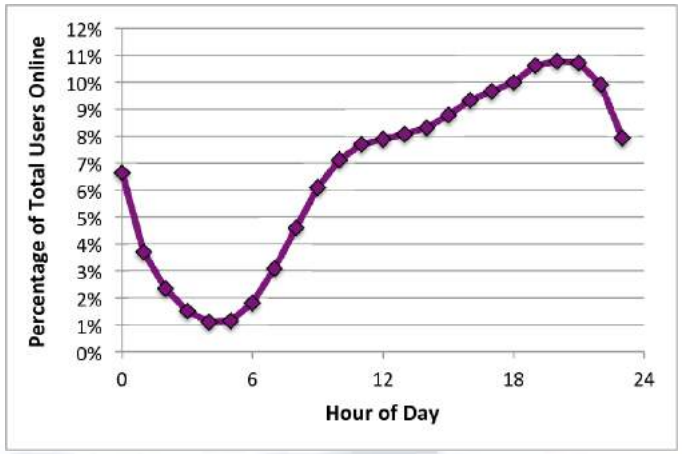
- ✔ Why Anonymity?
- ✔ Current State of the Art
- **Grand Challenges in Anonymity**
  - ✔ Global traffic analysis
  - ✔ Active interference attacks
  - **Intersection attacks**
  - De-anonymizing exploits
  - Accountability provisions
- Status and Ongoing Work

How anonymous are you *really?*

- Bob in Dictatopia posts via Tor to blog hosted in “The Free World”™
- Tor Metrics: 50,000 users/day connect from Dictatopia
  - Good anonymity, right?



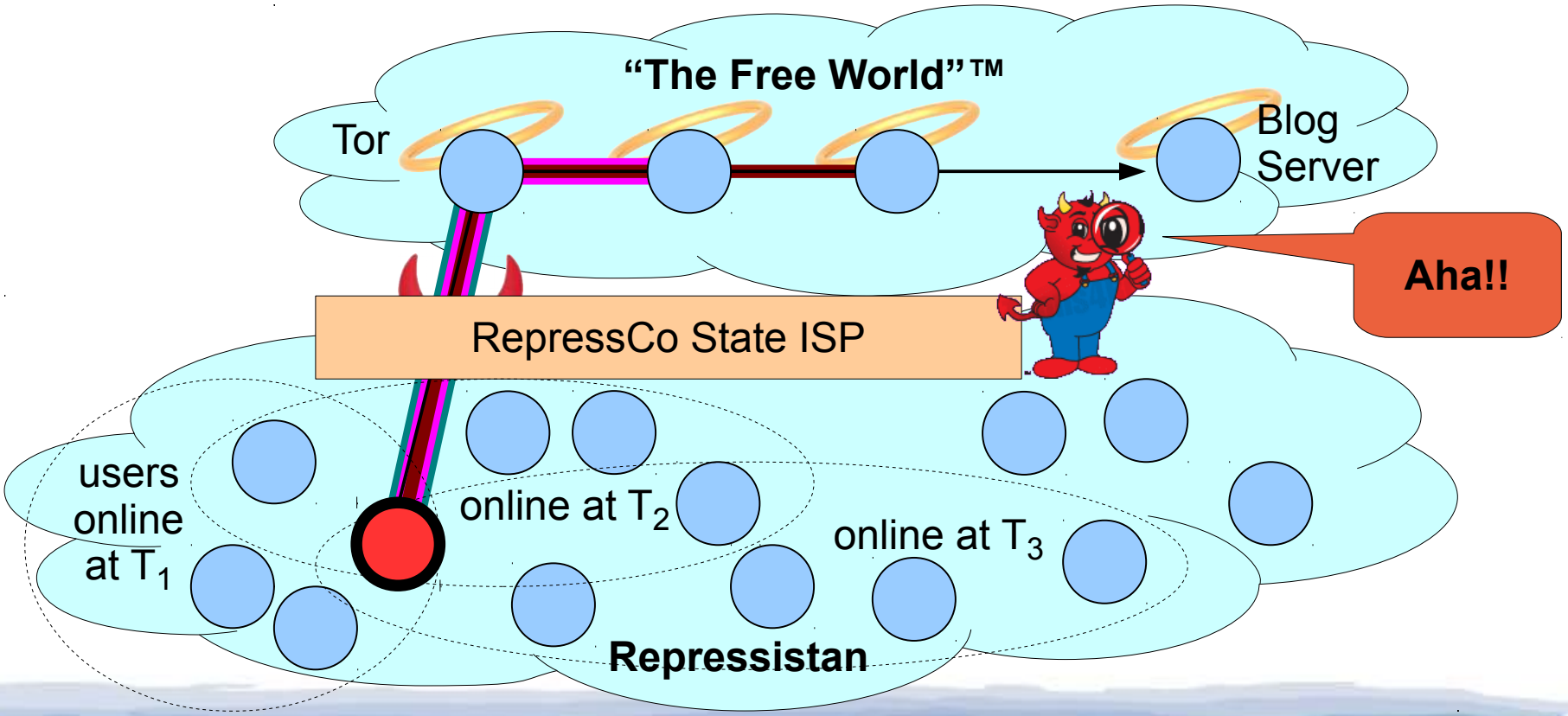
- But ISP logs tell police when users are online; blog post has timestamp
  - How many users are online ***at same time Bob posts?***
    - ~5,000 at 7PM?
      ~500 at 5AM?



The Intersection Attack Problem

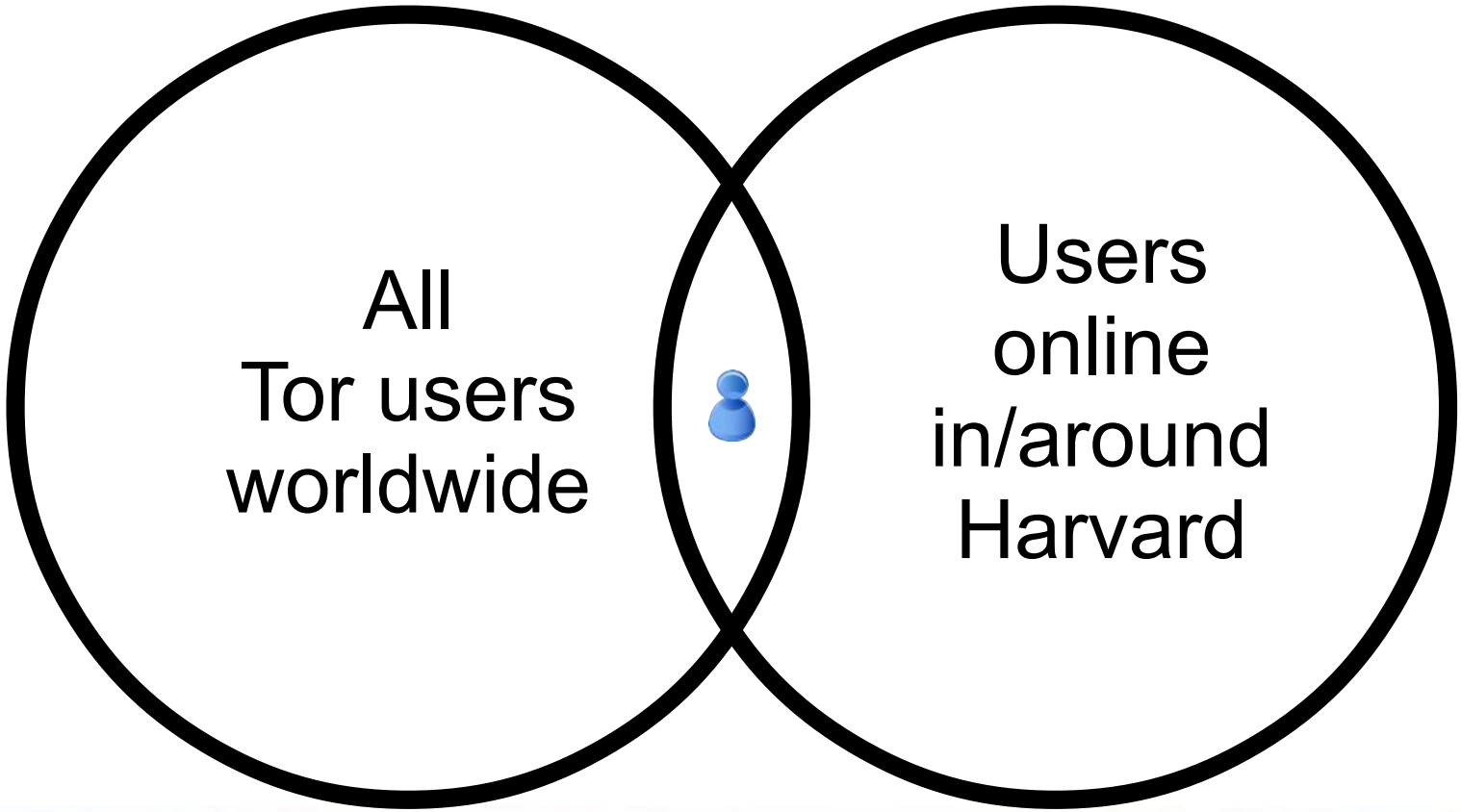
Kate signs posts with pseudonym “Bob”

- Posts signed messages at times T_1, T_2, T_3
- Police **intersects** user sets online each time



The Bomb Hoax Attack

The Harvard bomb hoaxer was de-anonymized by a particularly trivial intersection attack



Buddies [CCS '13]

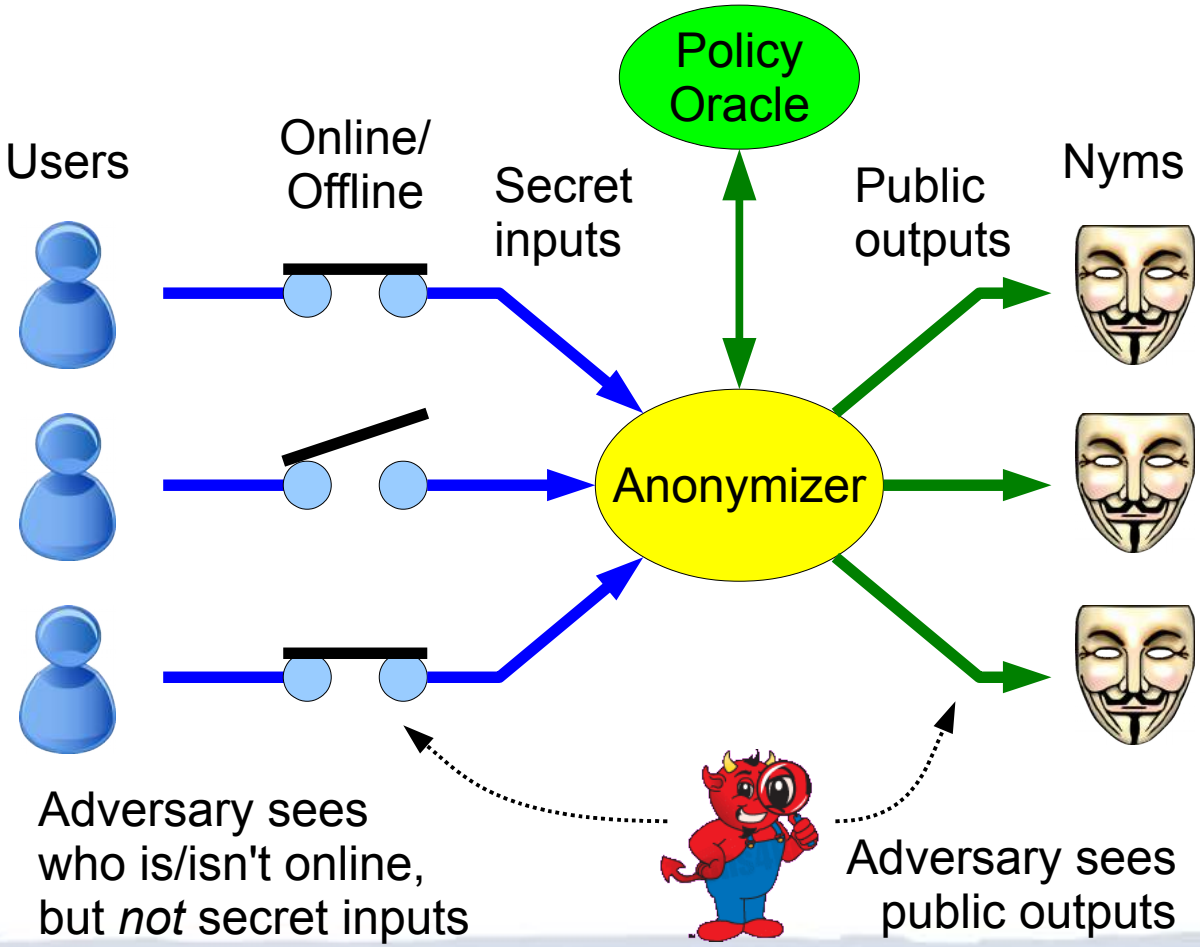
First attempt at building intersection attack resistance into a practical anonymity system

Goals:

- *Measure* anonymity under intersection attack
- Actively *mitigate* anonymity loss
- Enforce *lower bounds* by trading availability

Buddies Conceptual Model

Focus: what adversary learns from *online status*



Computing Anonymity Metrics

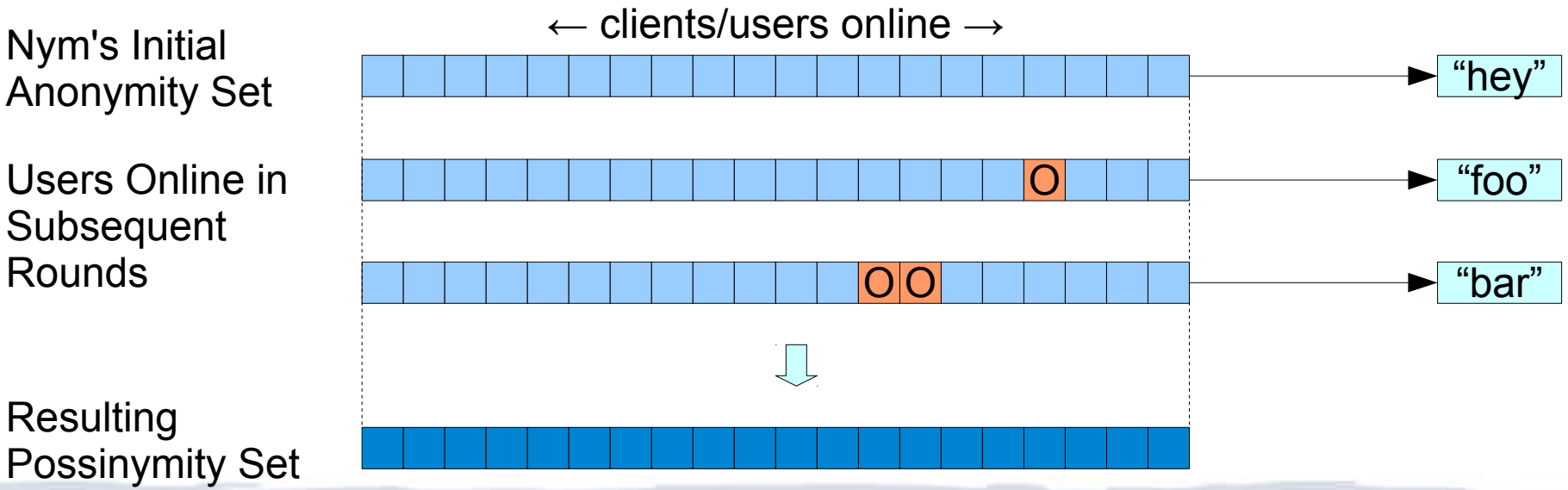
Policy Oracle *simulates an adversary's view*

- Knows who's online each round (via “tags”)
- Simulates “intersection attacks” against Nyms
- Computes anonymity metrics
    - **Possinymity:** “possibilistic deniability”
    - **Indinymity:** “probabilistic indistinguishability”
- Reports metrics, uses them in policy decisions

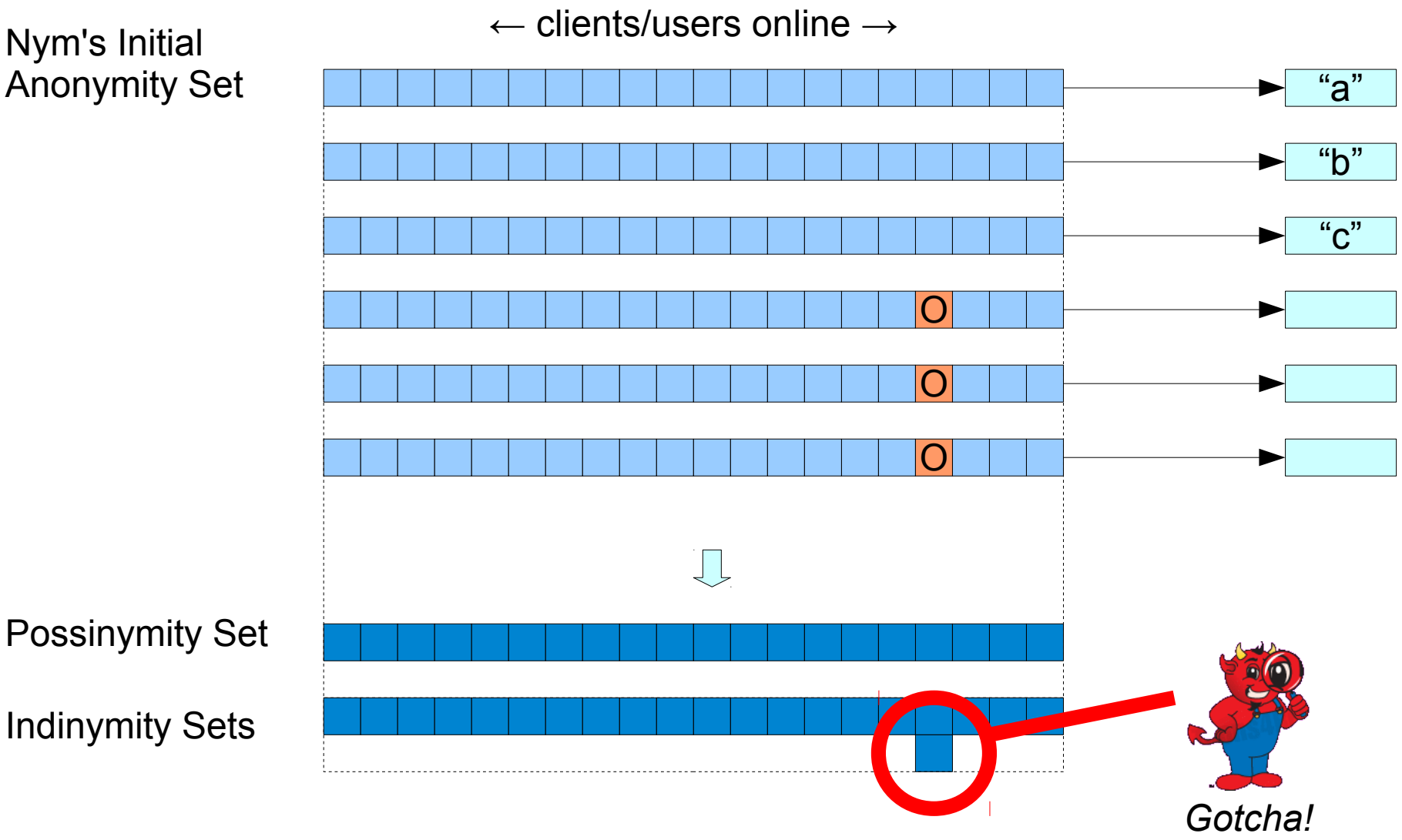
Possinymity: Possibilistic Deniability

Set of users who *could conceivably* own Nym

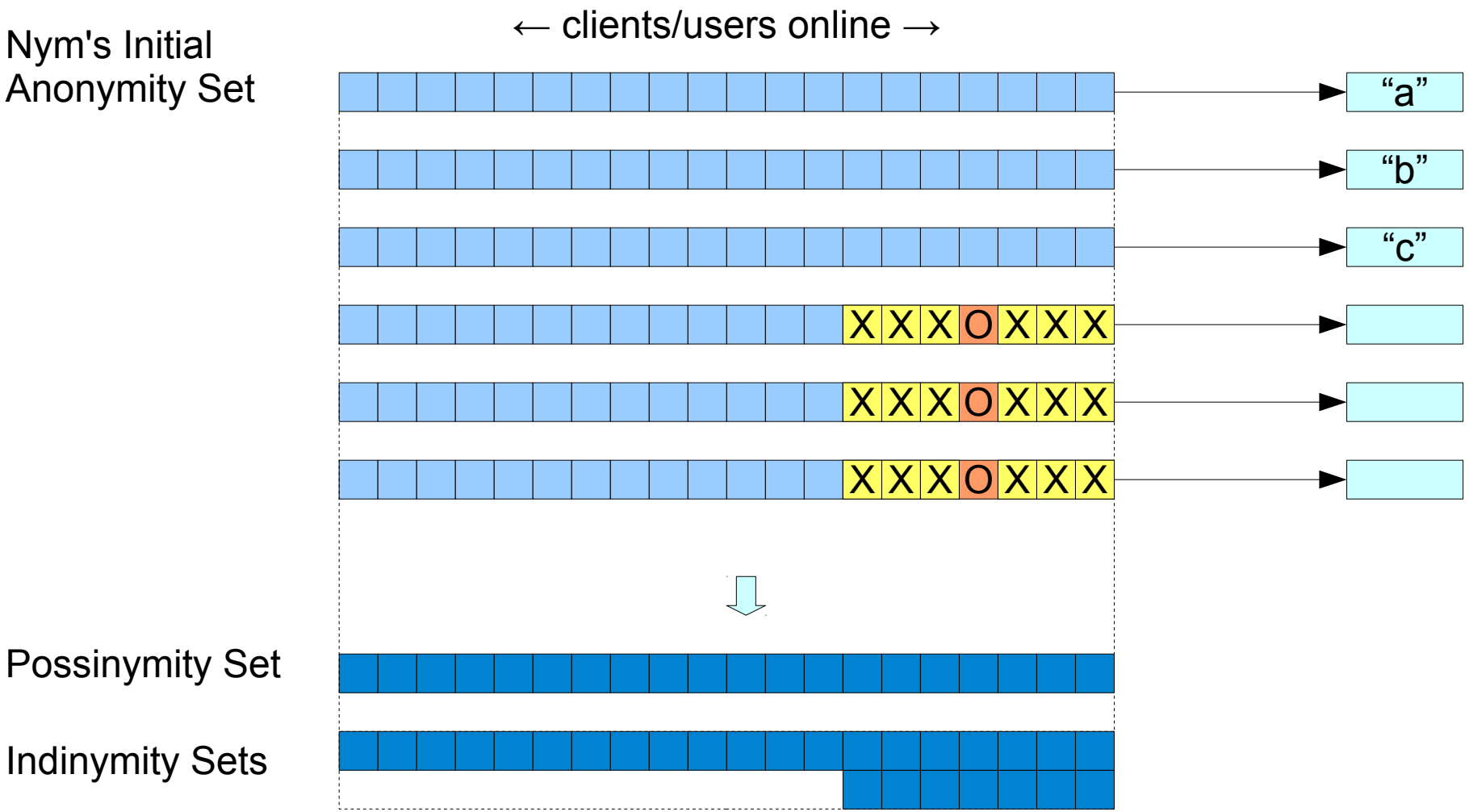
- Intersection of sets of all users *online and unfiltered* in rounds where *a message appears*
- Simplistic, but may build “reasonable doubt”



The “Statistical Disclosure” Problem

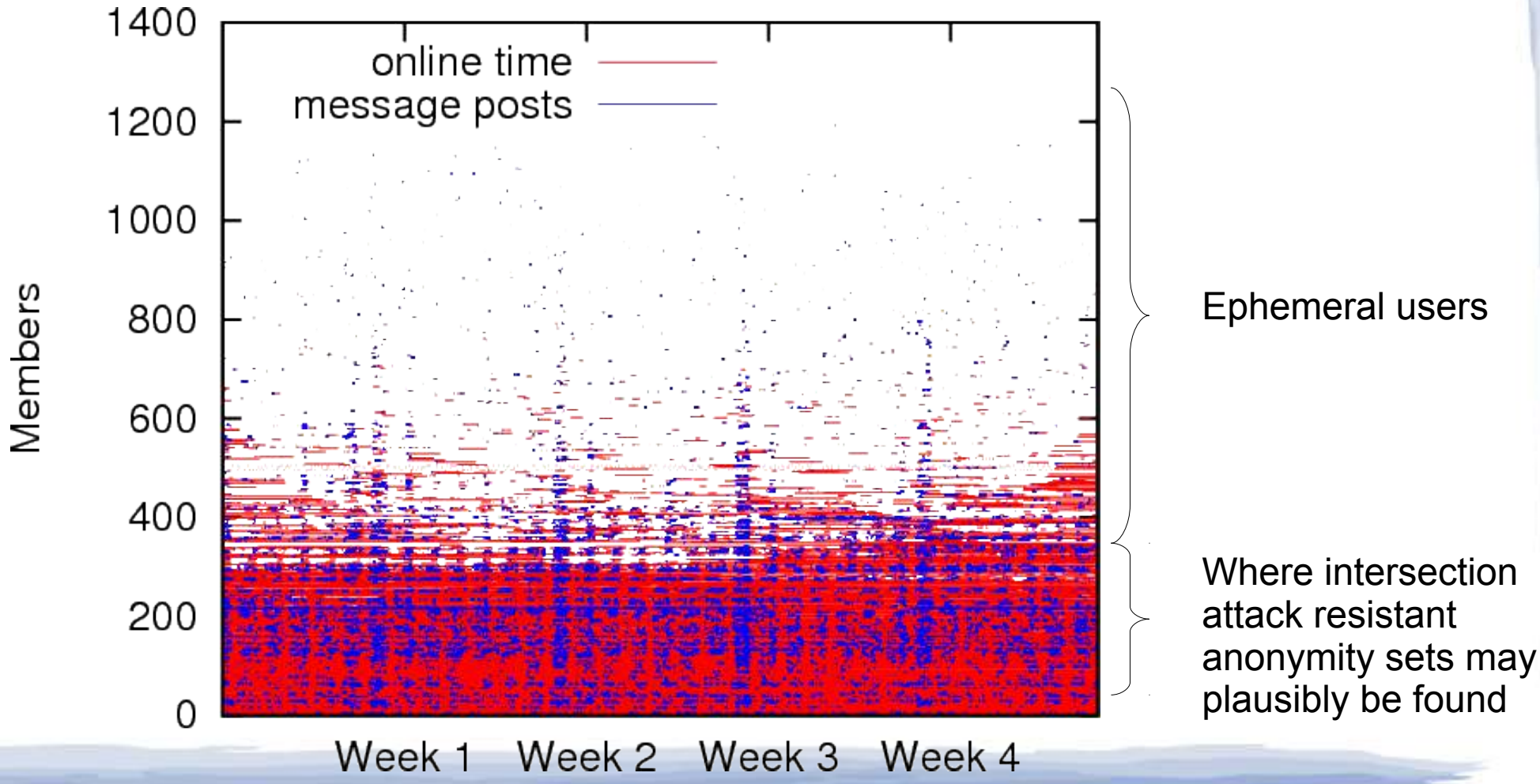


How Dissent Preserves Indinymity

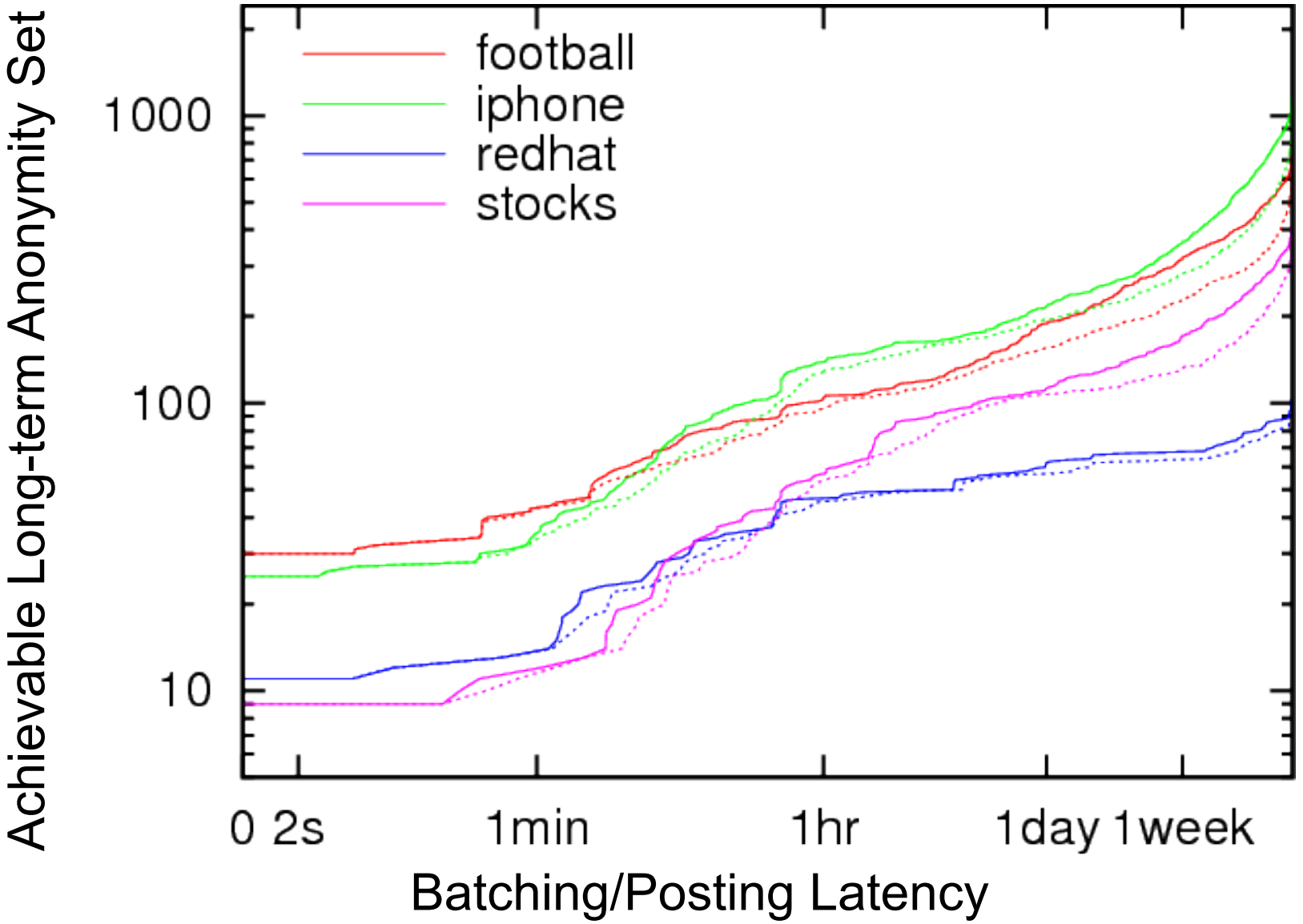


How effective? Depends on users...

Analysis based on IRC online status traces



Achievable anonymity fundamentally depends on *latency tolerance*



AnonRep

(separate slide deck)