IRBs and Computer Science Research

Simson L. Garfinkel
Associate Professor
Naval Postgraduate School

"The views expressed in this presentation do not necessarily reflect the position of the US Government or the Department of Defense."

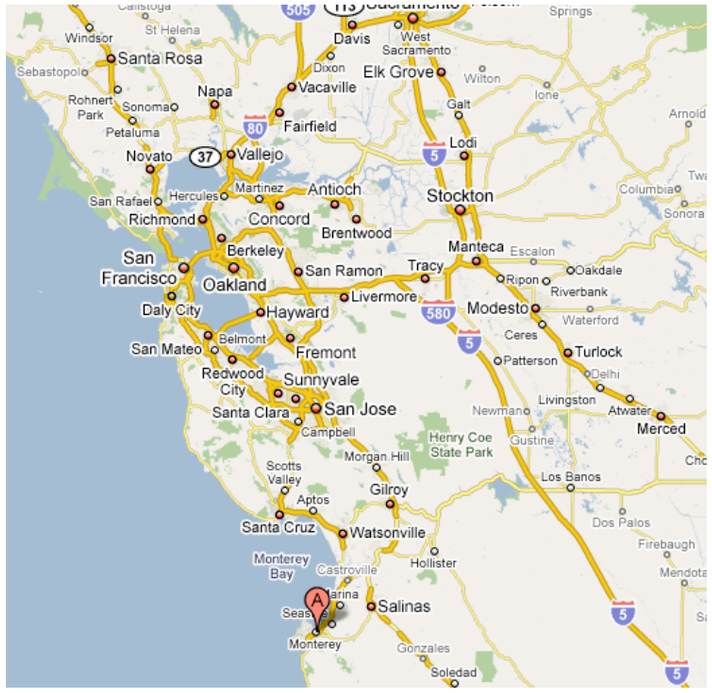
NPS is a research university operated by the Department of Defense.

Located in: Monterey, CA

Campus Size: 627 acres

Students: 1500

- US Military (All 5 services)
- US Civilian (Scholarship for Service & SMART)
- Foreign Military (30 countries)
- *All students are fully funded!*



Schools:

- Business & Public Policy
- Engineering & Applied Sciences
- Operational & Information Sciences
- International Graduate Studies

Key points in this presentation:

Many Computer Science Researchers:

- Don’t realize they are working with human subject data.
- Don’t know that there are special rules for this data.

What’s wrong with this picture?

Key points in this presentation:

Many Computer Science Researchers:

- Don’t realize they are working with human subject data.
- Don’t know that there are special rules for this data.

Much Computer *Security* Research:

- Is minimal risk
- Requires deception for external validity.
- Debriefing can cause *real harm.*

Many IRBs:

- Aren’t qualified to evaluate CS applications.

Is this “individually identifiable?”

70.134.86.19

Is this “individually identifiable?”

70.134.86.19

IP Address assigned to AT&T DSL service.

Is this “individually identifiable?”

70.134.86.19

5:29pm EST

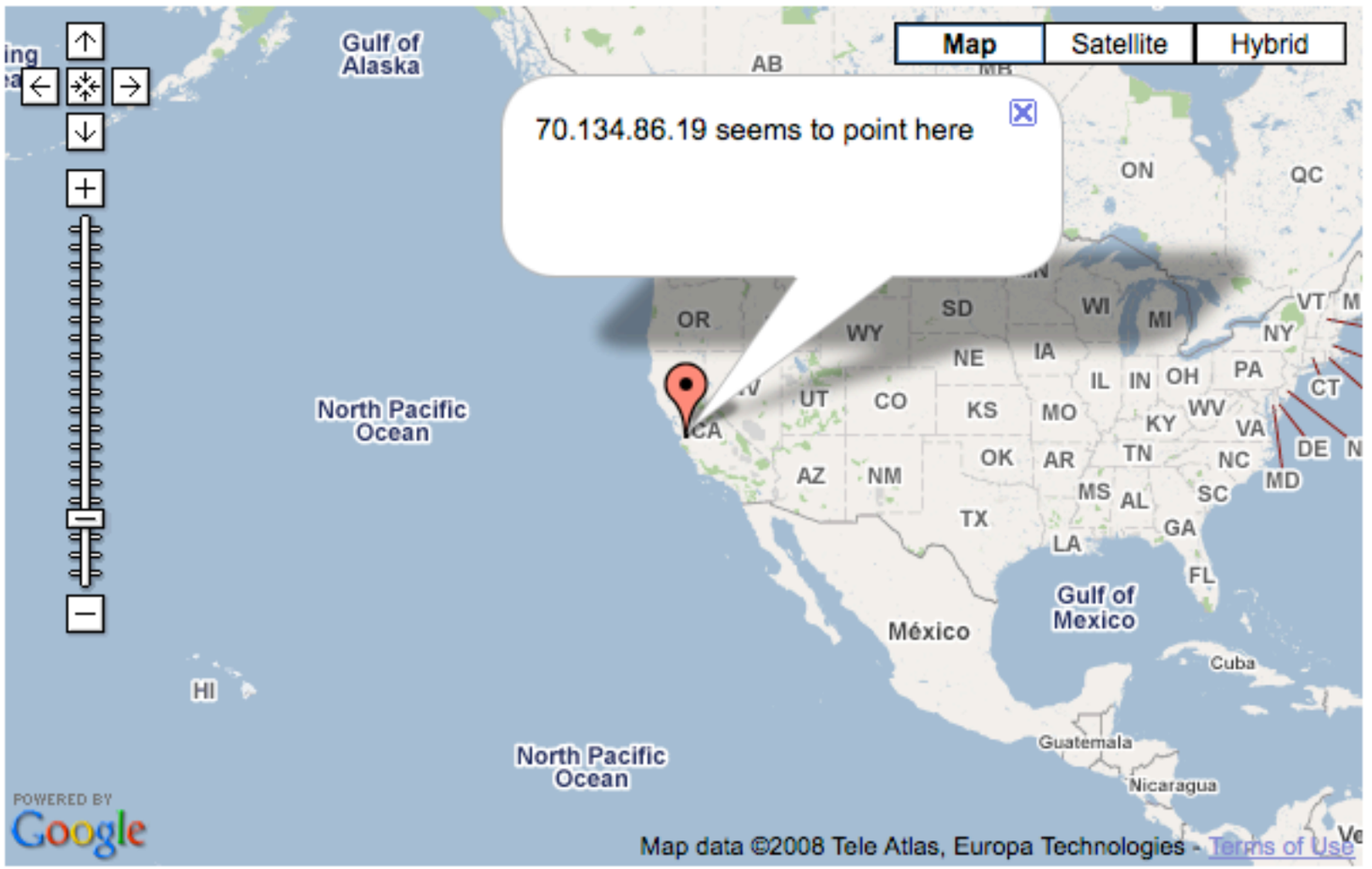
November 13, 2008

70.134.86.19 @ 5:29pm EST November 13, 2008

Geographical **Location** for IP address **70.134.86.19**

Location Info	Value
Country Code	US
Country Name	United States
Region Code	CA
Region Name	California
City	**Salinas**
Postal Code	
Latitude	36.685898
Longitude	-121.6293
DMA Code	828
Area Code	831
TimeZone	America/Los_Angeles

Map

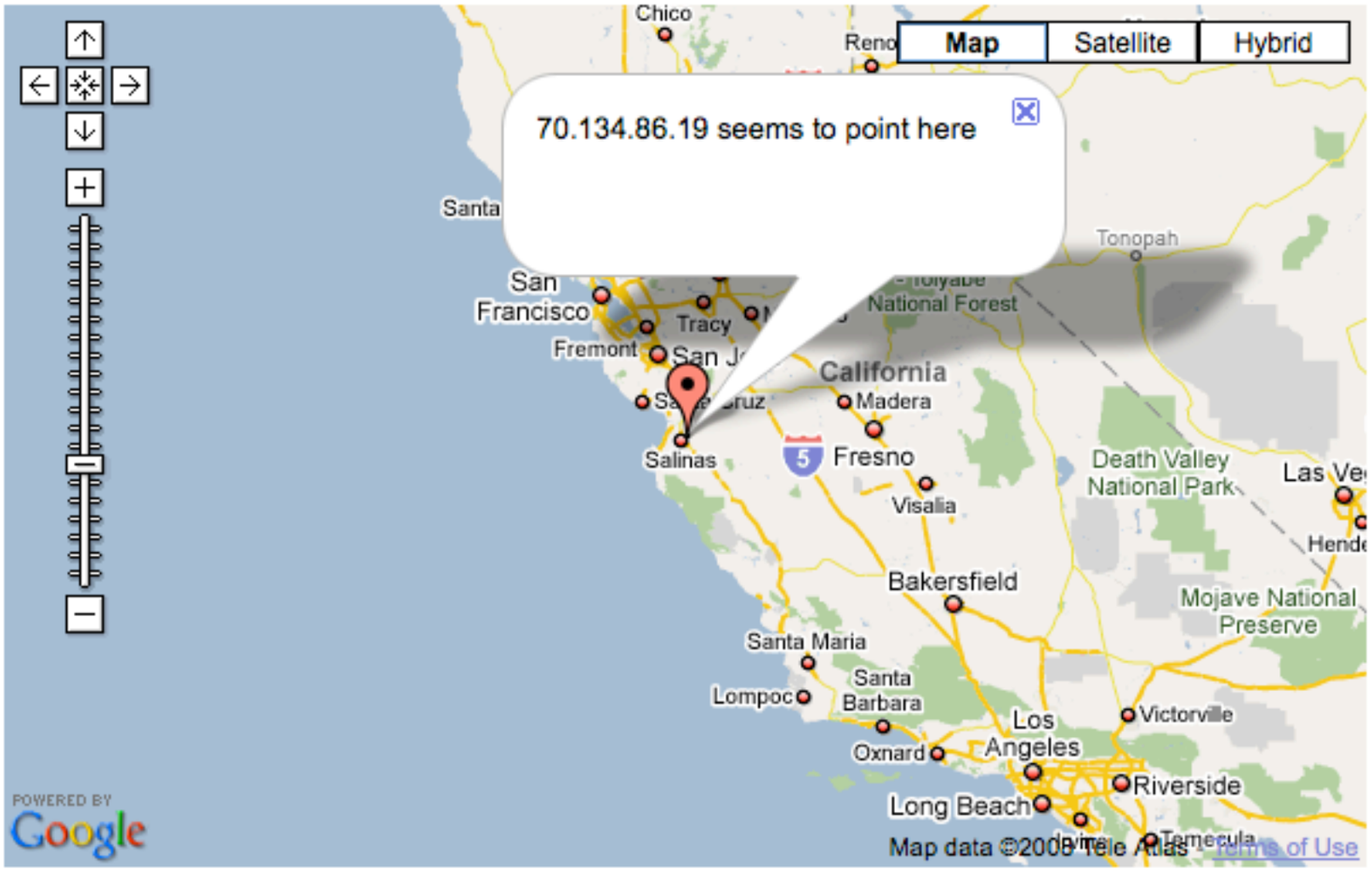


70.134.86.19 @ 5:29pm EST November 13, 2008

Geographical **Location** for IP address **70.134.86.19**

Location Info	Value
Country Code	US
Country Name	United States
Region Code	CA
Region Name	California
City	**Salinas**
Postal Code	
Latitude	36.685898
Longitude	-121.6293
DMA Code	828
Area Code	831
TimeZone	America/Los_Angeles

Map

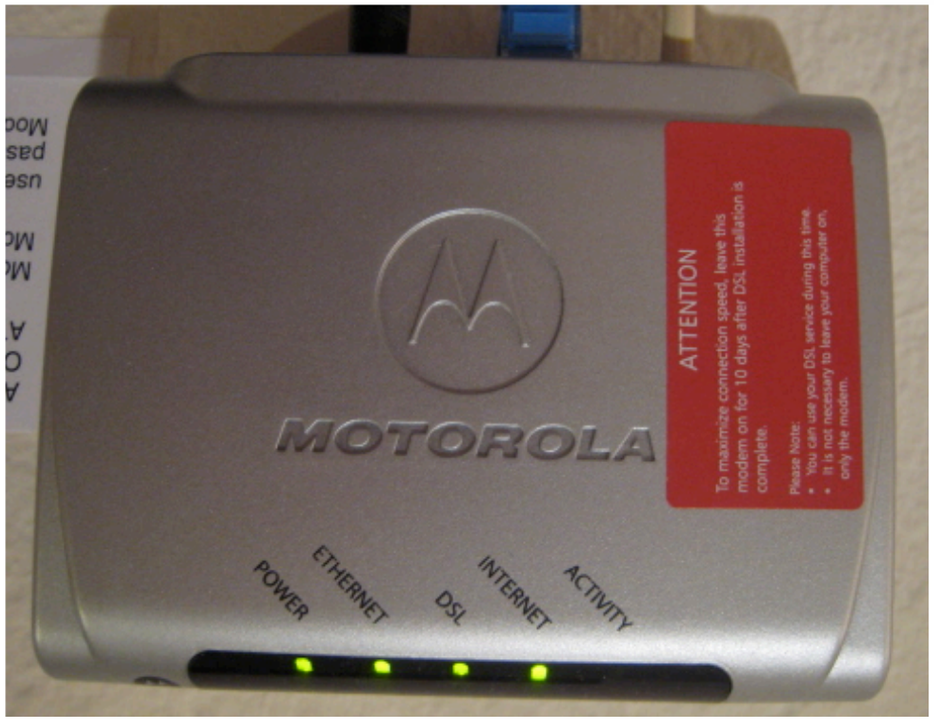


70.134.86.19 @ 5:29pm EST November 13, 2008

70.134.86.19 @ 5:29pm EST November 13, 2008

70.134.86.19 @ 5:29pm EST November 13, 2008

at&t



If you are AT&T or Vonage,
70.134.86.19 @ 5:29pm is individually identifiable.

How about this one?

64.7.15.234

26 Aug 2004
19:23:43

64.7.15.234 was my static IP address from 2001 till 2005.

Google 64.7.15.234 Search Advanced Preference

Web

Web Security, Privacy & Commerce, 2nd Edition: Chapter 8: The ...
... XX+/**64.7.15.234**/2.72.222.209.in-addr. arpa/PTR/IN Apr 20 13:18:20 <local2. info> r2 named[50916]: XX+/**64.7.15.234**/234.15.7.64.in-addr.arpa/PTR/IN Apr 20 ...
oreilly.com/catalog/websec2/chapter/ch08.html - 73k - Cached - Similar pages - Note this

ports/73363: isc-dhcp3-server doesn't honor hard-coded ethernet ... - 11:37pm
... option broadcast-address 192.168.1.255; option subnet-mask 255.255.255.0; option domain-name-servers **64.7.15.234**, # was 192.168.1.2, 216.200.176.4, ...
lists.freebsd.org/pipermail/freebsd-ports-
Cached - Similar pages - Note this

Web | Images | Video | Local | Shopping | more

64.7.15.234 Search Options Customize YAHOO!

1 - 7 of 7 for **64.7.15.234** (About) - 0.68 s | SearchScan BETA

[PPT] www.simson.net/ref/2004/csci_e-170/handouts/L10.ppt
674k - Microsoft Powerpoint - View as html
... H=ns.simson.net (**64.7.15.234**) [**64.7.15.234**] F=<ruxnezze@swissonline.ch> ...
251.dct.al.charter.com) [**64.7.15.234**] F=<mvceubrfvsrm@charter.com> rejected RCPT ...
www.**simson.net**/ref/2004/csci_e-170/handouts/L10.ppt

[PDF] CSCI E-170 Lecture 11: Redux of papers, Logging, The Law, Integrity ...
3472k - Adobe PDF - View as html
2004-11-13 23:51:35 H=ns.simson.net (**64.7.15.234**) [**64.7.15.234**] ... [**64.7.15.234**]
F=<mvceubrfvsrm@charter.com> rejected RCPT <gayda@ex.com>: Unknown user ...
www.**simson.net**/ref/2005/csci_e-170/slides/L11.pdf

Search 64.7.15.234 on Google or Yahoo, and you'll find me.

Are IP addresses PII?

"Sometimes... it depends on the context"

- **Peter Fleischer, Google, addressing the European Parliament Civil Liberties Committee, 21 January 2008.**

"In most cases IP addresses have to be seen as personal related and therefore the European Directive on Data Protection covers also the use of IP addresses."

- **Peter Schaar,German Federal Data Protection Commissioner, 31 Jan. 08**

"Dynamic IP Addresses Are Not Personally Identifiable Information Because They Are Anonymous, Temporary, And Only identify internet devices."

- **TRUSTe *Amicus Curiae*, Jeffrey Klimas v. Comcast, 465 F.3d 271, 273 (2006)**

Is your IP address PII?
It depends on who is looking.

OHRP says "private information must be individually identifiable in order for obtaining the information to constitute research involving human subjects."

Guidance on Research Involving Coded Private Information on Biological Specimens, October 16, 2008

http://www.dhhs.gov/ohrp/humansubjects/guidance/cdebiol.htm

Increasingly:

If there are individually records

Then the individuals can be identified (with effort).

Main message…

A significant amount of Computer Science research:

- Directly involves human beings or data.
- Is minimal risk.
- Cannot be reasonably performed with informed consent.

This work requires IRB approval under 45 CFR 46.

That's because:

- It involves human subjects.
- Experimenters are not allowed to exempt their own research!

AOL Search History Fiasco: August 2006

8/4/06 — AOL posted a link for CS researchers

- 20 million “anonymized” searches
- 658,000 customers, March 2006-May 2006

AnonID	Query	Querytime	Click URL	RANK
4417749	Movies for Dogs	2006-03-02 09:24:14		
4417749	blue book	2006-03-02 11:48:52	http://kbb.com	1
4417749	best dog for older owner	2006-03-02 11:48:24	http://caismajor.com	1

8/7/06 — AOL removed the link following complaints.

8/9/06 — NYT identifies AOL search user.

- Despite anonymization, some users identified by their search terms.

Netflix “prize” fiasco:
$1M for 10% improvement in recommendations.

Netflix released for researchers:

- Title of movie
- Year of movie release
- Ranking

Researchers Correlate with:

- Reviews posted to Internet Movie Database (IMDb)

Resulting Scientific Publication:

- “Robust De-Anonymization of Large Sparse Datasets,”
  Arvind Narayanan and Vitaly Shmatikov,
  IEEE Symposium on Security and Privacy 2008

Open question for IRB chairs:

If publicly available data can be reverse-engineered to reveal privacy sensitive facts,

but nobody knows that this reverse engineering is possible,

does the research require IRB approval?

Identity Trail: Covert Surveillance Using DNS
Guha & Francis, PET2007

Many companies provide "Dynamic DNS" updates.

- DynDNS, No-IP, TZO and others

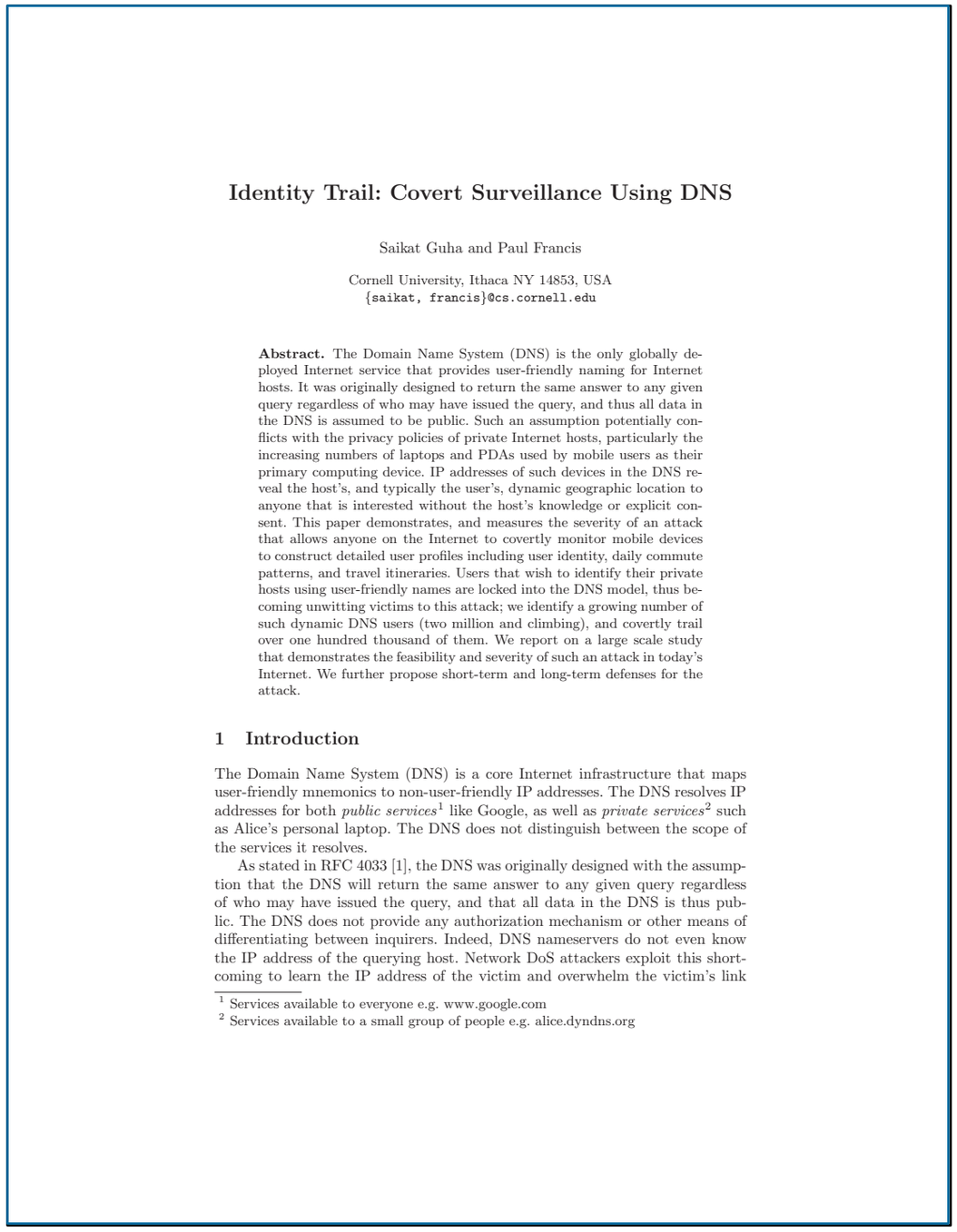
Typical vanity DNS address:

- laptop.simson.net = 172.20.0.43

Typical uses:

- Private web servers
- Gaming
- FTP
- Web cams

By design, DNS is a public service: no access controls.



Identity Trail: Covert Surveillance Using DNS

Saikat Guha and Paul Francis

Cornell University, Ithaca NY 14853, USA
{saikat, francis}@cs.cornell.edu

Abstract. The Domain Name System (DNS) is the only globally deployed Internet service that provides user-friendly naming for Internet hosts. It was originally designed to return the same answer to any given query regardless of who may have issued the query, and thus all data in the DNS is assumed to be public. Such an assumption potentially conflicts with the privacy policies of private Internet hosts, particularly the increasing numbers of laptops and PDAs used by mobile users as their primary computing device. IP addresses of such devices in the DNS reveal the host's, and typically the user's, dynamic geographic location to anyone that is interested without the host's knowledge or explicit consent. This paper demonstrates, and measures the severity of an attack that allows anyone on the Internet to covertly monitor mobile devices to construct detailed user profiles including user identity, daily commute patterns, and travel itineraries. Users that wish to identify their private hosts using user-friendly names are locked into the DNS model, thus becoming unwitting victims to this attack; we identify a growing number of such dynamic DNS users (two million and climbing), and covertly trail over one hundred thousand of them. We report on a large scale study that demonstrates the feasibility and severity of such an attack in today's Internet. We further propose short-term and long-term defenses for the attack.

1 Introduction

The Domain Name System (DNS) is a core Internet infrastructure that maps user-friendly mnemonics to non-user-friendly IP addresses. The DNS resolves IP addresses for both *public services*[1] like Google, as well as *private services*[2] such as Alice's personal laptop. The DNS does not distinguish between the scope of the services it resolves.

As stated in RFC 4033 [1], the DNS was originally designed with the assumption that the DNS will return the same answer to any given query regardless of who may have issued the query, and that all data in the DNS is thus public. The DNS does not provide any authorization mechanism or other means of differentiating between inquirers. Indeed, DNS nameservers do not even know the IP address of the querying host. Network DoS attackers exploit this shortcoming to learn the IP address of the victim and overwhelm the victim's link

[1] Services available to everyone e.g. www.google.com
[2] Services available to a small group of people e.g. alice.dyndns.org

"We discovered 36,011 potential victims through a variety of methods."

Users identified using public information:

- Web pages, mailing lists archives, etc.: 4,351 DNS names
- Dictionary attack of common names: 31,660 DNS names

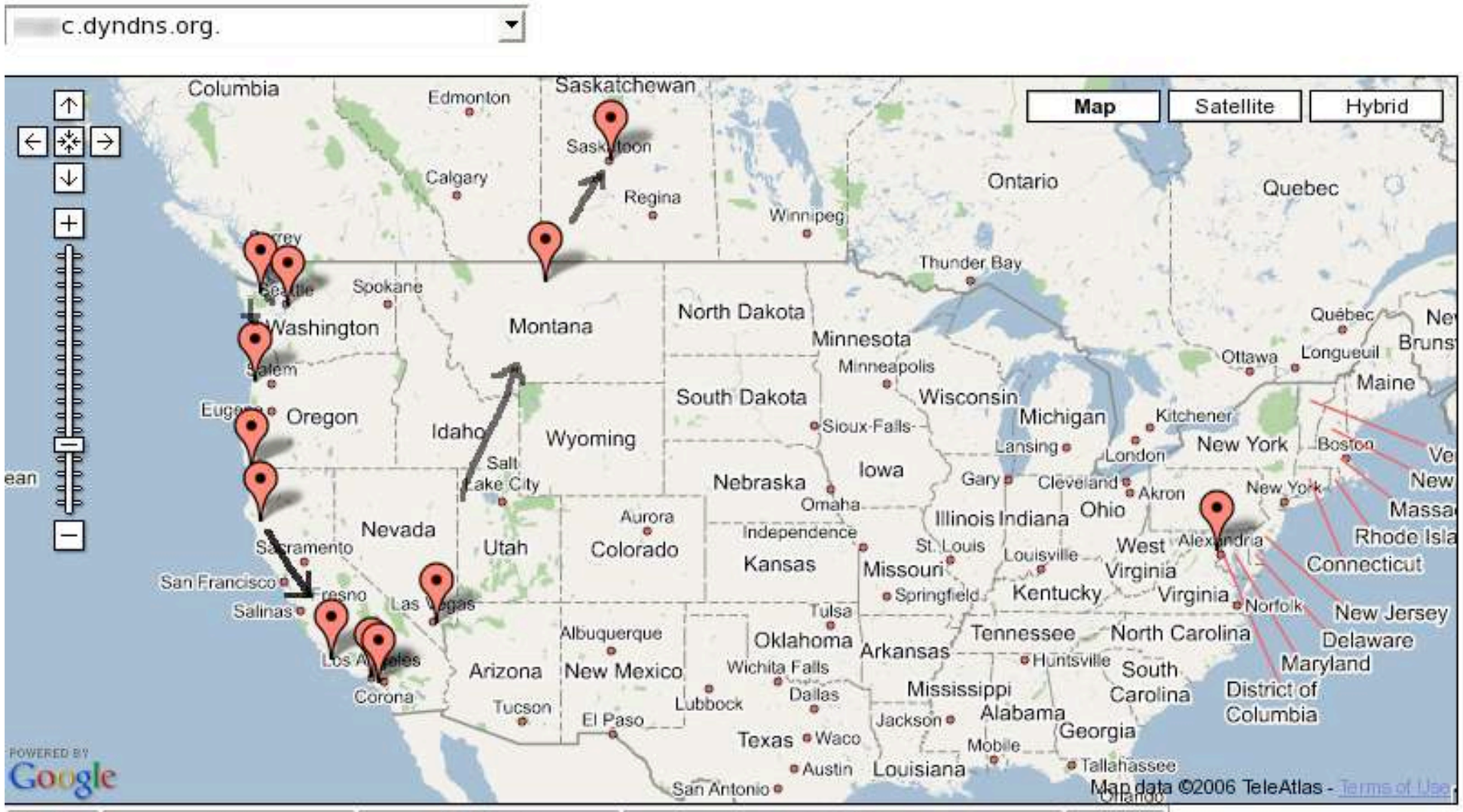
Users monitored with DNS queries:

- "We monitored 18,720 hosts from July 20, 2006 to August 8, 2006"

Ground truth determination:

- Some "victims" were *contacted* to verify accuracy of surveillance.

"Figure 2 is a screenshot of a summer road-trip taken by user M as tracked by our application."



Date	ip	Block	Location	AS
Jul 20	24.[redacted]	24.22.192.0/18	Seattle, Washington, US	33650
Jul 20	24.[redacted]	24.22.192.0/18	Seattle, Washington, US	33650

Fig. 2. Tracking a user's summer road-trip through the DNS

"Figure 3 plots the mobility of one of the authors from August 18, 2006 to November 2, 2006."

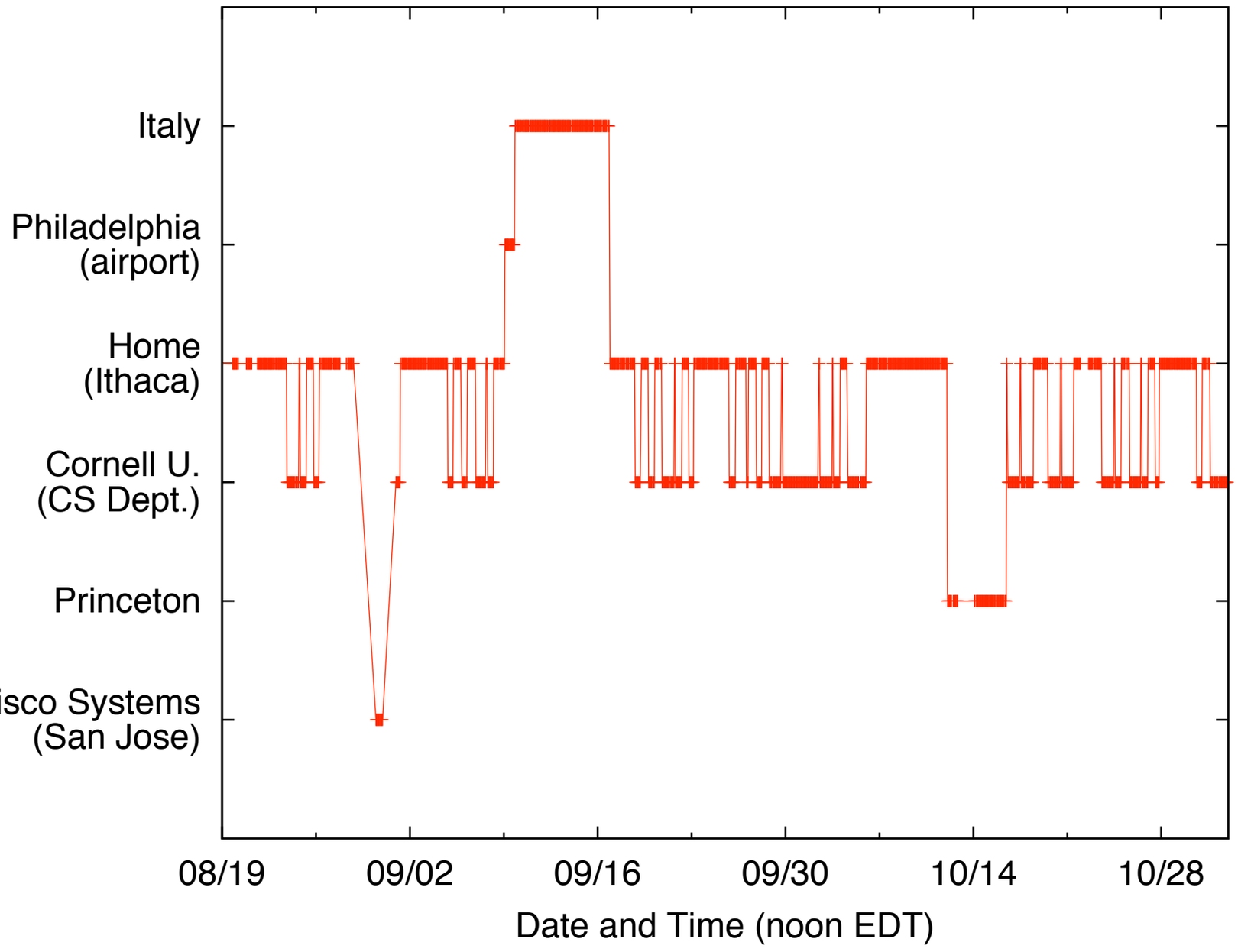


Fig. 3. Tracking a user's daily commute and travel through the DNS

Is IRB approval required for Self-Experimentation?

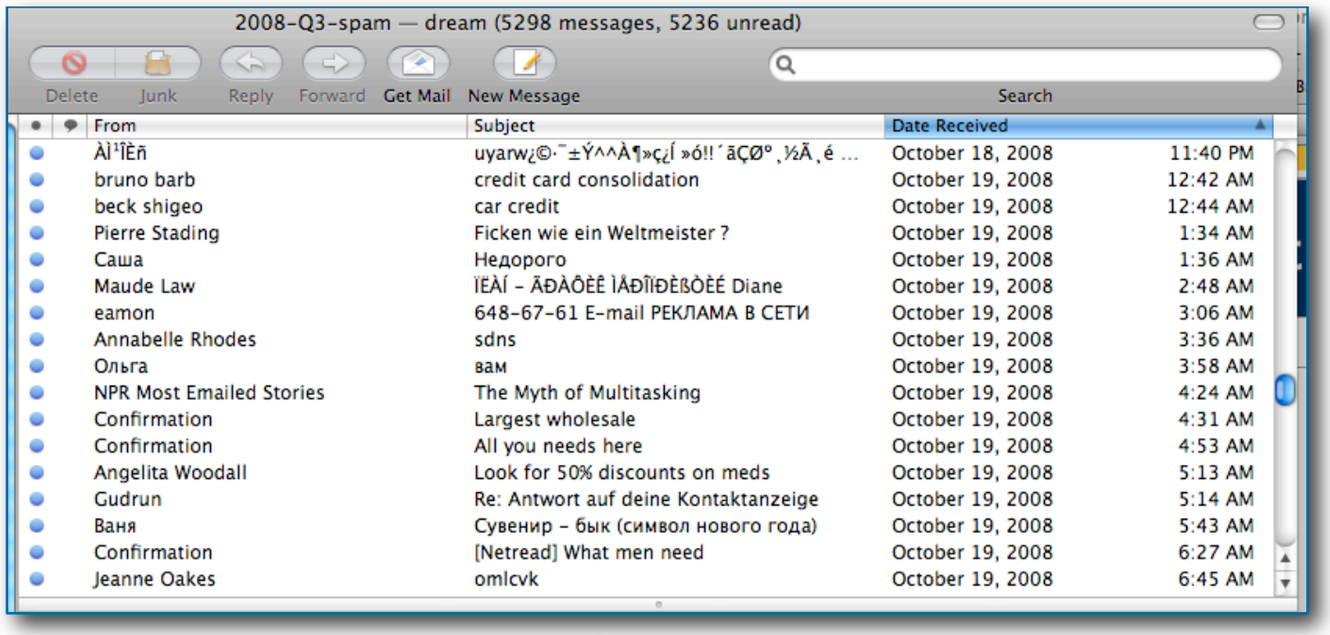
"In the case of self-experimentation... IRBs are at a considerable disadvantage. **There are no specific guidelines to cover their actions.** The OPRR does not include self-experimenters among its special classes of research subjects. On the other hand, the OPRR considers normal volunteers to be a vulnerable population..."

"Nevertheless, **we believe that the IRB must exercise its same best judgement on self-experimentation, as it does on all other human research...**"

- "A Case of Self-Experimentation," Editorial, Cancer Epidemiology, Biomarkers & Prevention, Vol. 6, 475-476, July 1997.

(Office for Protection from Research Risks became Office for Human Research Protections in 2000.)

Security research increasingly involves human subjects.



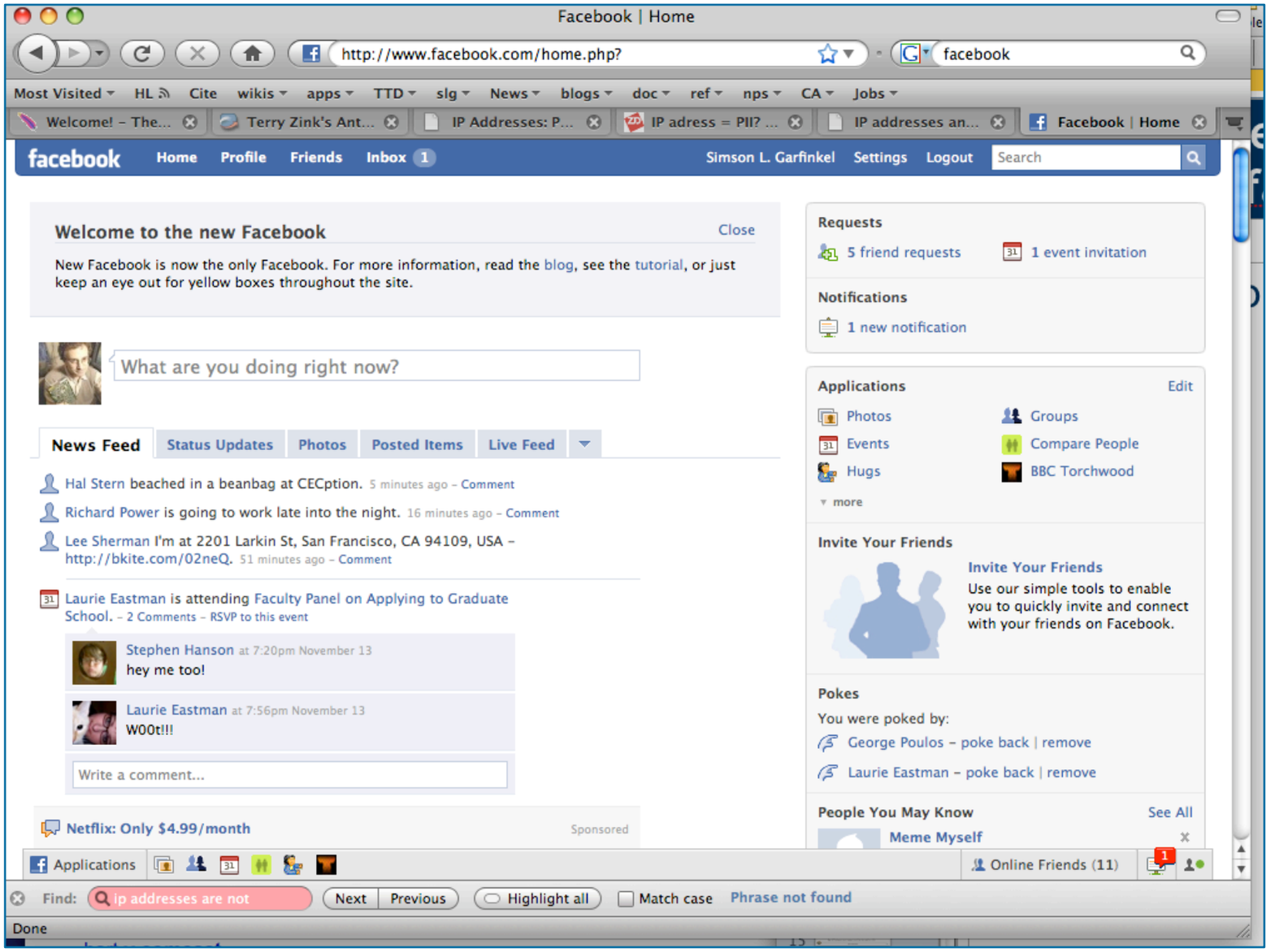
SPAM

Wireless Usability

phishing

This research is vital for our national interests.

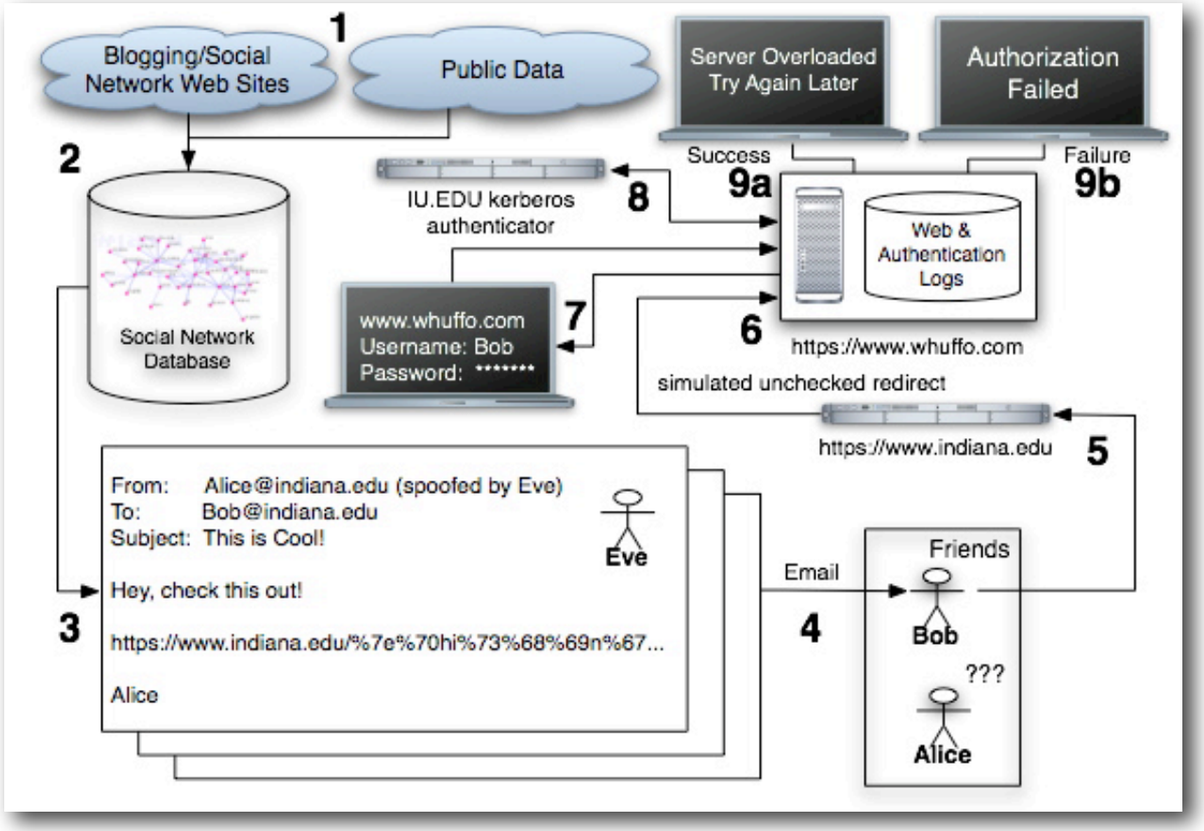
Research question:
Can data from Facebook improve phishing?



"Social Phishing," Jagatic, Johnson, Jakobsson, Menczer, Indiana University, 2005

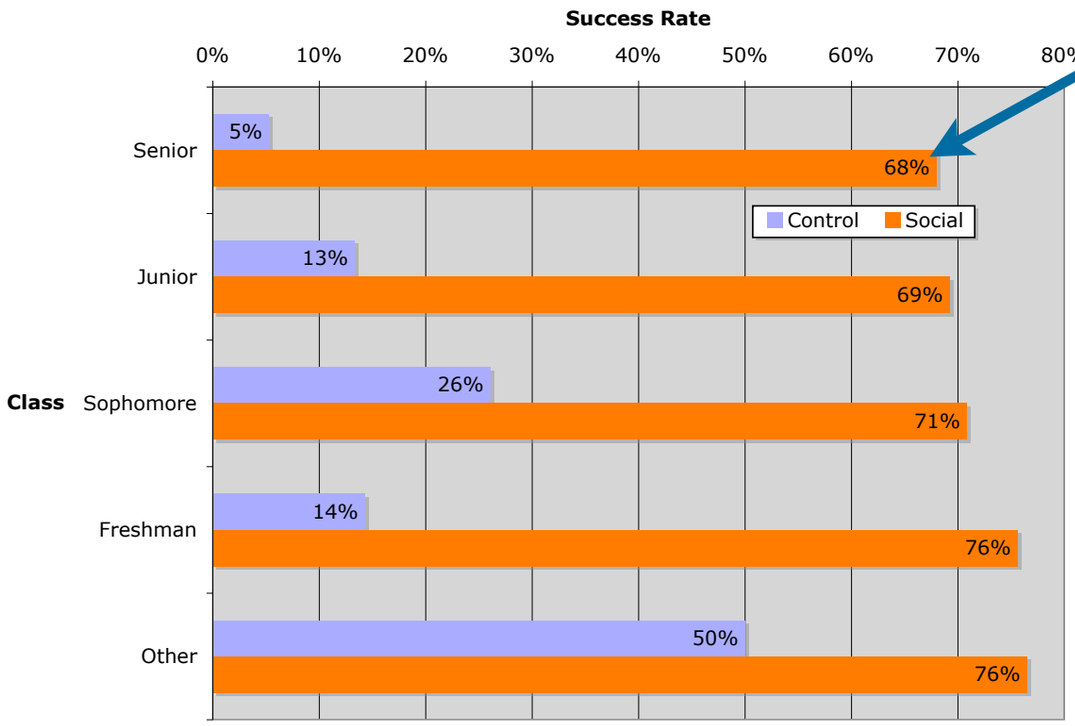
Protocol:

- Search Facebook for IU students.
- Email Alice with fake mail from "Bob."
- "Hey, check this out!"
- https://www.whuffo.com/
- Prompt students for IU username & password.



Facebook helps a lot!

With Facebook



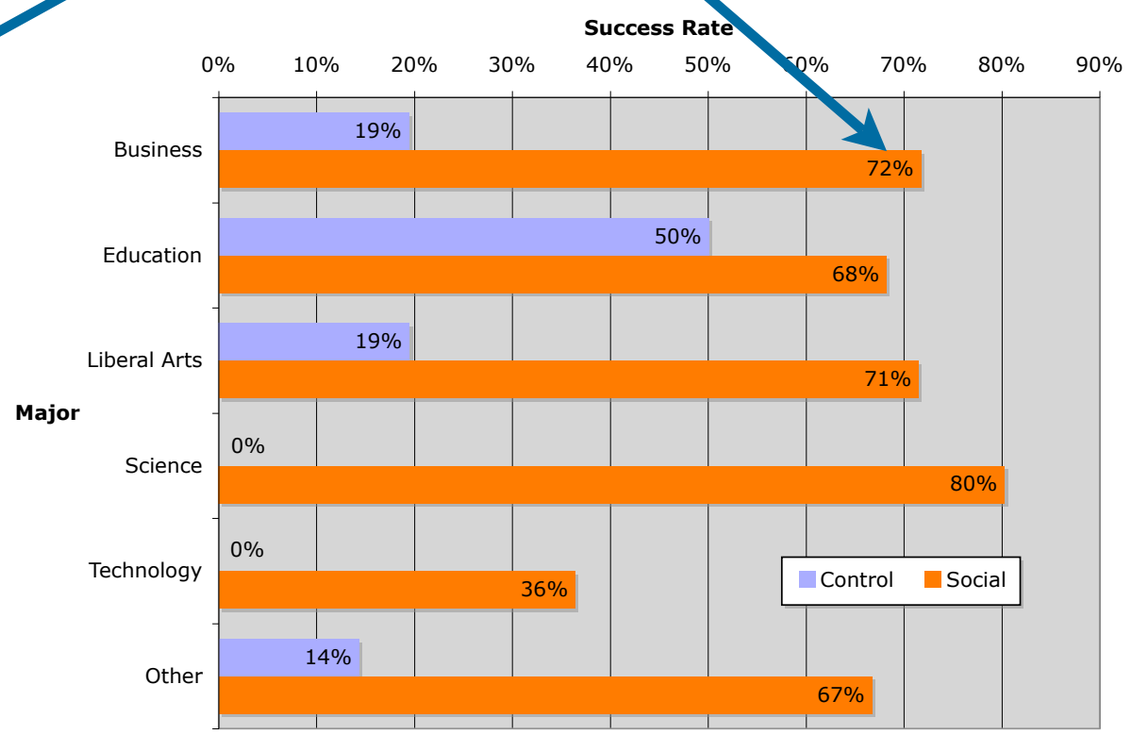


Figure 3: Success rate of phishing attack by target class.

Social Phishing study was heavily criticized.

Key aspects of study:

- Performed with cooperation of Indiana University Security Staff.
- Students involved *without their consent!*
- *Deception!*
- IRB Approval!

Major complaint:

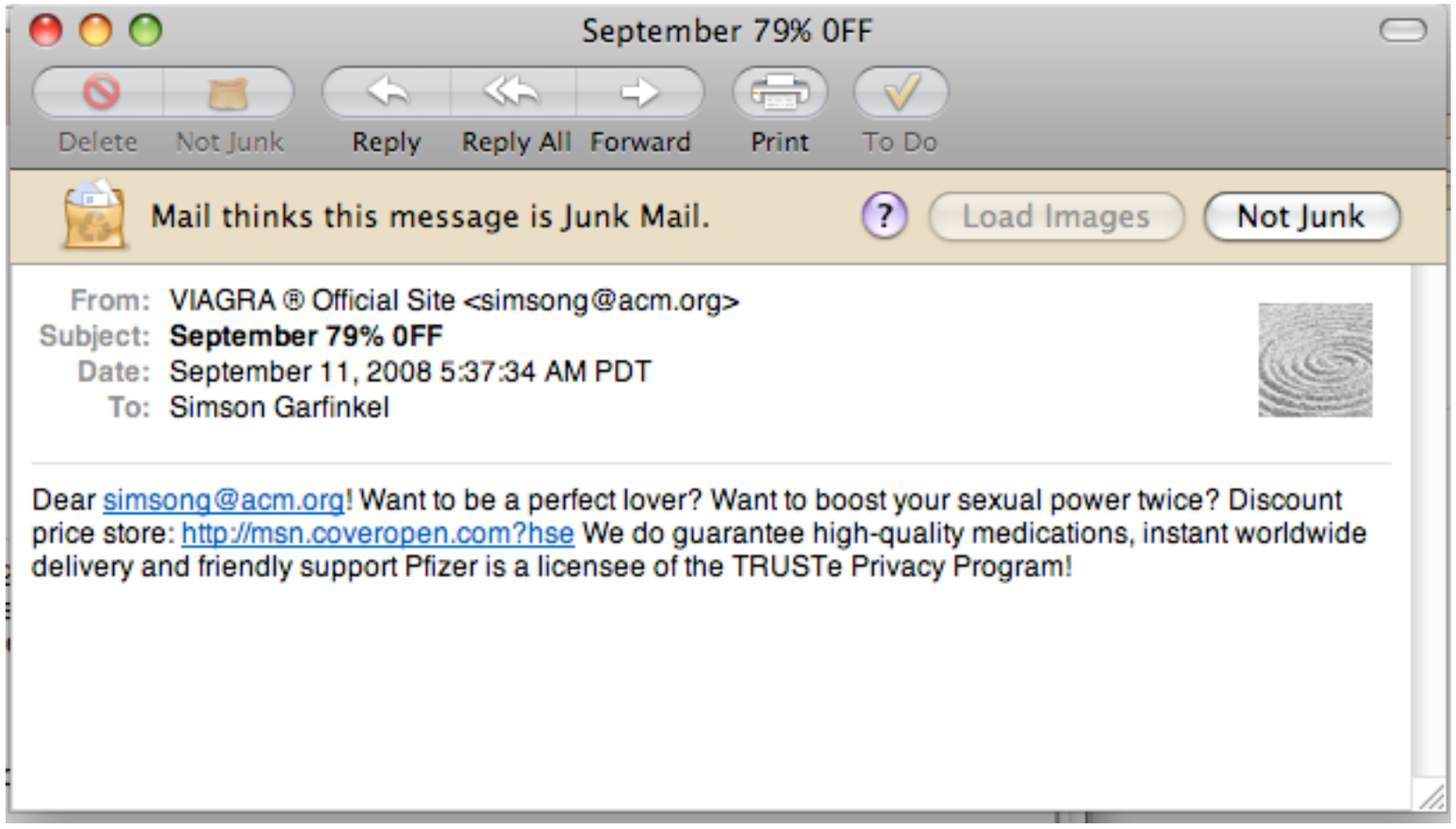
- Telling students that they had been successfully phished caused stress!

The IRB required the “debriefing.”

- Was this a mistake?

Phishing attacks are “minimal risk.” --- !

Research Question:
How do spammers make money?



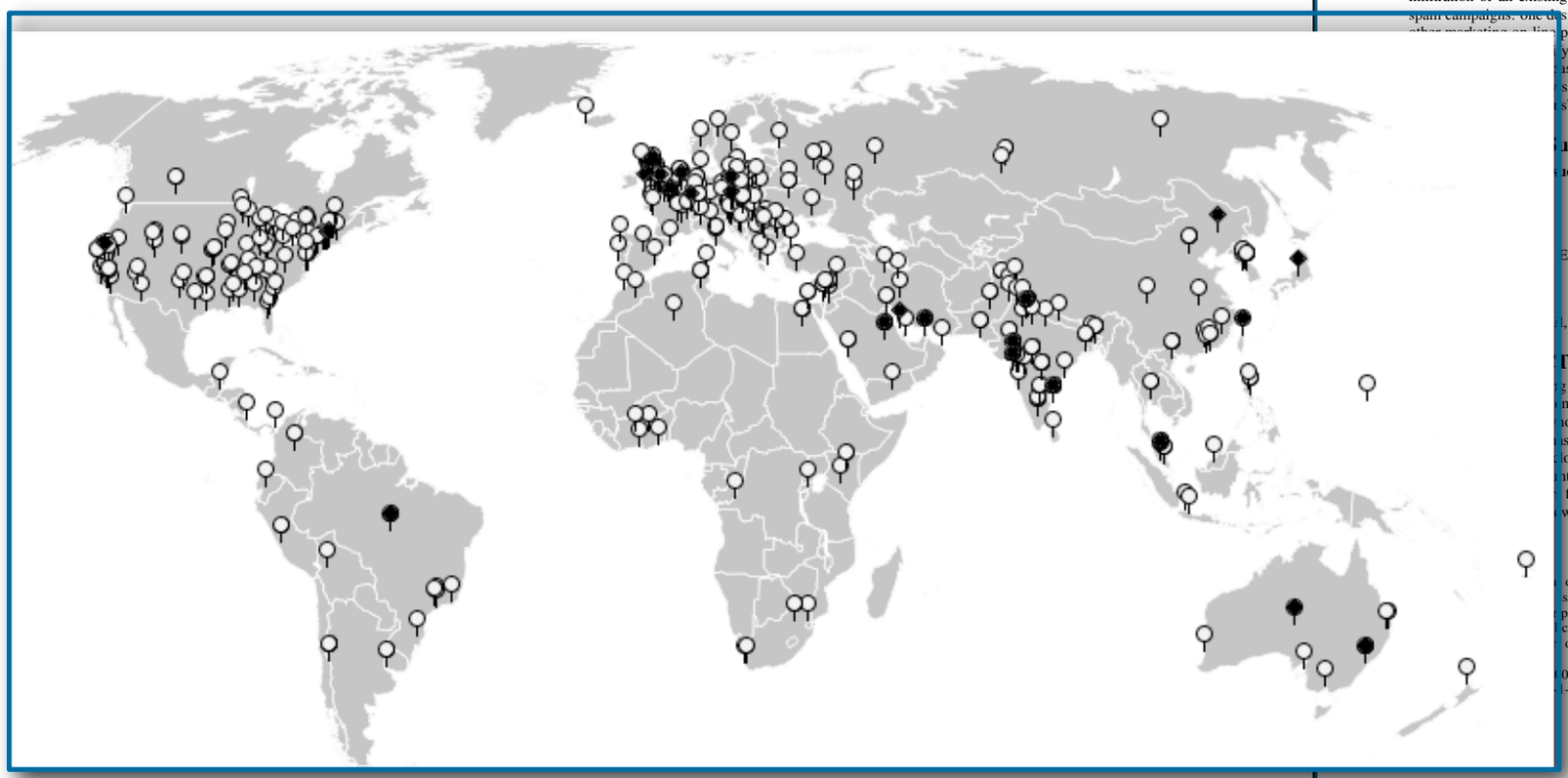
What are the economics of spam?

What is the response rate?

"Spamalytics: An Empirical Analysis of Spam Marketing Conversion," CCS 2008

Security Researchers:

- Infiltrated part of a botnet.
- Set up a fake online pharmacy.
- Redirected clicks for 469,906,992 spam messages.
- Converted 569 recipients!



Spamalytics: An Empirical Analysis of Spam Marketing Conversion

Chris Kanich* Christian Kreibich† Kirill Levchenko* Brandon Enright*
Geoffrey M. Voelker* Vern Paxson† Stefan Savage*

†International Computer Science Institute
Berkeley, USA
christian@icir.org,vern@cs.berkeley.edu

*Dept. of Computer Science and Engineering
University of California, San Diego, USA
{ckanich,klevchen,voelker,savage}@cs.ucsd.edu
bmenrigh@ucsd.edu

ABSTRACT

The "conversion rate" of spam — the probability that an unsolicited e-mail will ultimately elicit a "sale" — underlies the entire spam value proposition. However, our understanding of this critical behavior is quite limited, and the literature lacks any quantitative study concerning its true value. In this paper we present a methodology for measuring the conversion rate of spam. Using a parasitic infiltration of an existing botnet's infrastructure, we analyze two spam campaigns: one designed to propagate a malware Trojan, the other marketing on-line pharmaceuticals. For nearly a half billion ...y the number that are successfully deliv- ...ass through popular anti-spam filters, the ...sits to the advertised sites, and the number ...s" produced.

...ubject Descriptors

...ies]: ABUSE AND CRIME INVOLVING

...Economics

..., Conversion

...TION

...g is a curious beast. We all receive the ad- ...nt hardness is easy!" — but few of us have ...o admits to following through on this of- ...se. And yet, the relentlessness by which ...logs Internet inboxes, despite years of en- ...nti-spam technology, provides undeniable ... find their campaigns profitable. Someone ...w many, how often, and how much?

... 08, Alexandria, Virginia, USA.
...-59593-810-7/08/10 ...$5.00.

Unraveling such questions is *essential* for understanding the economic support for spam and hence where any structural weaknesses may lie. Unfortunately, spammers do not file quarterly financial reports, and the underground nature of their activities makes third-party data gathering a challenge at best. Absent an empirical foundation, defenders are often left to speculate as to how successful spam campaigns are and to what degree they are profitable. For example, IBM's Joshua Corman was widely quoted as claiming that spam sent by the Storm worm alone was generating "millions and millions of dollars every day" [2]. While this claim could in fact be true, we are unaware of any public data or methodology capable of confirming or refuting it.

The key problem is our limited visibility into the three basic parameters of the spam value proposition: the cost to send spam, offset by the "conversion rate" (probability that an e-mail sent will ultimately yield a "sale"), and the marginal profit per sale. The first and last of these are self-contained and can at least be estimated based on the costs charged by third-party spam senders and through the pricing and gross margins offered by various Internet marketing "affiliate programs".[1] However, the conversion rate depends fundamentally on group actions — on what hundreds of millions of Internet users do when confronted with a new piece of spam — and is much harder to obtain. While a range of anecdotal numbers exist, we are unaware of any well-documented measurement of the spam conversion rate.[2]

In part, this problem is methodological. There are no apparent methods for indirectly measuring spam conversion. Thus, the only obvious way to extract this data is to build an e-commerce site, market it via spam, and then record the number of sales. Moreover, to capture the spammer's experience with full fidelity, such a study must also mimic their use of illicit botnets for distributing e-mail and proxying user responses. In effect, the best way to measure spam is to be a spammer.

In this paper, we have effectively conducted this study, though *sidestepping* the obvious legal and ethical problems associated with sending spam.[3] Critically, our study makes use of an *existing* spam-

[1] Our cursory investigations suggest that commissions on pharmaceutical affiliate programs tend to hover around 40-50%, while the *retail* cost for spam delivery has been estimated at under $80 per million [22].

[2] The best known among these anecdotal figures comes from the Wall Street Journal's 2003 investigation of Howard Carmack (a.k.a the "Buffalo Spammer"), revealing that he obtained a 0.00036 conversion rate on ten million messages marketing an herbal stimulant [4].

[3] We conducted our study under the ethical criteria of ensuring *neutral actions* so that users should never be worse off due to our ac-

How can we make better spam filters? How about "personal keywords?"

Inputs:

- **Father name: Marvin**
- **Town: Pacific Grove**

Rules:

```
body  my_father /Marvin/
score my_father -55.0
body  my_town /Pacific Grove/i
score my_town -55.0
```

Protocol:

- **Run this on my inbox and see how well it works.**
- **Post ideas to a mailing list and get other people's experiences.**

What's wrong here?

45 CFR 46 never anticipated email.

Protocol:

- Run this on my inbox and see how well it works.
- Post ideas to a mailing list and get other people's experiences.

Problems:

- **I am experimenting on people who send me mail.**
- **Most email is not a public document.**
- **Senders did not give consent to be involved in my research.**
- **Under 45 CFR 46, I need IRB approval for this experiment.**

True Story: “Send me your file system statistics.”

A PhD student at a major university sent email to a list:

- **“Please download this program and run it.”**
- **The program collected *file system statistics* and *sends it back for analysis.***

I sent mail to the student:

- **“Did you get IRB approval?”**

Student response:

- **“What’s an IRB?”**

True Story: Check for updates.

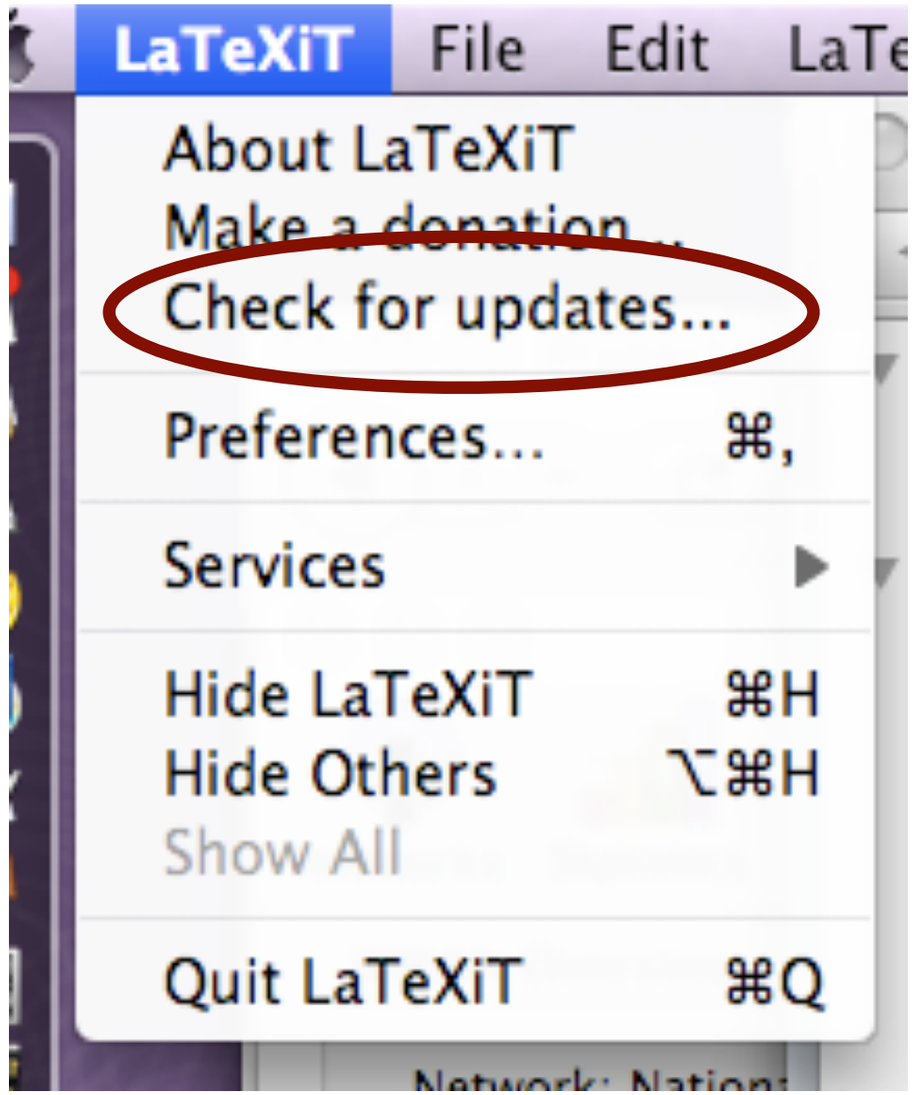
Software that checks for updates "phones home."

Information collected:

- Version number of client.
- IP address of client
- When the client is run

Question for IRB Chairs:

- Does the release of a program which checks for updates require IRB approval?



Computer Science and IRBs: on collision course.

Most Computer Scientists:

- **Don't know what an IRB is.**
- **Don't want IRB oversight.**
- **Do not have the training to formulate a "protocol" in advance.**
- **Don't know what they are going to find — are doing exploratory work**
  - —Just like ethnographers
- **Don't know if they are experimenting on humans!**

Conflict in the IRB rules:

- **Experimenters are not allowed to decide if research is exempt.**
- **But experimenters *are* allowed to decide if research involves humans!**

Computer Science and IRBs: on collision course.

Most IRBs:

- **Can't turn around an exempt or minimal risk protocol in 1-2 business days.**
  - —Some can't do it in 2 months!
  - —This is a real problem for student research.
- **Don't have much computer science knowledge.**
  - —The details matter a lot!
- **Some don't think that they even have jurisdiction.**

NIH was unwilling to issue a CoC to protect highly confidential experimental subject data.

- **"This project doesn't seem to fit within the mission of any of the Public Health Service agencies that have been delegated this authority."**

Recommendations

IRBs need dramatically streamlined review processes.

- CS students *routinely* work with human subject data for class projects.
- Approve “Minimal Risk” studies with self-certification via web.

IRBs need better understanding of computer science.

- What data collection is standard/appropriate/excessive?
- Can information be "anonymized." ?
- *DOES IT MATTER IF INFORMATION IS ANONYMIZED?*

PIs need to do a better job crafting IRB applications.

- Broadly written applications to avoid needless reviews.

45 CFR 46 needs updating.

- Address issue of retroactive approval and "pilot studies."