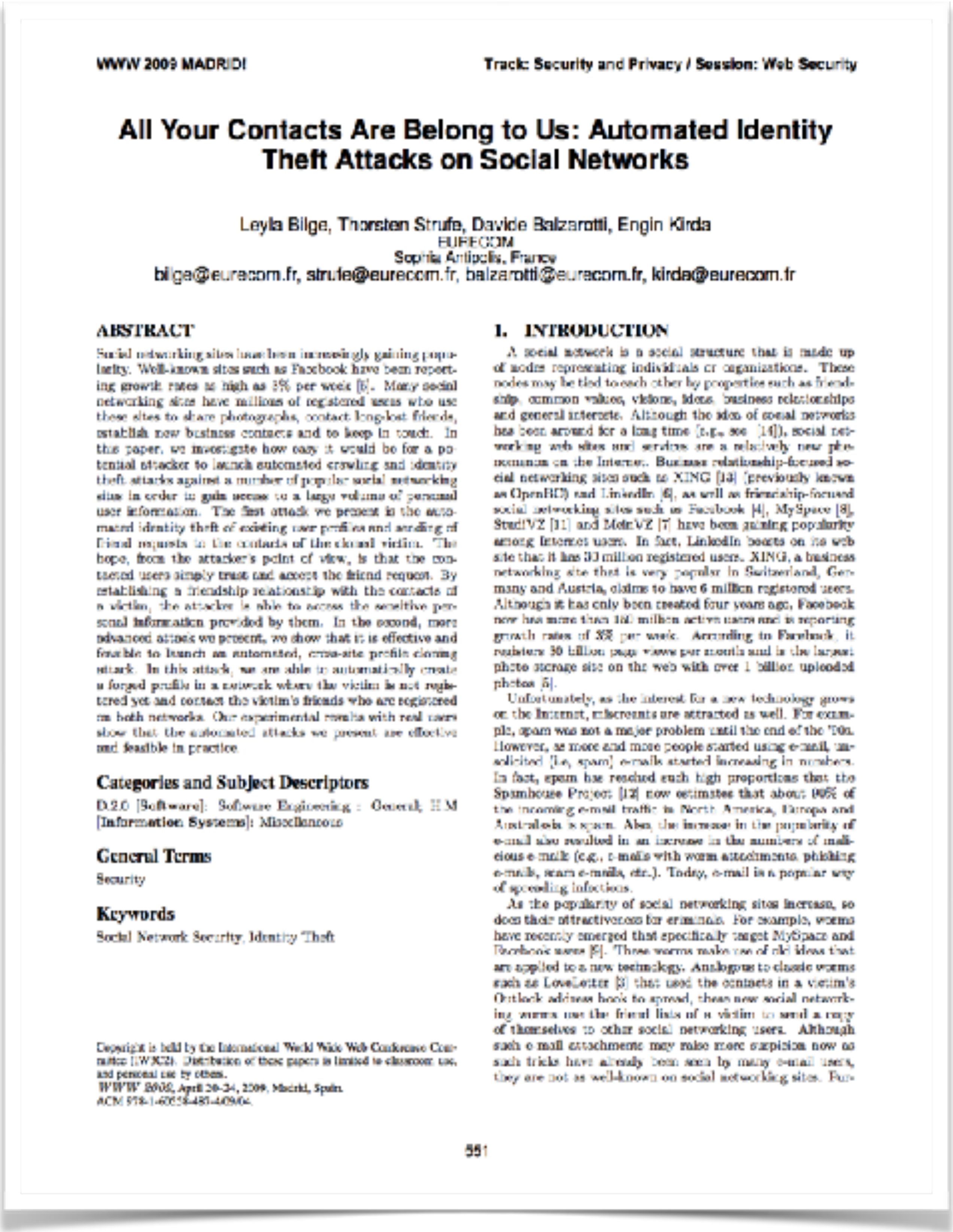
ETHICS
IN SECURITY RESEARCH

GRAD SEC

SEP 05 2017

TODAY'S PAPERS



All Your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Networks

Leyla Bilge, Thorsten Strufe, Davide Balzarotti, Engin Kirda
EURECOM
Sophia Antipolis, France
bilge@eurecom.fr, strufe@eurecom.fr, balzarotti@eurecom.fr, kirda@eurecom.fr

ABSTRACT

Social networking sites have been increasingly gaining popularity. Well-known sites such as Facebook have been reporting growth rates as high as 3% per week [5]. Many social networking sites have millions of registered users who use these sites to share photographs, contact long-lost friends, establish new business contacts and to keep in touch. In this paper, we investigate how easy it would be for a potential attacker to launch automated crawling and identity theft attacks against a number of popular social networking sites in order to gain access to a large volume of personal user information. The first attack we present is the automated identity theft of existing user profiles and sending of friend requests to the contacts of the cloned victim. The hope, from the attacker's point of view, is that the contacted users simply trust and accept the friend request. By establishing a friendship relationship with the contacts of a victim, the attacker is able to access the sensitive personal information provided by them. In the second, more advanced attack we present, we show that it is effective and feasible to launch an automated, cross-site profile cloning attack. In this attack, we are able to automatically create a forged profile in a network where the victim is not registered yet and contact the victim's friends who are registered on both networks. Our experimental results with real users show that the automated attacks we present are effective and feasible in practice.

Categories and Subject Descriptors

D.2.0 [Software]: Software Engineering : General; H.M [Information Systems]: Miscellaneous

General Terms

Security

Keywords

Social Network Security, Identity Theft

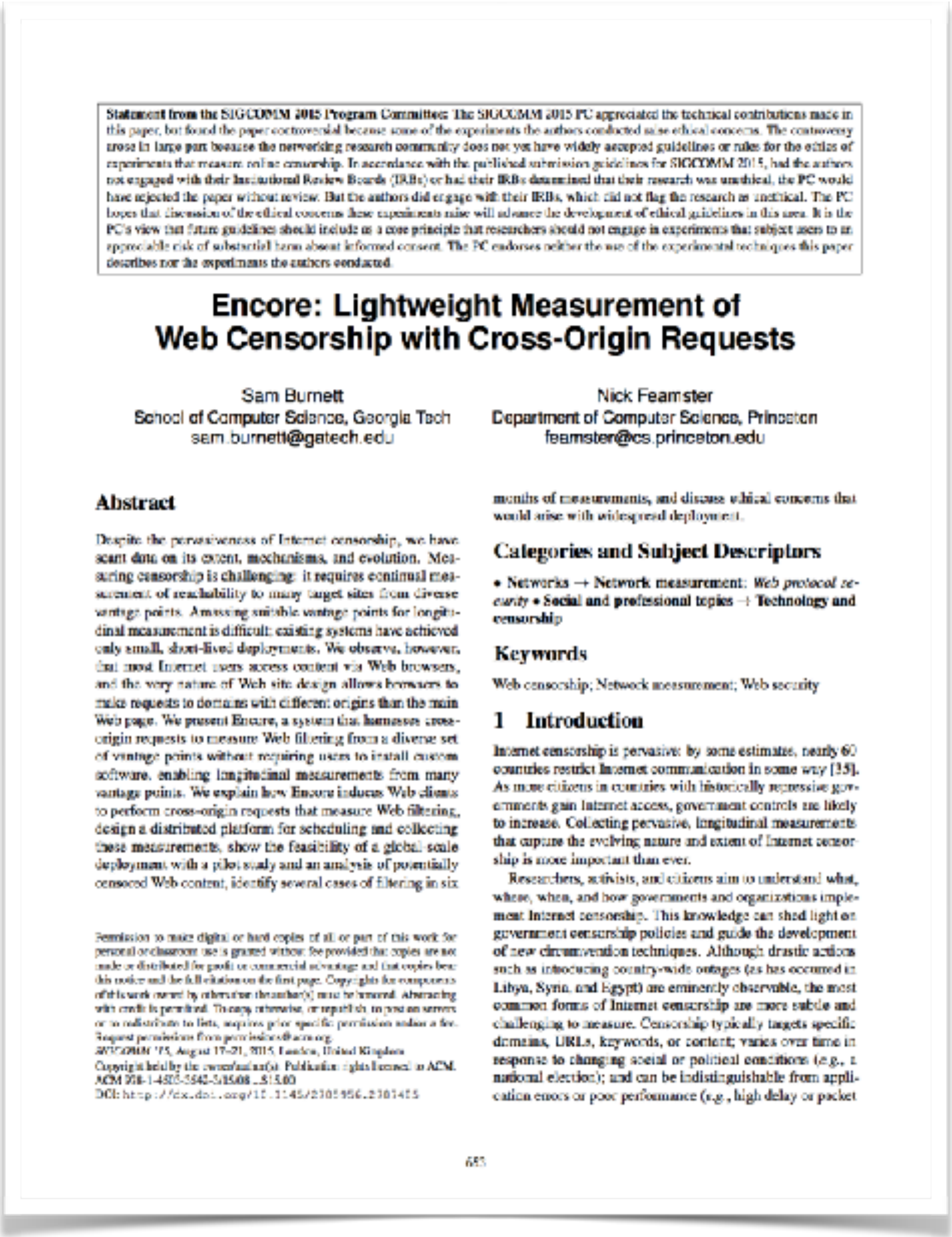
WWW 2009, April 20–24, 2009, Madrid, Spain.
ACM 978-1-60558-487-4/09/04.

1. INTRODUCTION

A social network is a social structure that is made up of nodes representing individuals or organizations. These nodes may be tied to each other by properties such as friendship, common values, visions, ideas, business relationships and general interests. Although the idea of social networks has been around for a long time (e.g., see [14]), social networking web sites and services are a relatively new phenomenon on the Internet. Business relationship-focused social networking sites such as XING [13] (previously known as OpenBC) and LinkedIn [6], as well as friendship-focused social networking sites such as Facebook [4], MySpace [8], StudiVZ [11] and MeinVZ [7] have been gaining popularity among Internet users. In fact, LinkedIn boasts on its web site that it has 30 million registered users. XING, a business networking site that is very popular in Switzerland, Germany and Austria, claims to have 6 million registered users. Although it has only been created four years ago, Facebook now has more than 160 million active users and is reporting growth rates of 3% per week. According to Facebook, it registers 30 billion page views per month and is the largest photo storage site on the web with over 1 billion uploaded photos [5].

Unfortunately, as the interest for a new technology grows on the Internet, miscreants are attracted as well. For example, spam was not a major problem until the end of the '90s. However, as more and more people started using e-mail, unsolicited (i.e., spam) e-mails started increasing in numbers. In fact, spam has reached such high proportions that the Spamhouse Project [12] now estimates that about 90% of the incoming e-mail traffic in North America, Europe and Australasia is spam. Also, the increase in the popularity of e-mail also resulted in an increase in the numbers of malicious e-mails (e.g., e-mails with worm attachments, phishing e-mails, scam e-mails, etc.). Today, e-mail is a popular way of spreading infections.

As the popularity of social networking sites increases, so does their attractiveness for criminals. For example, worms have recently emerged that specifically target MySpace and Facebook users [5]. These worms make use of old ideas that are applied to a new technology. Analogous to classic worms such as LoveLetter [3] that used the contacts in a victim's Outlook address book to spread, these new social networking worms use the friend lists of a victim to send a copy of themselves to other social networking users. Although such e-mail attachments may raise more suspicion now as such tricks have already been seen by many e-mail users, they are not as well-known on social networking sites. Fur-



Statement from the SIGCOMM 2015 Program Committee: The SIGCOMM 2015 PC appreciated the technical contributions made in this paper, but found the paper controversial because some of the experiments the authors conducted raise ethical concerns. The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship. In accordance with the published submission guidelines for SIGCOMM 2015, had the authors not engaged with their Institutional Review Boards (IRBs) or had their IRBs determined that their research was unethical, the PC would have rejected the paper without review. But the authors did engage with their IRBs, which did not flag the research as unethical. The PC hopes that discussion of the ethical concerns these experiments raise will advance the development of ethical guidelines in this area. It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent. The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted.

Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests

Sam Burnett
School of Computer Science, Georgia Tech
sam.burnett@gatech.edu

Nick Feamster
Department of Computer Science, Princeton
feamster@cs.princeton.edu

Abstract

Despite the pervasiveness of Internet censorship, we have scant data on its extent, mechanisms, and evolution. Measuring censorship is challenging: it requires continual measurement of reachability to many target sites from diverse vantage points. Amassing suitable vantage points for longitudinal measurement is difficult; existing systems have achieved only small, short-lived deployments. We observe, however, that most Internet users access content via Web browsers, and the very nature of Web site design allows browsers to make requests to domains with different origins than the main Web page. We present Encore, a system that harnesses cross-origin requests to measure Web filtering from a diverse set of vantage points without requiring users to install custom software, enabling longitudinal measurements from many vantage points. We explain how Encore induces Web clients to perform cross-origin requests that measure Web filtering, design a distributed platform for scheduling and collecting these measurements, show the feasibility of a global-scale deployment with a pilot study and an analysis of potentially censored Web content, identify several cases of filtering in six months of measurements, and discuss ethical concerns that would arise with widespread deployment.

SIGCOMM '15, August 17–21, 2015, London, United Kingdom

ACM 978-1-4503-3542-3/15/08 ...$15.00
DOI: http://dx.doi.org/10.1145/2785956.2787485

Categories and Subject Descriptors

• **Networks → Network measurement**: *Web protocol security* • **Social and professional topics → Technology and censorship**

Keywords

Web censorship; Network measurement; Web security

1 Introduction

Internet censorship is pervasive: by some estimates, nearly 60 countries restrict Internet communication in some way [35]. As more citizens in countries with historically repressive governments gain Internet access, government controls are likely to increase. Collecting pervasive, longitudinal measurements that capture the evolving nature and extent of Internet censorship is more important than ever.

Researchers, activists, and citizens aim to understand what, where, when, and how governments and organizations implement Internet censorship. This knowledge can shed light on government censorship policies and guide the development of new circumvention techniques. Although drastic actions such as introducing country-wide outages (as has occurred in Libya, Syria, and Egypt) are eminently observable, the most common forms of Internet censorship are more subtle and challenging to measure. Censorship typically targets specific domains, URLs, keywords, or content; varies over time in response to changing social or political conditions (e.g., a national election); and can be indistinguishable from application errors or poor performance (e.g., high delay or packet

HUMAN SUBJECT RESEARCH

There is a history of harmful experimentation

TUSKEGEE SYPHILIS STUDY
(1932–1972)

Goal: Study the effects of untreated syphilis

Subjects: African-American men in rural Alabama, told they were receiving free health care

MILGRAM EXPERIMENT
(1961)

Goal: Understand “obedience to authority”

Effect: Extreme emotional stress on participants

STANFORD PRISON EXPERIMENT
(1971)

Goal: Understand

Effect: “Guards” psychologically abused “prisoners”

BELMONT REPORT

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research

Lays out the ethical principles by which to guide research

RESPECT FOR PERSONS

Participation is voluntary; they get to decide

Participation follows from **informed consent**

Protect those who are incapable of deciding

BENEFICENCE

Do no harm

Maximize probable benefits; minimize probable harm

JUSTICE

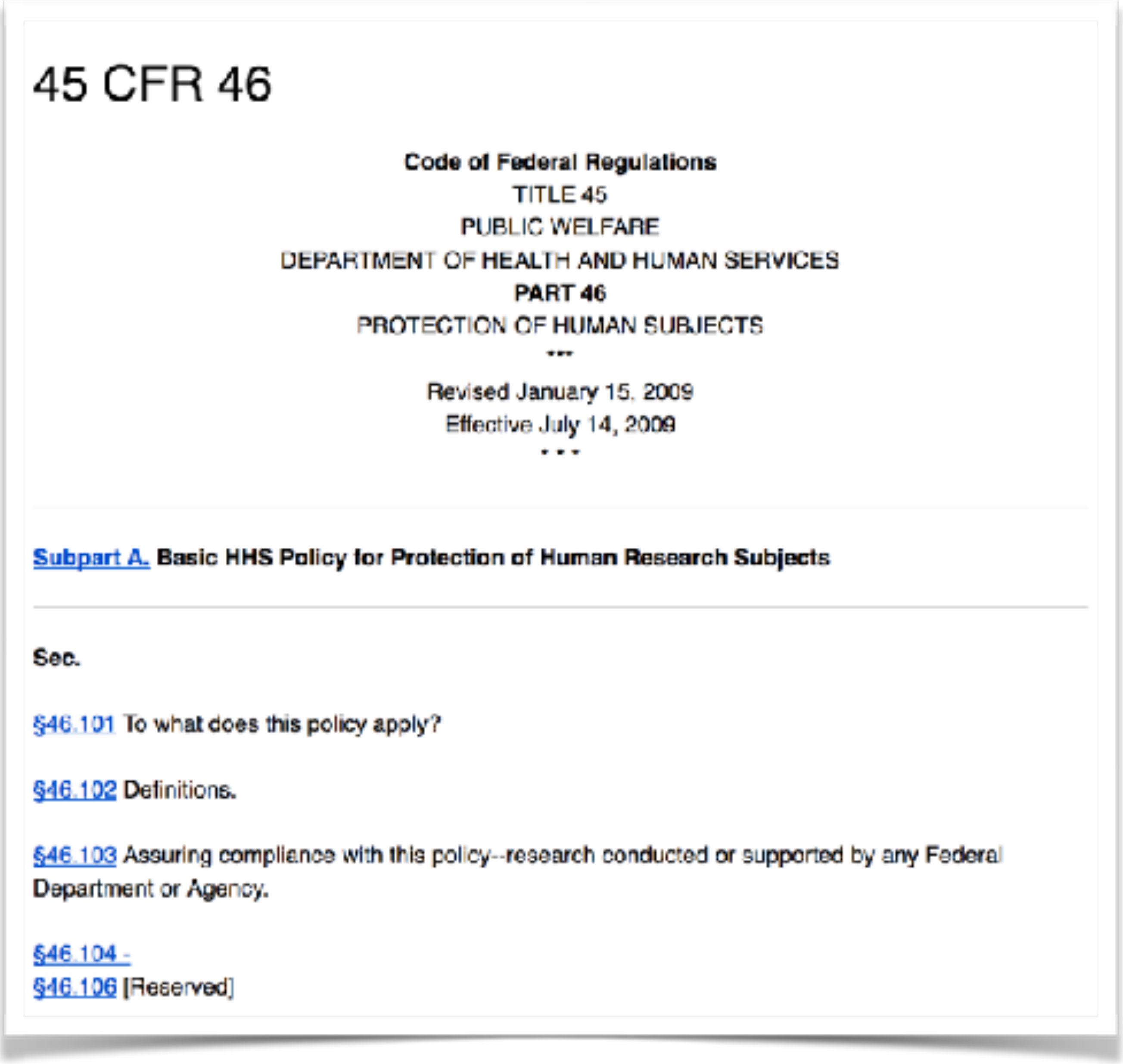
Treat all people fairly

The benefits of research should be fairly distributed

Menlo report adds: Engage in legal due diligence; be transparent

INSTITUTIONAL REVIEW BOARD (IRB)

Codified by the Health and Human Services (HHS)



45 CFR 46

Code of Federal Regulations
TITLE 45
PUBLIC WELFARE
DEPARTMENT OF HEALTH AND HUMAN SERVICES
PART 46
PROTECTION OF HUMAN SUBJECTS
***
Revised January 15, 2009
Effective July 14, 2009
* * *

Subpart A. Basic HHS Policy for Protection of Human Research Subjects

Sec.

§46.101 To what does this policy apply?

§46.102 Definitions.

§46.103 Assuring compliance with this policy--research conducted or supported by any Federal Department or Agency.

§46.104 -
§46.106 [Reserved]

An institution's IRB seeks to protect subjects

- reviews research plans
- may require modifications
- may disapprove
- can expedite research involving "no more than minimal risk

DEFINITIONS

HUMAN SUBJECT RESEARCH

Research on living individuals that collects:

- Data through interaction or intervention
- Identifiable private information

INFORMED CONSENT

Written

Participation is always voluntary

MINIMAL RISK

The probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests

EXEMPT REVIEW

Research that does not require IRB review

Very specific categories

Example: passively observing public records

CRITERIA FOR IRB APPROVAL

Risks to subjects are minimized
by using procedures consistent with sound research design

Risks to subjects are reasonable
in relation to anticipated benefits

Selection of subjects is equitable
taking into account the purpose of the research

Informed consent is sought
taking into account the purpose of the research

Data kept safe and private

BASIC ELEMENTS OF INFORMED CONSENT

Research statement

“This is research”, purposes, duration, and the procedure

Statement that participation is voluntary

taking into account the purpose of the research

Description of any foreseeable risks and benefits

to the subject or others

Exception: Some research cannot be done with informed consent

No more than minimal risk

Waiver will not adversely affect the subjects’ rights & welfare

When appropriate, subjects will be informed

LEGAL ASPECTS OF SECURITY RESEARCH

Conducting Cybersecurity Research Legally and Ethically [Burnstein]

WIRETAP ACT

Prohibits real-time interception of ***contents*** of electronic communication

Contains a "provider exception"

Distinction unclear

PEN/TRAP STATUTE

Prohibits real-time interception of the ***non-content*** portions of electronic communications

Contains a "provider exception"

STORED COMMUNICATIONS ACT (SCA)

Prohibits providers of "electronic communications service to the public" from knowingly disclosing the contents of customers' communication

None of these contain research exceptions (unlike, say, HIPAA)

QUESTIONS

How do we get informed consent?
Are banners enough?

How do we share data once we have collected it?
Is data anonymization enough?

Even if we legally *can*, who is to say we *should*?
Whom on campus should we ask?

Can we run infected hosts?
Can we stop attackers' computers?

RUNNING INFECTED HOSTS

Likely largely covered by Computer Fraud and Abuse Act (CFAA)

CFAA

Prohibits conduct directed against virtually any computer

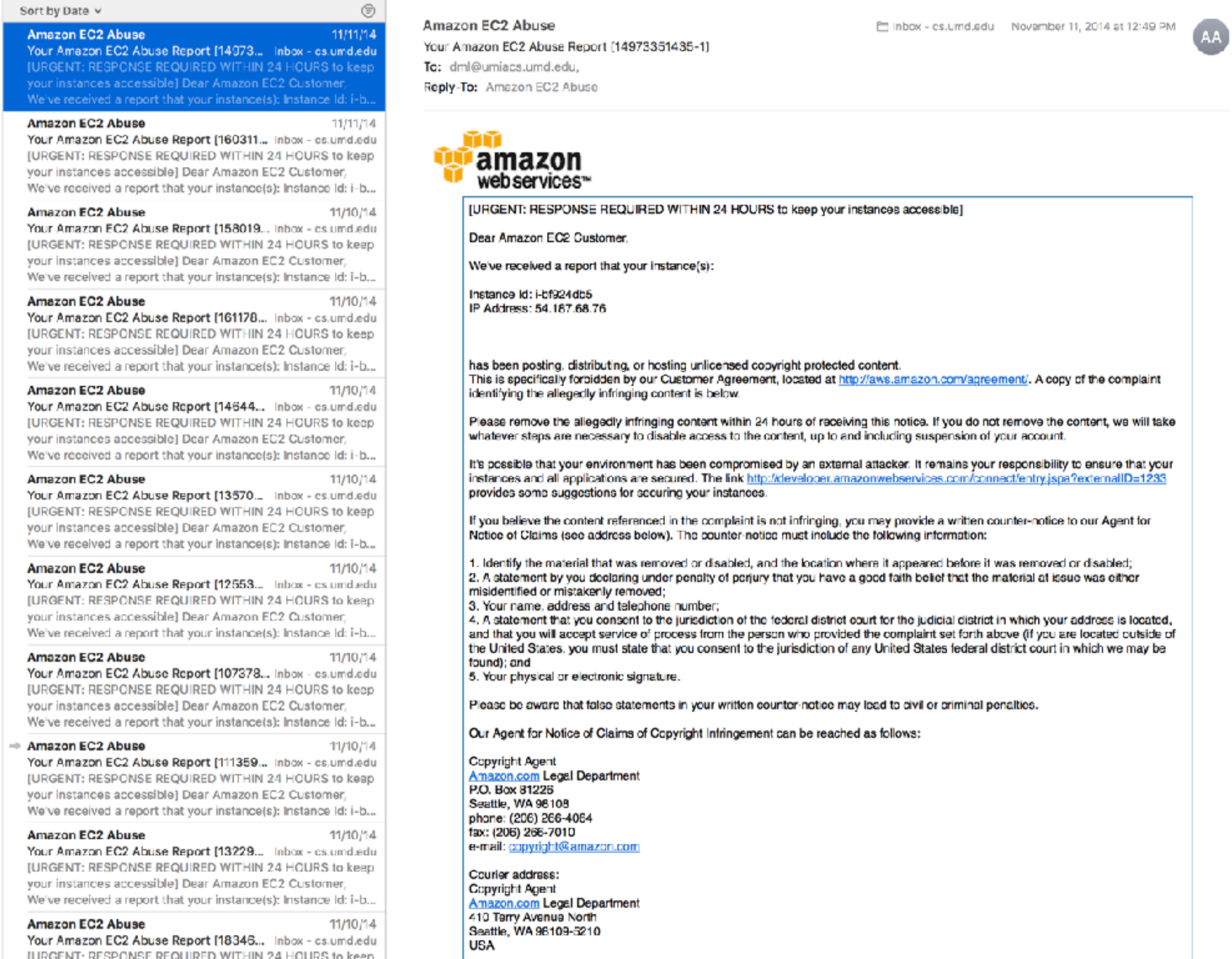
> Prohibits "[K]nowingly caus[ing] the transmission of a program, information, code, or command, and as a result of such conduct, intentionally caus[ing] damage without authorization, to a protected computer."

Can we let computers join a botnet?

Can we take down an infected host on the Internet?

Are researchers liable for storing copyrighted or illegal data?

"Federal law makes it a crime to knowingly possess any image of child pornography".



Amazon EC2 Abuse 11/11/14
Your Amazon EC2 Abuse Report [14973... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/11/14
Your Amazon EC2 Abuse Report [160311... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [158019... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [161178... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [14644... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [13570... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [12553... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [107378... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [111359... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [13229... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible] Dear Amazon EC2 Customer, We've received a report that your instance(s): Instance Id: i-b...

Amazon EC2 Abuse 11/10/14
Your Amazon EC2 Abuse Report [18346... Inbox - cs.umd.edu
[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep

Amazon EC2 Abuse Inbox - cs.umd.edu November 11, 2014 at 12:49 PM
Your Amazon EC2 Abuse Report [14973351435-1]
To: dml@umiacs.umd.edu,
Reply-To: Amazon EC2 Abuse

[URGENT: RESPONSE REQUIRED WITHIN 24 HOURS to keep your instances accessible]

Dear Amazon EC2 Customer,

We've received a report that your instance(s):

Instance Id: i-bf924db5
IP Address: 54.167.68.76

has been posting, distributing, or hosting unlicensed copyright protected content.
This is specifically forbidden by our Customer Agreement, located at http://aws.amazon.com/agreement/. A copy of the complaint identifying the allegedly infringing content is below.

Please remove the allegedly infringing content within 24 hours of receiving this notice. If you do not remove the content, we will take whatever steps are necessary to disable access to the content, up to and including suspension of your account.

It's possible that your environment has been compromised by an external attacker. It remains your responsibility to ensure that your instances and all applications are secured. The link http://developer.amazonwebservices.com/connect/entry.jspa?externalID=1233 provides some suggestions for securing your instances.

If you believe the content referenced in the complaint is not infringing, you may provide a written counter-notice to our Agent for Notice of Claims (see address below). The counter-notice must include the following information:

1. Identify the material that was removed or disabled, and the location where it appeared before it was removed or disabled;
2. A statement by you declaring under penalty of perjury that you have a good faith belief that the material at issue was either misidentified or mistakenly removed;
3. Your name, address and telephone number;
4. A statement that you consent to the jurisdiction of the federal district court for the judicial district in which your address is located, and that you will accept service of process from the person who provided the complaint set forth above (if you are located outside of the United States, you must state that you consent to the jurisdiction of any United States federal district court in which we may be found); and
5. Your physical or electronic signature.

Please be aware that false statements in your written counter-notice may lead to civil or criminal penalties.

Our Agent for Notice of Claims of Copyright Infringement can be reached as follows:

Copyright Agent
Amazon.com Legal Department
P.O. Box 81226
Seattle, WA 98108
phone: (206) 266-4064
fax: (206) 266-7010
e-mail: copyright@amazon.com

Courier address:
Copyright Agent
Amazon.com Legal Department
410 Terry Avenue North
Seattle, WA 98109-5210
USA

LEGAL ASPECTS OF SECURITY RESEARCH

Carefully design your experiments

Consult experts and campus representatives

Be transparent and react quickly

Let someone know if and when something happens

If it's for this class, talk to me first!

I can help create a safe environment to work in

RECAP: ALL YOUR CONTACTS

QUESTION	Do users reciprocate friend requests?
CONTEXT	Sybil attacks in social networks

CONTEXT: SYBIL ATTACKS IN SOCIAL NETWORKS

SYBIL ATTACKS Multiple illegitimate identities, or “Sybils”

[Douceur’02] Today, sometimes referred to as “bots”

SOCIAL NETWORKS Attacker’s goal: Influence

Upvotes, comments, reviews, public opinion, …

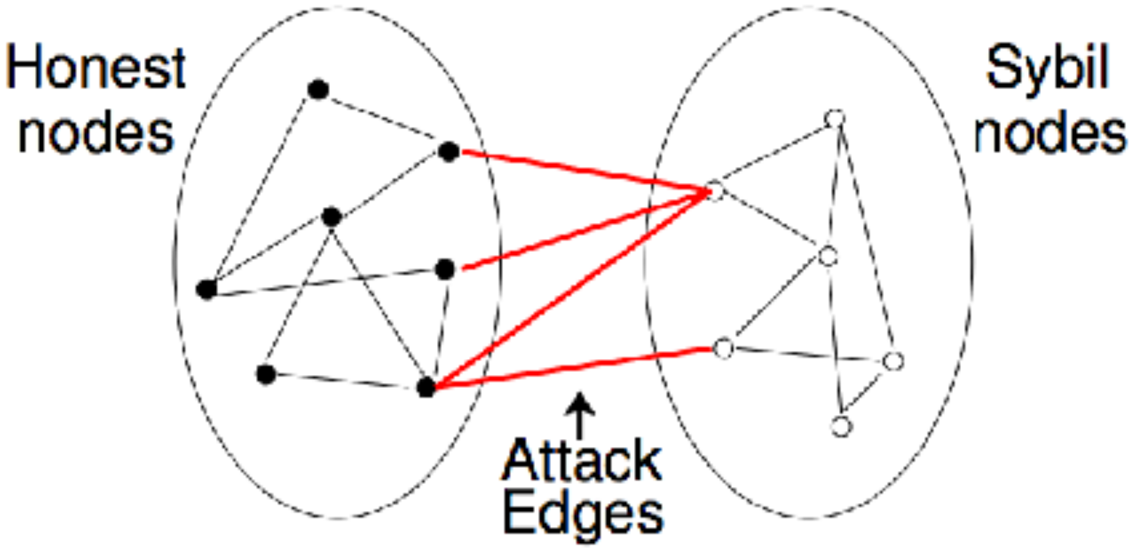
Papers to mitigate Sybils

SybilGuard SybilLimit SybilInfer

SoK: The Evolution of Sybil Defense via Social Networks

Make a common assumption

ATTACK EDGES



Basic idea: Use network flow to constrain how much influence any subgraph has on the rest of the graph

Common assumption:

“Attack edges” hard to make

If users reciprocate friend requests, then this is a bad assumption

ATTACKS

PROFILE CLONING

"many users will not get suspicious if a friend request comes from someone they know, even if this person is already on their contact list."

CROSS-SITE PROFILE CLONING

1. **identify victims** who are registered in one social network, but not in another
2. **steal their identities:** create accounts for them in the network where they are not registered
3. **friend the friends** of the victim in the original network who are registered in the target network

How do you launch these attacks at scale?

HOW DO YOU LAUNCH THESE ATTACKS AT SCALE?

CAPTCHA SOLVERS

Completely Automated Turing test to tell Computers and Humans Apart

Primary line of defense against automated acct creation

WEBSITE CRAWLER

- Pull identities off of a website
- Compare two different identities across websites
  - Cross-site cloning: identify friends

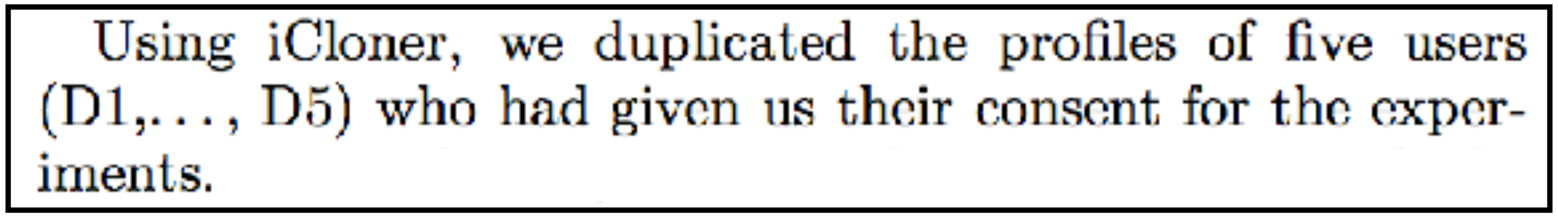
The barrier to scalability is low

HOW DO YOU EVALUATE THESE ATTACKS?

REAL USERS

705 in total

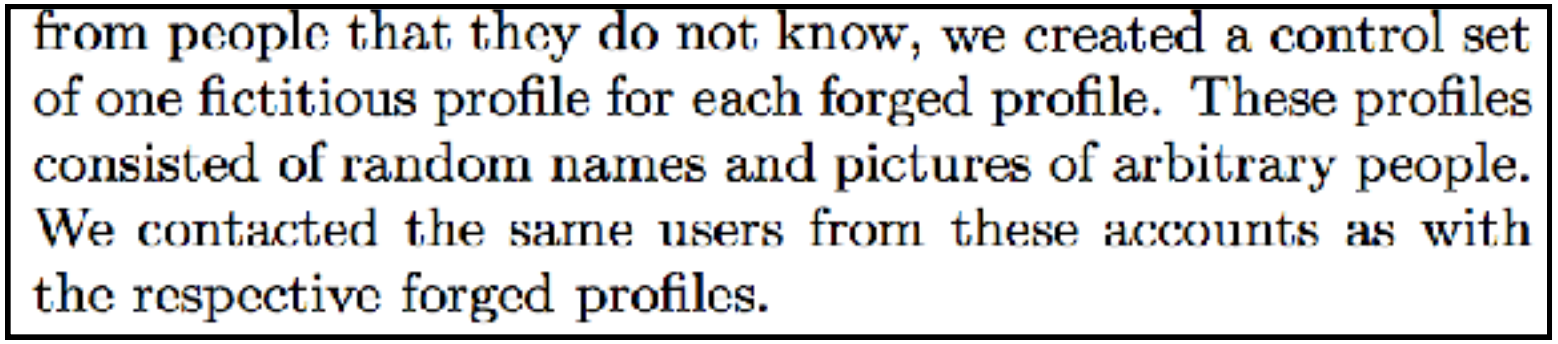
Unsuspecting; on real sites



Using iCloner, we duplicated the profiles of five users (D1,. . . , D5) who had given us their consent for the experiments.

CONTROL

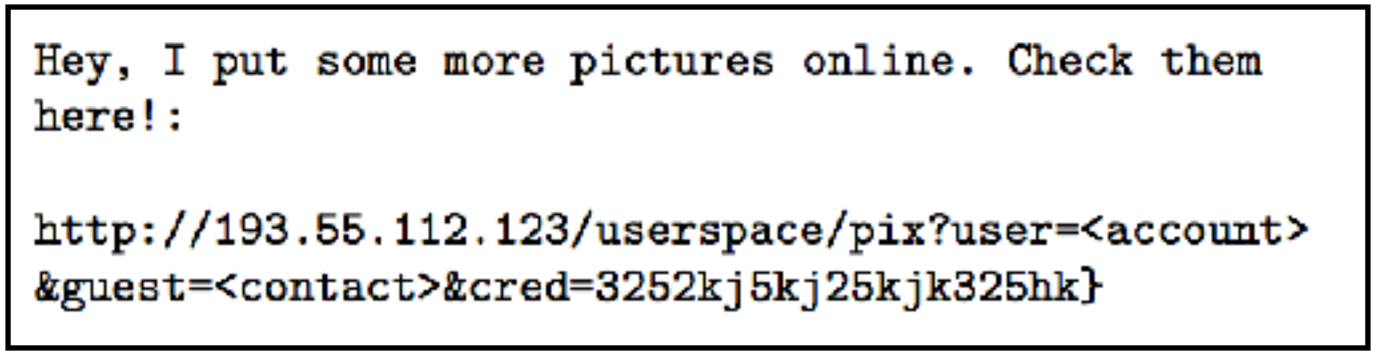
Are they reciprocating because of the attacks or do they just reciprocate everything?



from people that they do not know, we created a control set of one fictitious profile for each forged profile. These profiles consisted of random names and pictures of arbitrary people. We contacted the same users from these accounts as with the respective forged profiles.

MEANING

Does reciprocation imply trust?



```
Hey, I put some more pictures online. Check them
here!:

http://193.55.112.123/userspace/pix?user=<account>
&guest=<contact>&cred=3252kj5kj25kjk325hk}
```

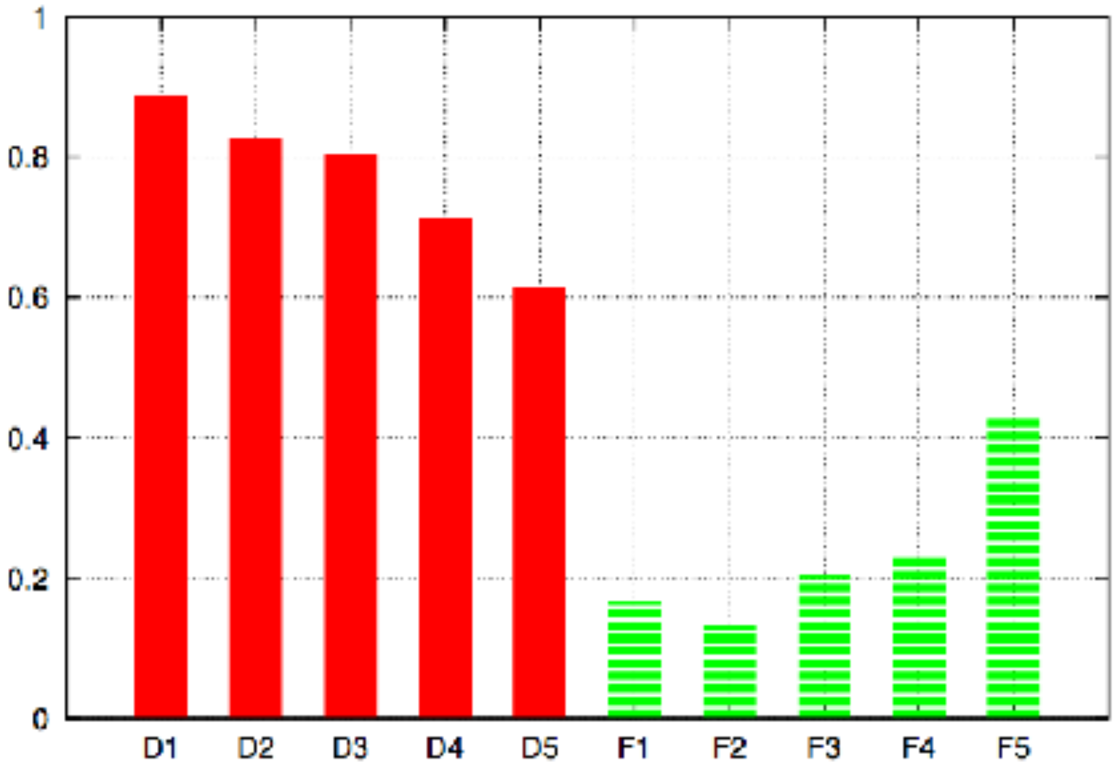
RESULTS

FORGED

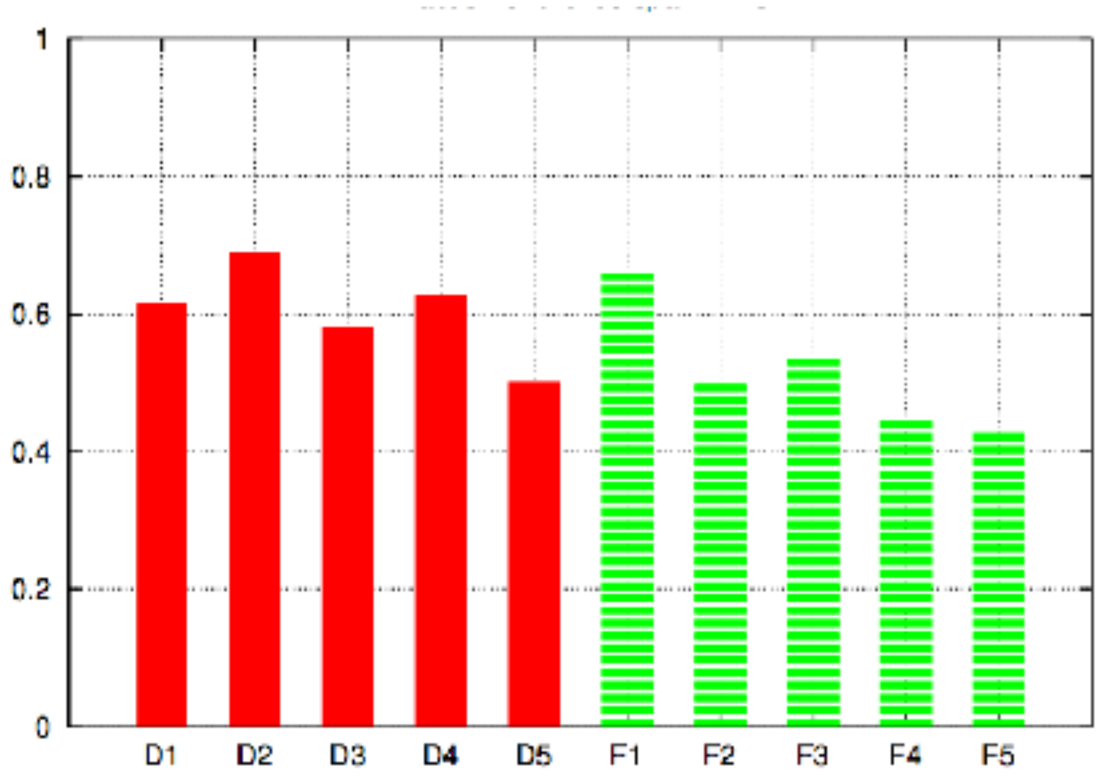
Profile cloning

FICTITIOUS

Control



Reciprocation rates



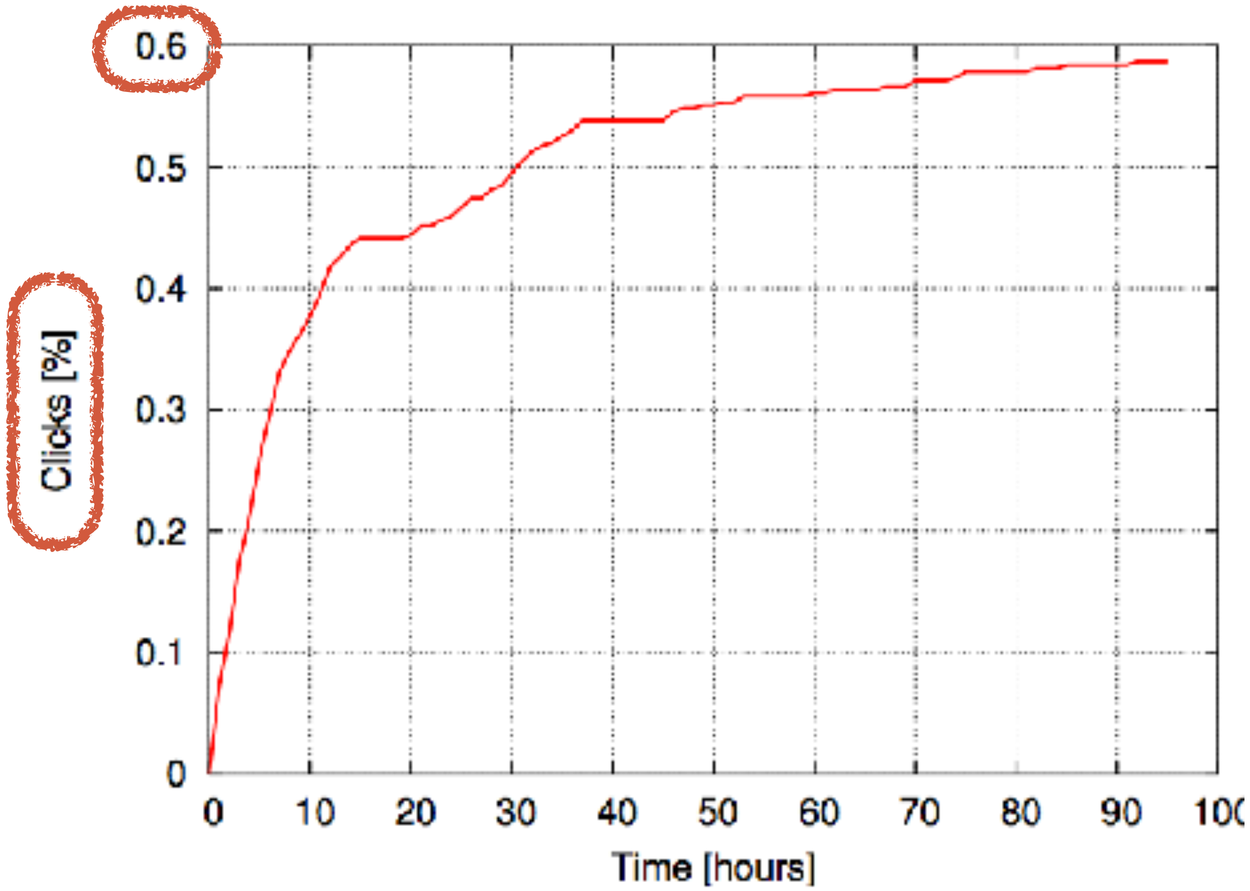
Click-through rates

How would the numbers compare on LinkedIn? Twitter?

RESULTS

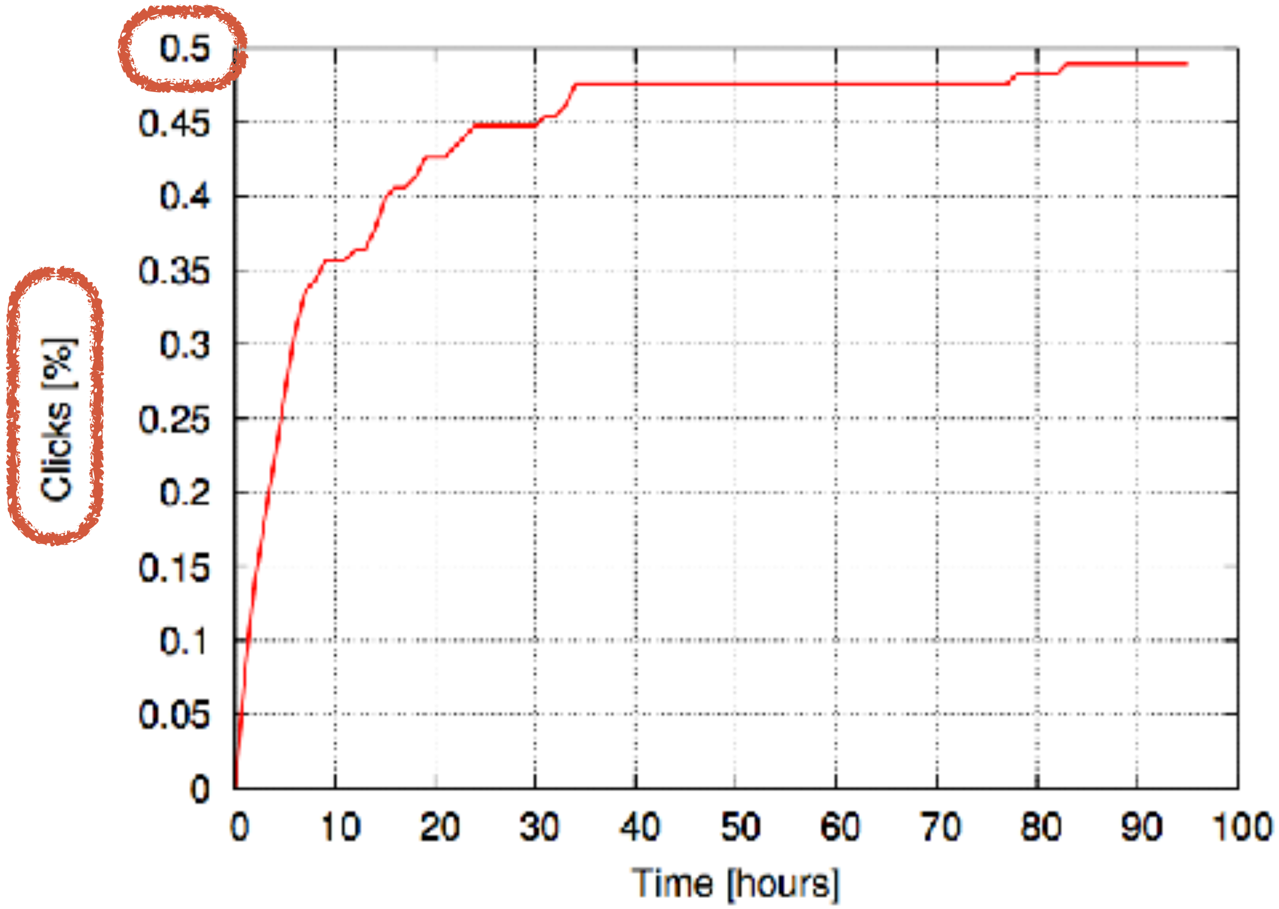
FORGED

Profile cloning



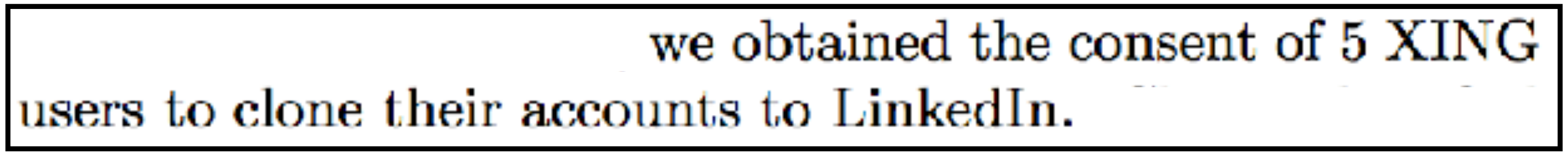
FICTITIOUS

Control



How would the numbers compare on LinkedIn? Twitter?

RESULTS

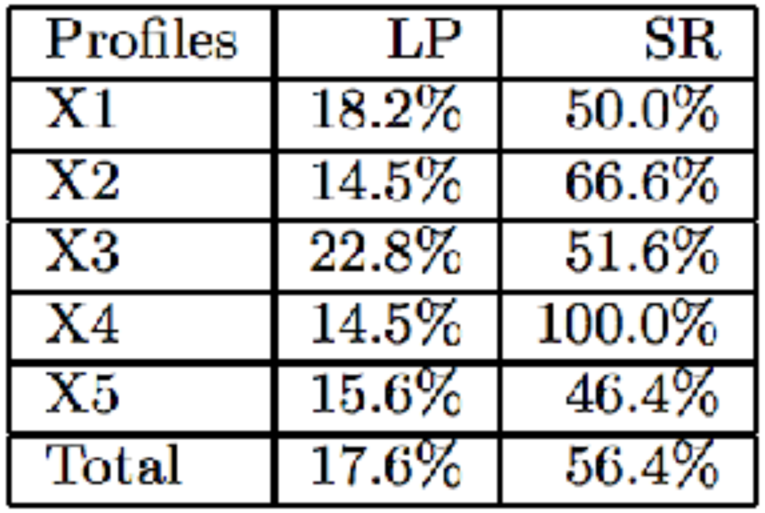


we obtained the consent of 5 XING users to clone their accounts to LinkedIn.

CROSS-SITE PROFILE CLONING

5 XING users **consented** to having their accounts cloned on LinkedIn

443 XING user "victims"



Profiles	LP	SR
X1	18.2%	50.0%
X2	14.5%	66.6%
X3	22.8%	51.6%
X4	14.5%	100.0%
X5	15.6%	46.4%
Total	17.6%	56.4%

Table 1: Percentage of XING profiles found in LinkedIn (LP) and the success rate (SR) of the contact requests

Facebook

How would the numbers compare on ~~LinkedIn? Twitter~~?

Does not permit a head-to-head comparison of attacks

WAS “ALL YOUR CONTACTS” DONE ETHICALLY?

HUMAN SUBJECT RESEARCH

Research on living individuals that collects:

- Data through interaction or intervention
- Identifiable private information

INFORMED CONSENT

Written

Participation is always voluntary

MINIMAL RISK

The probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests

EXEMPT REVIEW

Research that does not require IRB review

Very specific categories

Example: passively observing public records

SOME OF YOUR THOUGHTS

Pouya *there is no mention that the friends of the victims who have received friend requests from fake profiles, were notified regarding the study or not*

there is no mention of IRB being approved of this study or not

Benjamin *The authors of this paper make little effort to obscure the techniques used; they describe the CAPTCHA breaking software in particular detail*

Jo *However, I do not think it is a publisher/conference's responsibility to deny papers with ethical problems of this kind, within reason. As far as I'm concerned,* ***by the time this paper was submitted for review the damage was already done****. It was the responsibility of the researchers' IRBs to safeguard potential experimental participants' privacy before the experiment even got approved, right?*

Nirat *with a new discovered attack, there must be more insights/techniques about how to counter such attacks.*

RECAP: ENCORE

QUESTION How are users being censored worldwide?

CONTEXT Censorship measurement tools

CENSORSHIP MEASUREMENT TOOLS

How do you measure the prevalence of online censorship?

SURVEY Ask users what is being blocked

VANTAGE POINTS Deploy measurement nodes in users’ homes

REMOTE MEASUREMENT Probe networks and infer

NEWS Wait to hear about big outage events

None of these are: **Longitudinal and broad (in users & sites)**

CROSS-ORIGIN REQUESTS

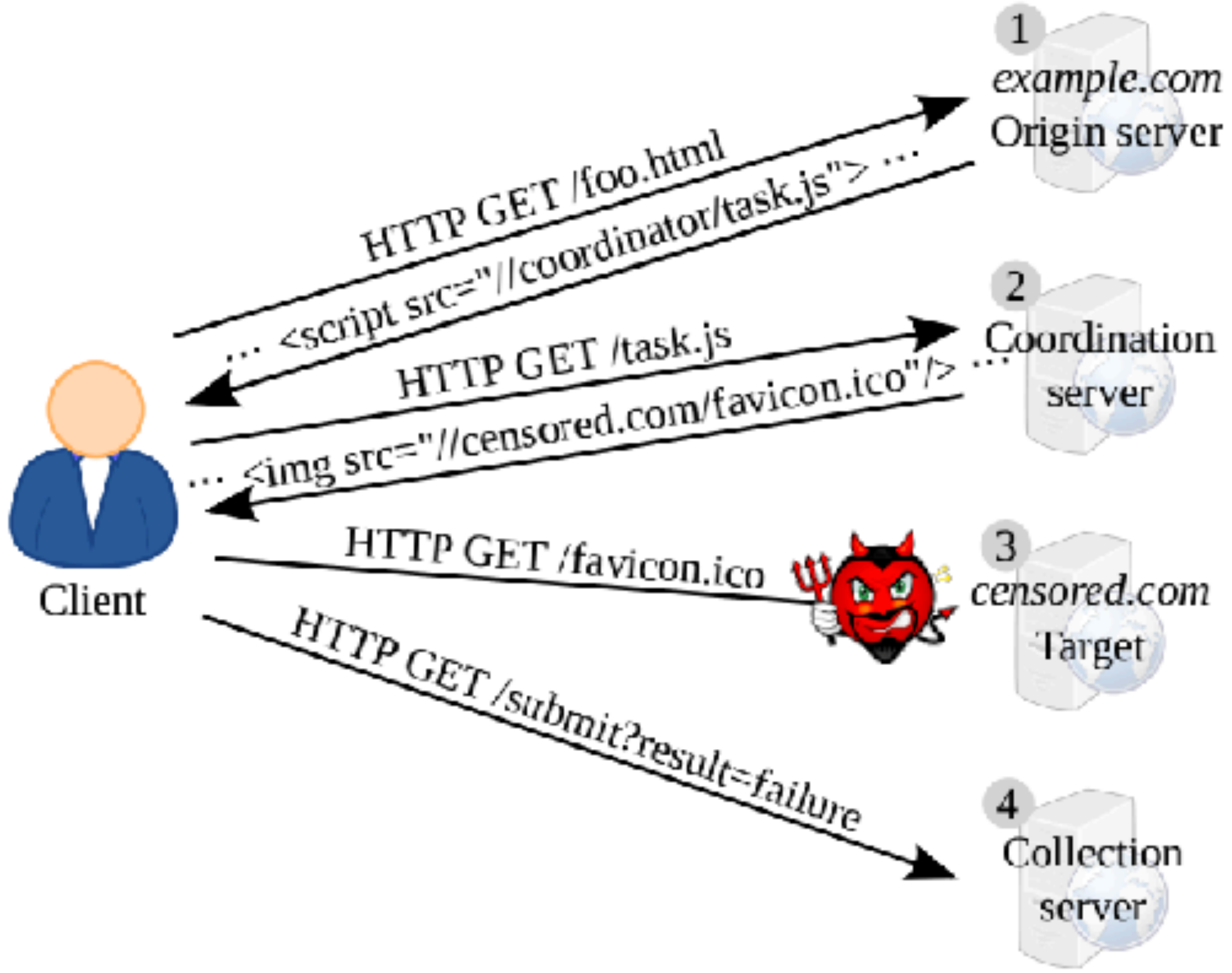
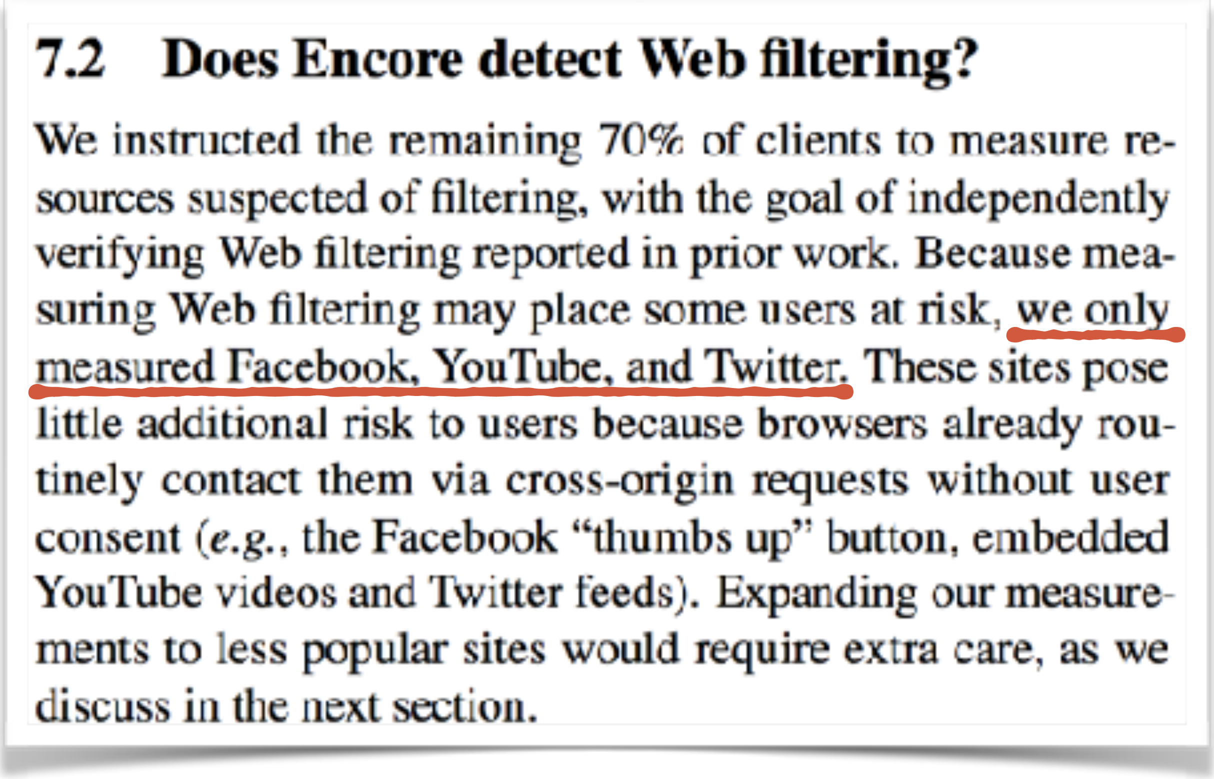


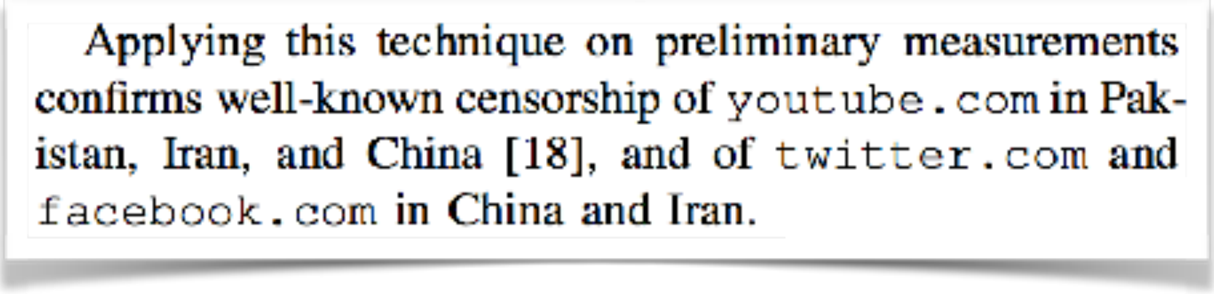
Figure 2: *An example of observing Web filtering with Encore. The origin Web page includes Encore's measurement script, which the coordinator decides should test filtering of* `censored.com` *by attempting to fetch an image. The request for this image fails so the client notifies the collection server.*

RESULTS



7.2 Does Encore detect Web filtering?

We instructed the remaining 70% of clients to measure resources suspected of filtering, with the goal of independently verifying Web filtering reported in prior work. Because measuring Web filtering may place some users at risk, we only measured Facebook, YouTube, and Twitter. These sites pose little additional risk to users because browsers already routinely contact them via cross-origin requests without user consent (*e.g.*, the Facebook "thumbs up" button, embedded YouTube videos and Twitter feeds). Expanding our measurements to less popular sites would require extra care, as we discuss in the next section.



Applying this technique on preliminary measurements confirms well-known censorship of `youtube.com` in Pakistan, Iran, and China [18], and of `twitter.com` and `facebook.com` in China and Iran.

WAS "ENCORE" DONE ETHICALLY?

HUMAN SUBJECT RESEARCH

Research on living individuals that collects:

- Data through interaction or intervention
- Identifiable private information

INFORMED CONSENT

Written

Participation is always voluntary

MINIMAL RISK

The probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests

EXEMPT REVIEW

Research that does not require IRB review

Very specific categories

Example: passively observing public records

WAS "ENCORE" DONE ETHICALLY?

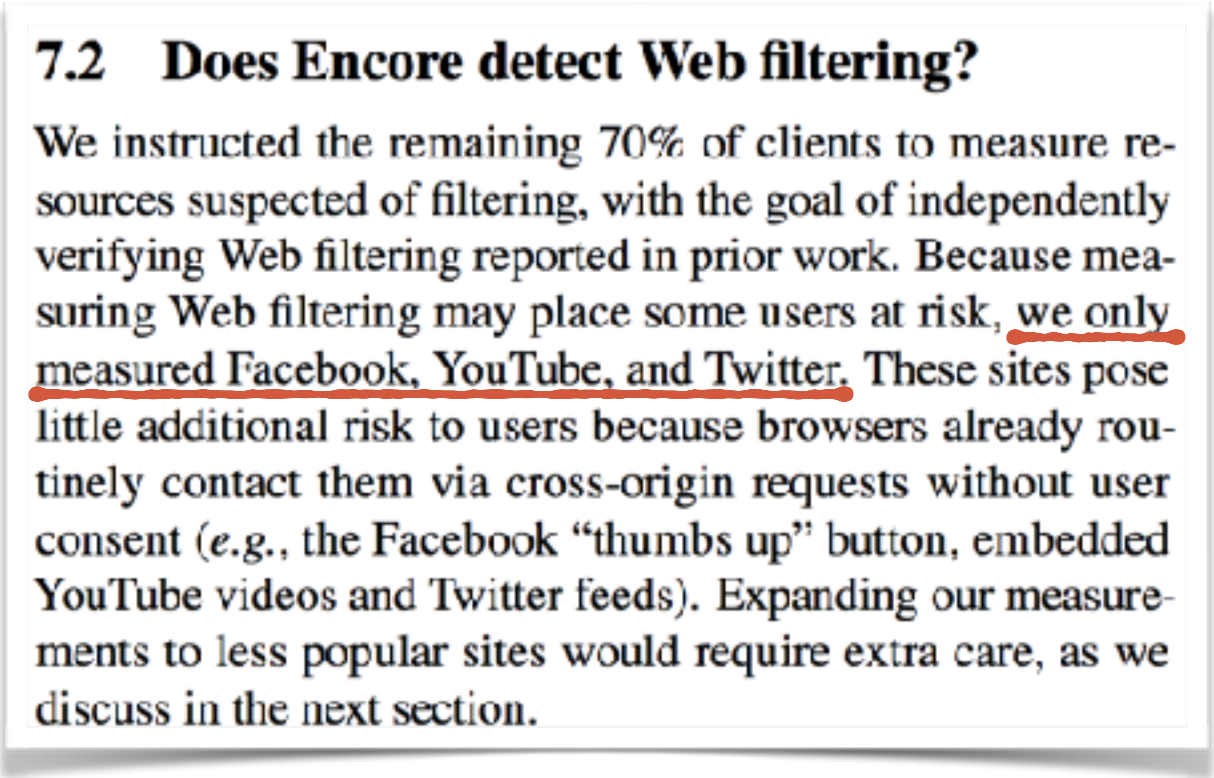
HUMAN SUBJECT RESEARCH

- "Informal discussions" indicated no human subjects
- Sought IRB approval at 2 institutions after rejections

INFORMED CONSENT

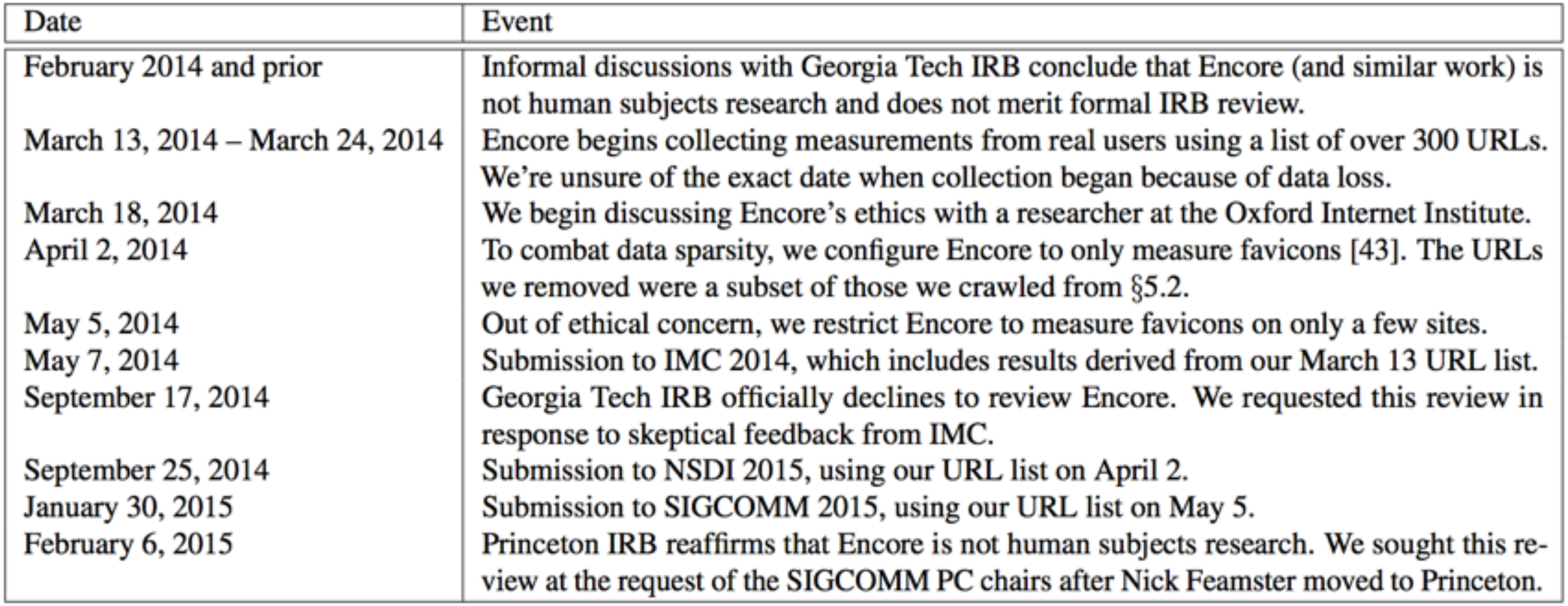
- "not always appropriate"
- "would require apprising a user about nuanced technical concepts"
- **Users would probably have not consented**
  "would dramatically reduce the scale and scope of measurements"

WHAT DID THEY MEASURE?



7.2 Does Encore detect Web filtering?

We instructed the remaining 70% of clients to measure resources suspected of filtering, with the goal of independently verifying Web filtering reported in prior work. Because measuring Web filtering may place some users at risk, we only measured Facebook, YouTube, and Twitter. These sites pose little additional risk to users because browsers already routinely contact them via cross-origin requests without user consent (*e.g.*, the Facebook "thumbs up" button, embedded YouTube videos and Twitter feeds). Expanding our measurements to less popular sites would require extra care, as we discuss in the next section.



Date	Event
February 2014 and prior	Informal discussions with Georgia Tech IRB conclude that Encore (and similar work) is not human subjects research and does not merit formal IRB review.
March 13, 2014 – March 24, 2014	Encore begins collecting measurements from real users using a list of over 300 URLs. We're unsure of the exact date when collection began because of data loss.
March 18, 2014	We begin discussing Encore's ethics with a researcher at the Oxford Internet Institute.
April 2, 2014	To combat data sparsity, we configure Encore to only measure favicons [43]. The URLs we removed were a subset of those we crawled from §5.2.
May 5, 2014	Out of ethical concern, we restrict Encore to measure favicons on only a few sites.
May 7, 2014	Submission to IMC 2014, which includes results derived from our March 13 URL list.
September 17, 2014	Georgia Tech IRB officially declines to review Encore. We requested this review in response to skeptical feedback from IMC.
September 25, 2014	Submission to NSDI 2015, using our URL list on April 2.
January 30, 2015	Submission to SIGCOMM 2015, using our URL list on May 5.
February 6, 2015	Princeton IRB reaffirms that Encore is not human subjects research. We sought this review at the request of the SIGCOMM PC chairs after Nick Feamster moved to Princeton.

WAS “ENCORE” DONE ETHICALLY?

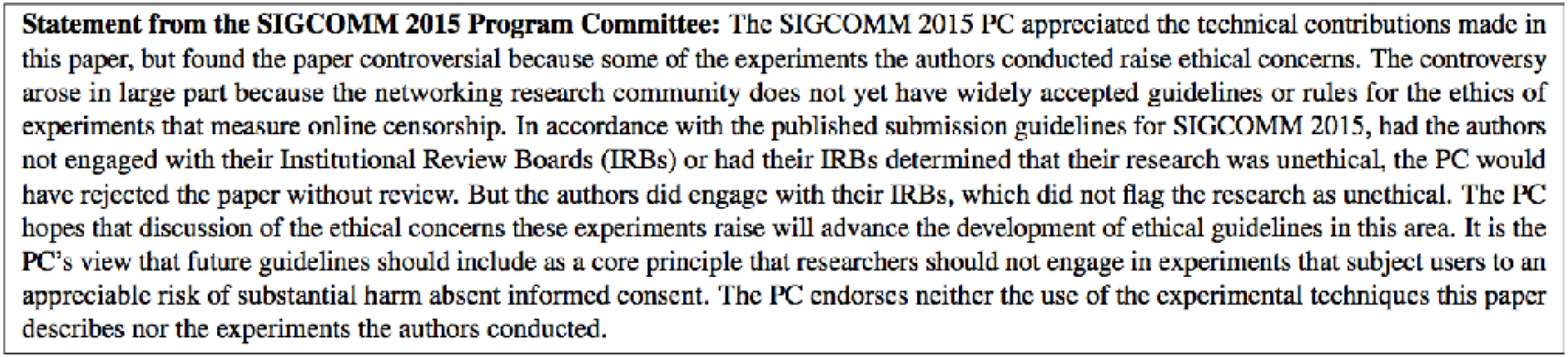
HUMAN SUBJECT RESEARCH

- “Informal discussions” indicated no human subjects
- Sought IRB approval at 2 institutions after rejections

INFORMED CONSENT

- “not always appropriate”
- “would require apprising a user about nuanced technical concepts”
- **Users would probably have not consented**
  “would dramatically reduce the scale and scope of measurements”
- “informed consent does not *ever* decrease risk to users; it only alleviates researchers from some responsibility for that risk, and may even increase risk to users by removing any traces of plausible deniability”

STATEMENT FROM THE SIGCOMM'15 PC



Statement from the SIGCOMM 2015 Program Committee: The SIGCOMM 2015 PC appreciated the technical contributions made in this paper, but found the paper controversial because some of the experiments the authors conducted raise ethical concerns. The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship. In accordance with the published submission guidelines for SIGCOMM 2015, had the authors not engaged with their Institutional Review Boards (IRBs) or had their IRBs determined that their research was unethical, the PC would have rejected the paper without review. But the authors did engage with their IRBs, which did not flag the research as unethical. The PC hopes that discussion of the ethical concerns these experiments raise will advance the development of ethical guidelines in this area. It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent. The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted.

…the networking research community does not yet have **widely accepted guidelines or rules** for the ethics of experiments that measure online censorship…

…had the authors not engaged with their IRBs,
or had their IRBs determined that their research was unethical,
the PC would have rejected the paper without review.

The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted.

SOME OF YOUR THOUGHTS

Brook *Political climates may change, or investigative agencies may decide to crack down on illegal information. The authors can't predict with enough reliability that nobody will be harmed by this, especially since it applies to people all over the world.*

Richard *I take issue with the fact that the entire source code for Encore was made public.*