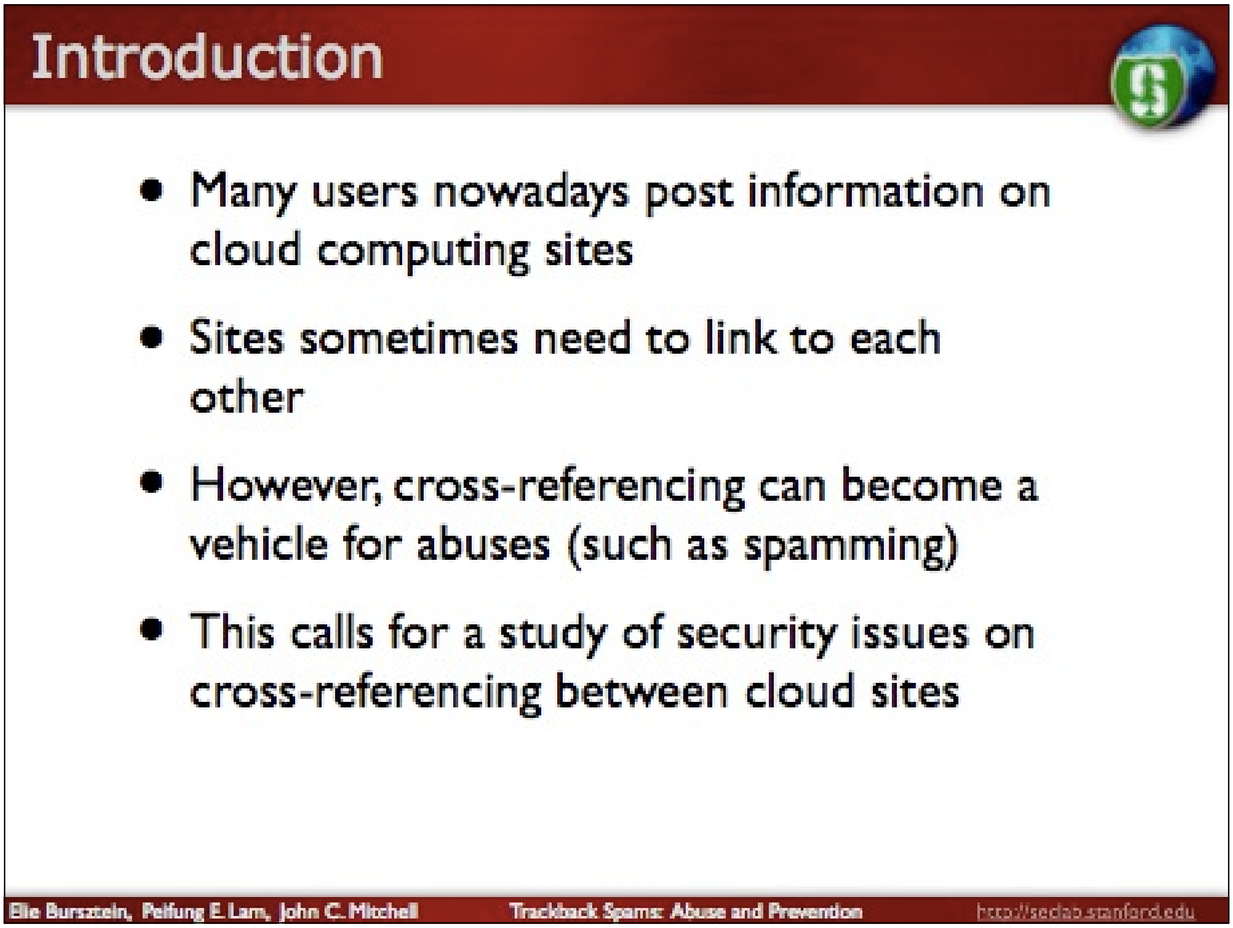
# TrackBack Spam: Abuse and Prevention

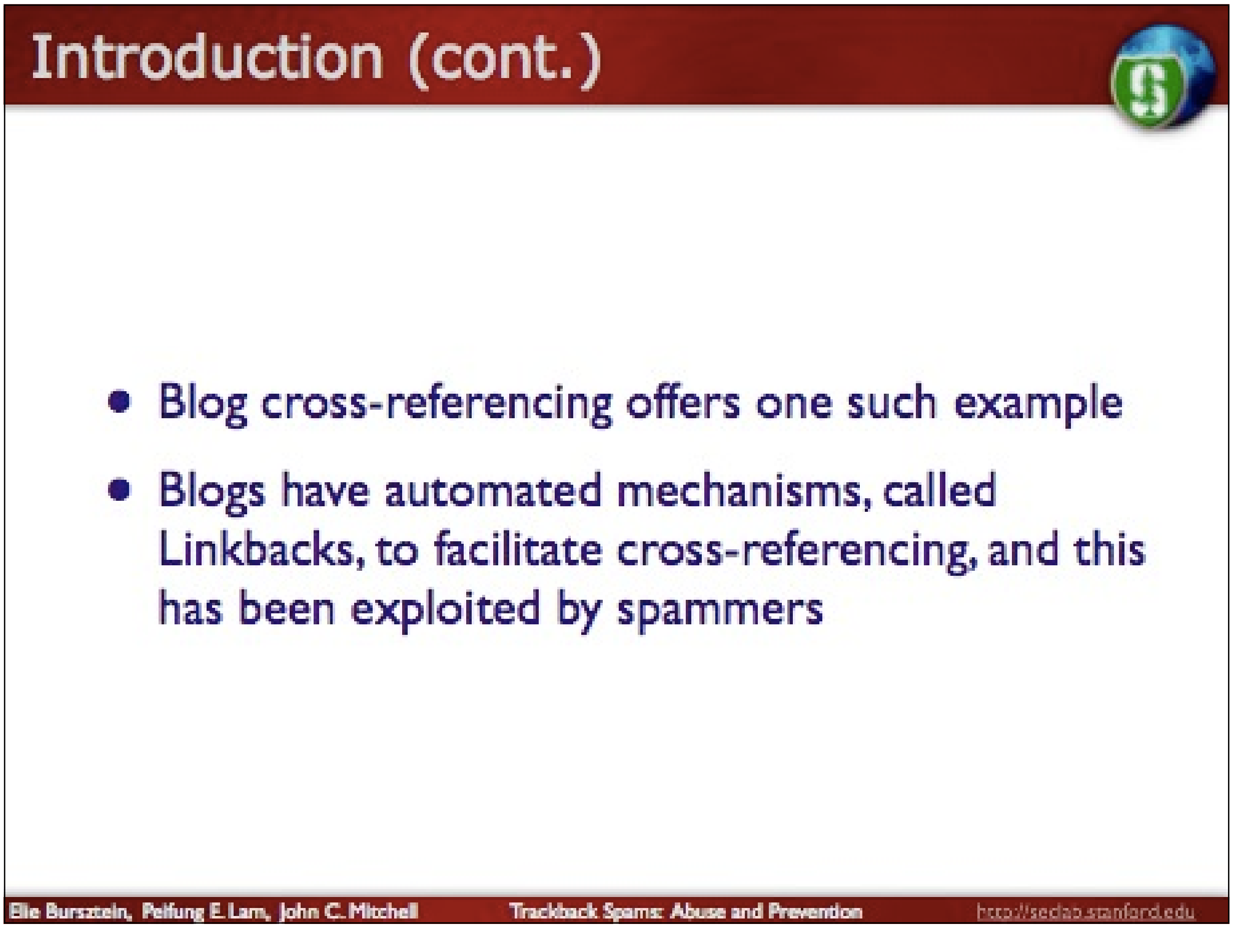
Elie Bursztein, **Peifung E. Lam**, John C. Mitchell
Stanford University

Stanford Computer Security Lab



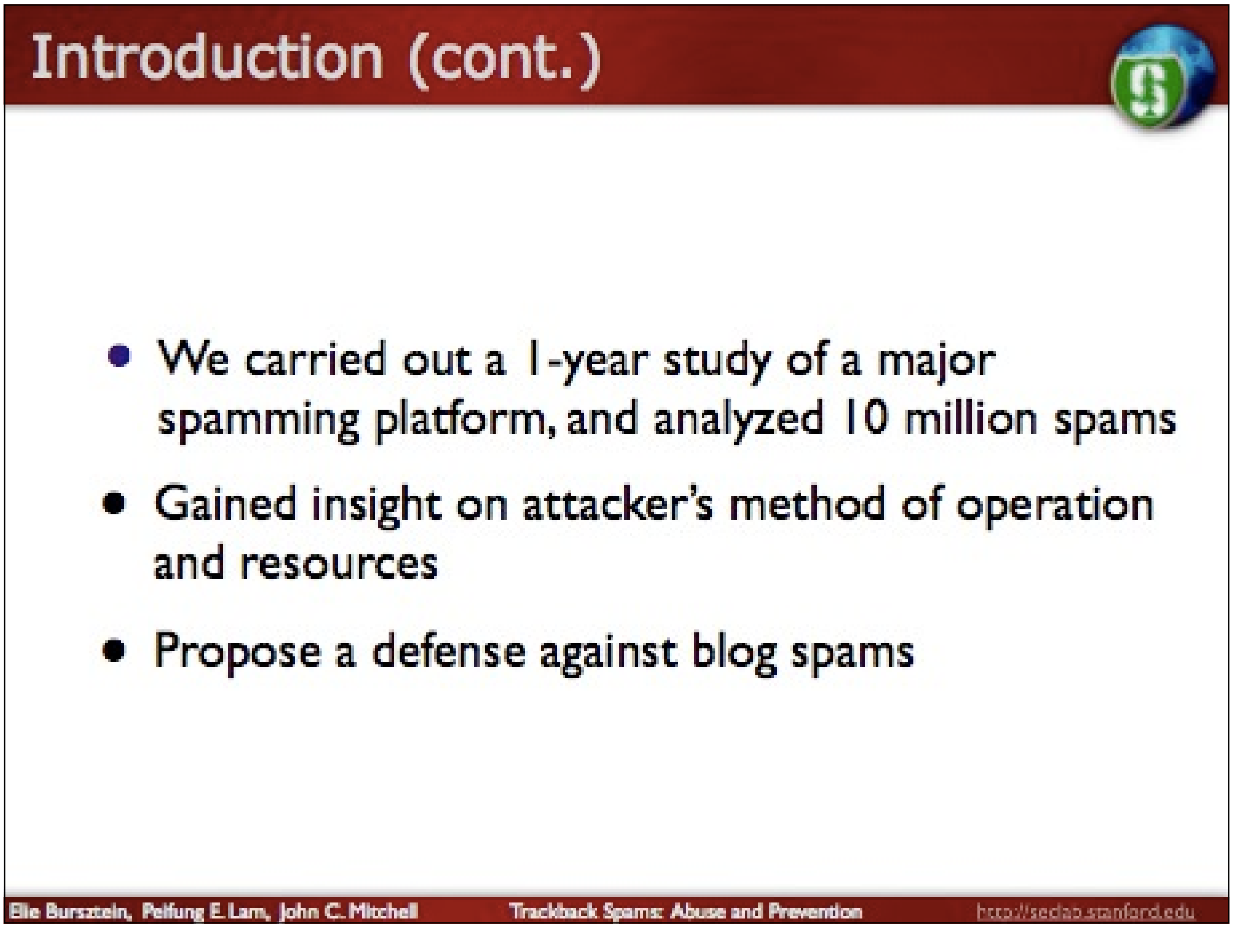
## Introduction

- Many users nowadays post information on cloud computing sites
- Sites sometimes need to link to each other
- However, cross-referencing can become a vehicle for abuses (such as spamming)
- This calls for a study of security issues on cross-referencing between cloud sites



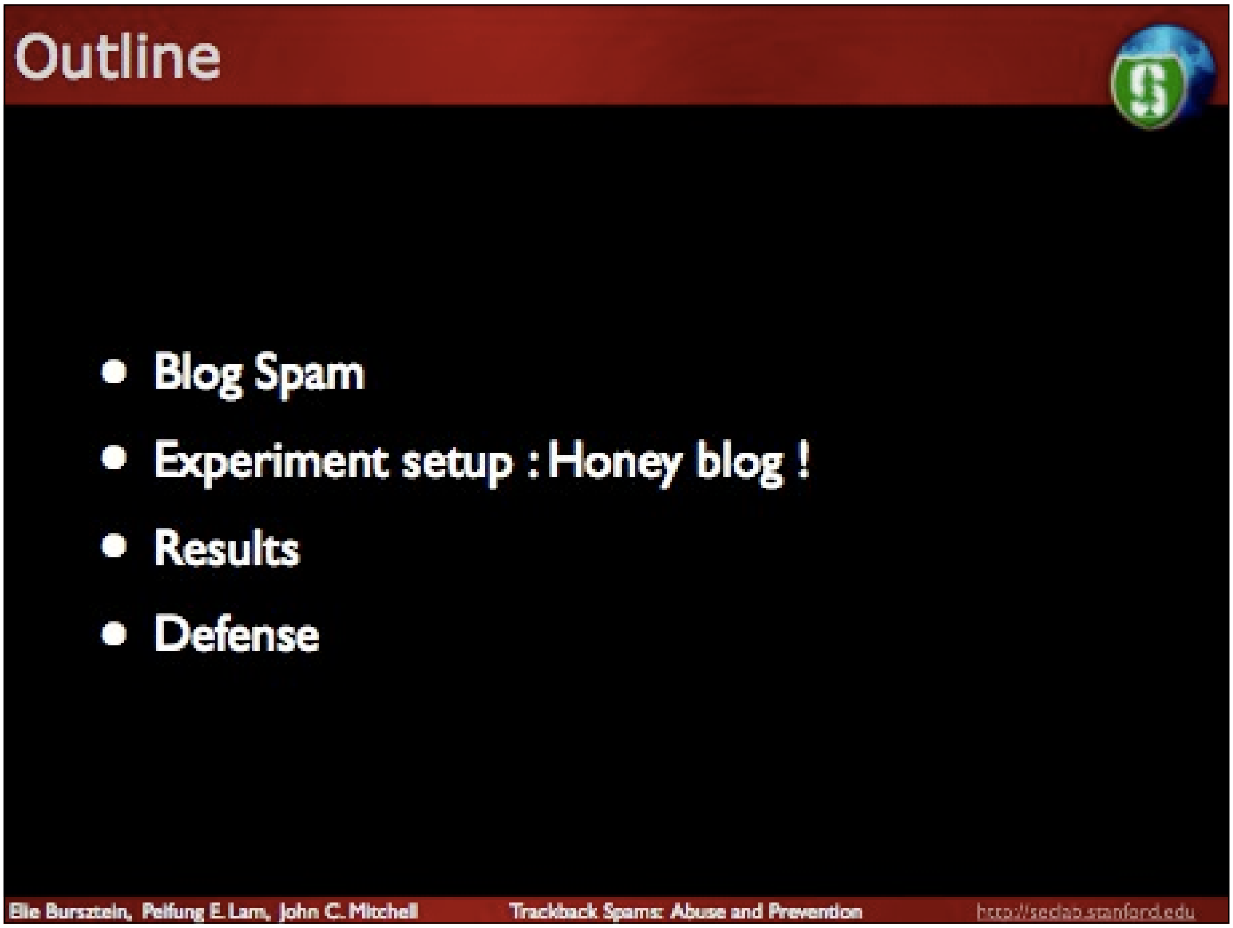
## Introduction (cont.)

- Blog cross-referencing offers one such example
- Blogs have automated mechanisms, called Linkbacks, to facilitate cross-referencing, and this has been exploited by spammers



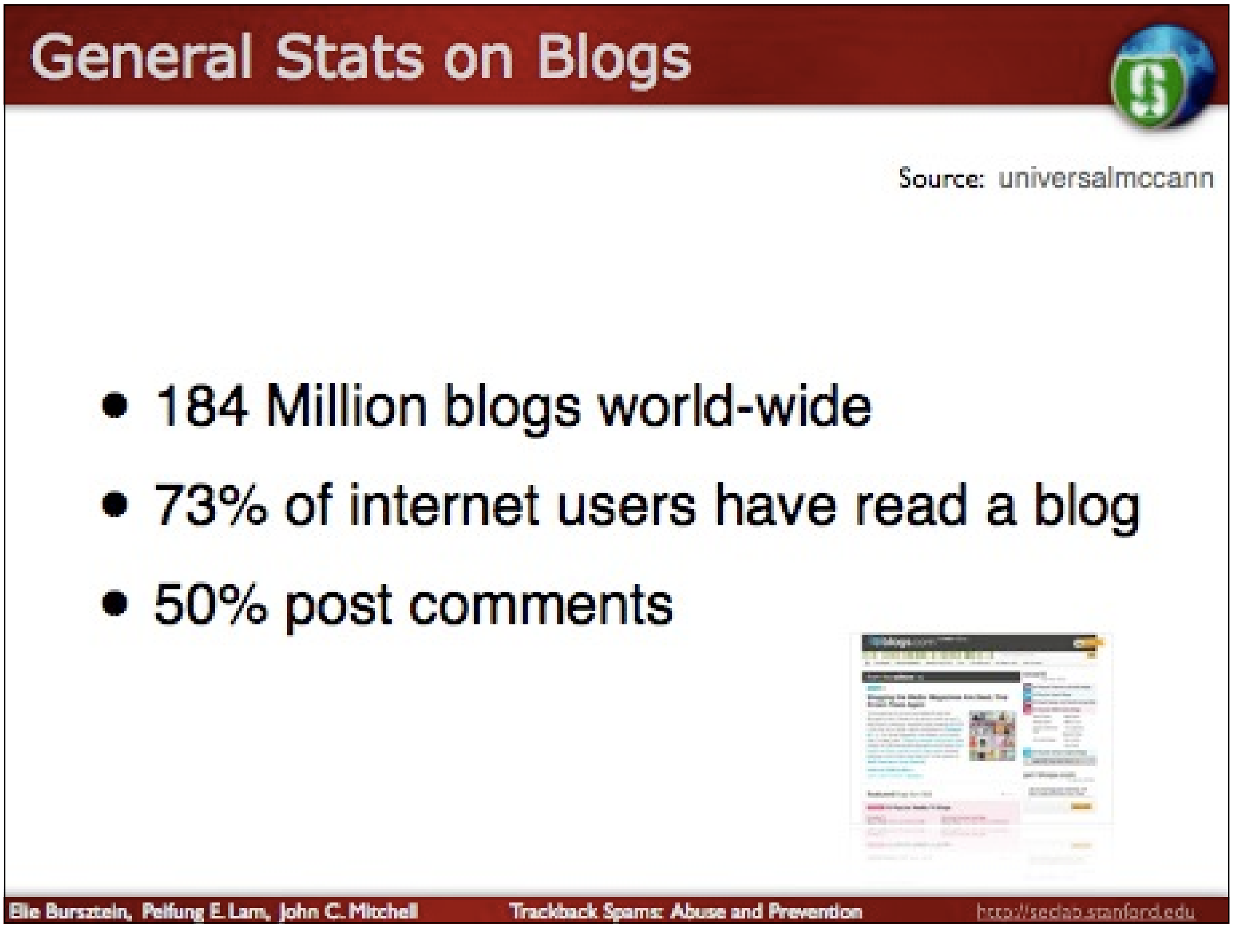
## Introduction (cont.)

- We carried out a 1-year study of a major spamming platform, and analyzed 10 million spams
- Gained insight on attacker's method of operation and resources
- Propose a defense against blog spams



## Outline

- Blog Spam
- Experiment setup : Honey blog !
- Results
- Defense



## General Stats on Blogs

Source: universalmccann

- 184 Million blogs world-wide
- 73% of internet users have read a blog
- 50% post comments

## Common Blog Platforms

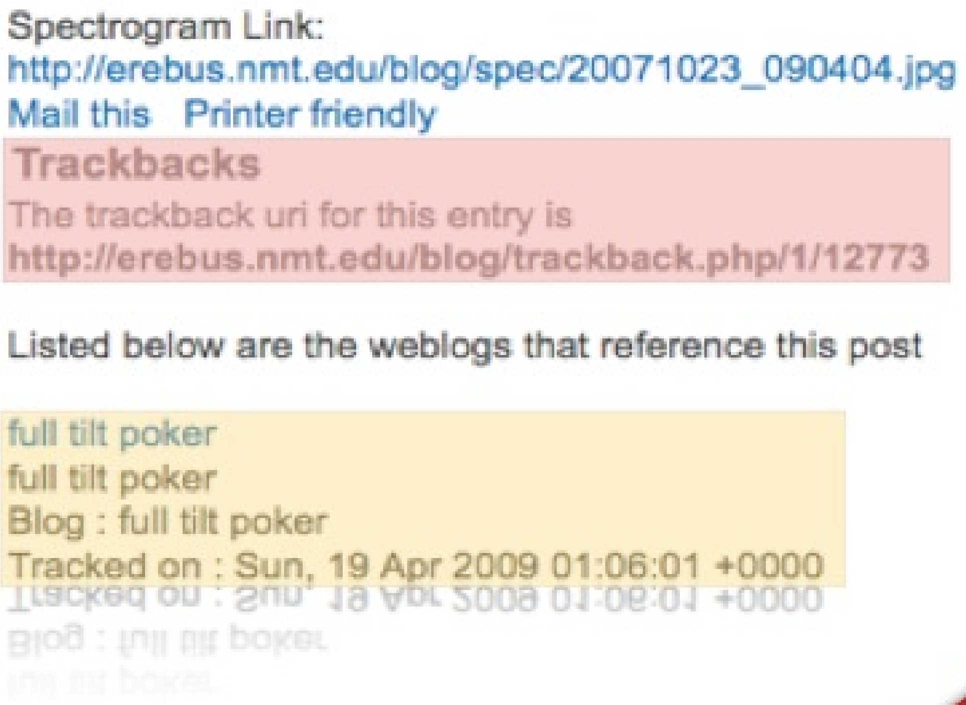
## Why blogs are special

- Blog are designed around the idea of user pushing content
- As an example, Linkbacks allow cross-linking between blogs.
- More specifically, when blog A cites another blog B, a notification of the citation can be sent to B, which can then link back to blog A automatically.

# TrackBack - a type of LinkBack

| | |
|---|---|
| TrackBack URL | The URL of TrackBack capture script |
| Auto discovery of TrackBack URL | Resource Description Framework (RDF) |
| Trigger | Code on blog site extracts citations to other blogs |
| Notification | HTTP Post |

# TrackBack URL and Blog Comments



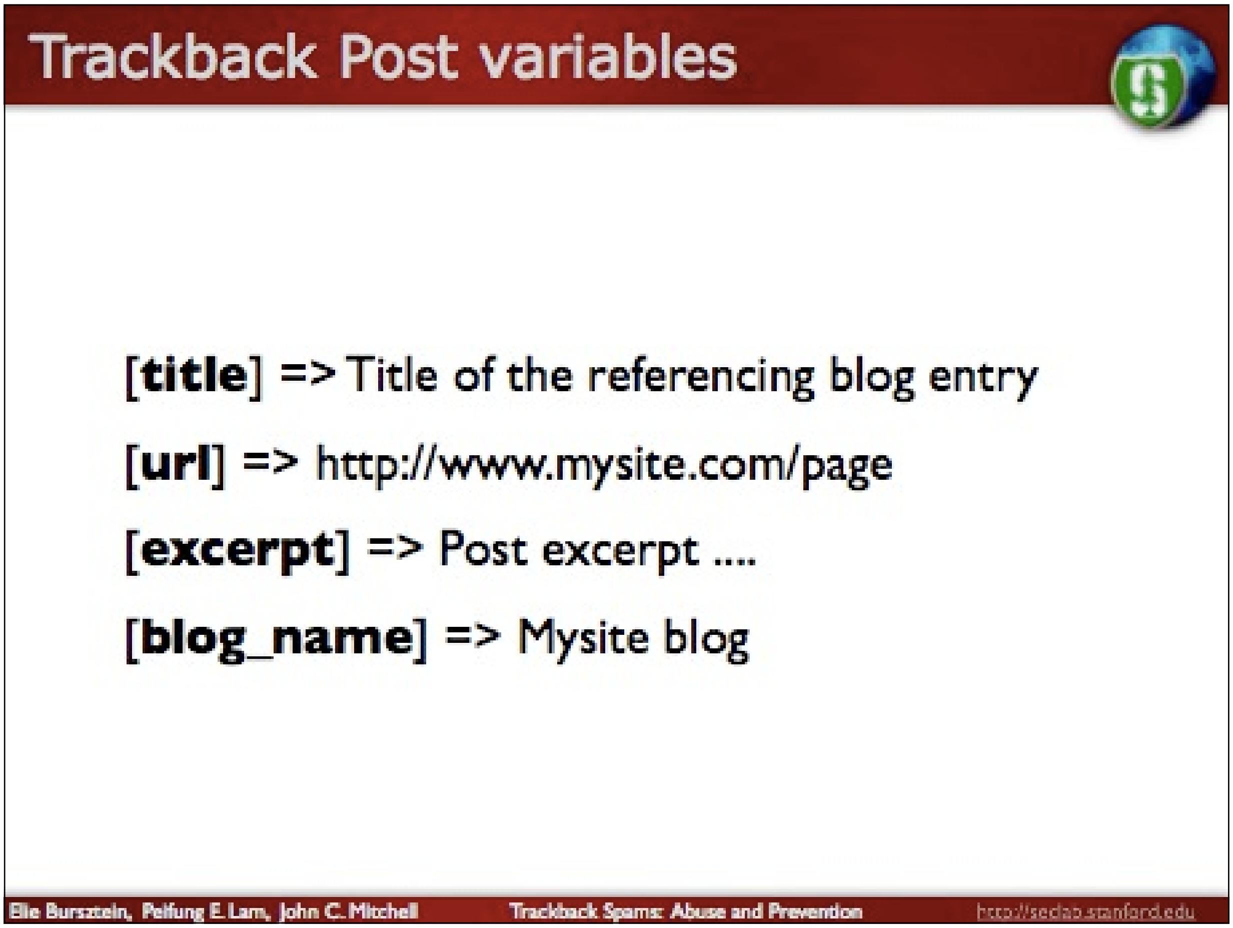
Spectrogram Link:
http://erebus.nmt.edu/blog/spec/20071023_090404.jpg
Mail this Printer friendly

**Trackbacks**

The trackback uri for this entry is
**http://erebus.nmt.edu/blog/trackback.php/1/12773**

Listed below are the weblogs that reference this post

full tilt poker
full tilt poker
Blog : full tilt poker
Tracked on : Sun, 19 Apr 2009 01:06:01 +0000



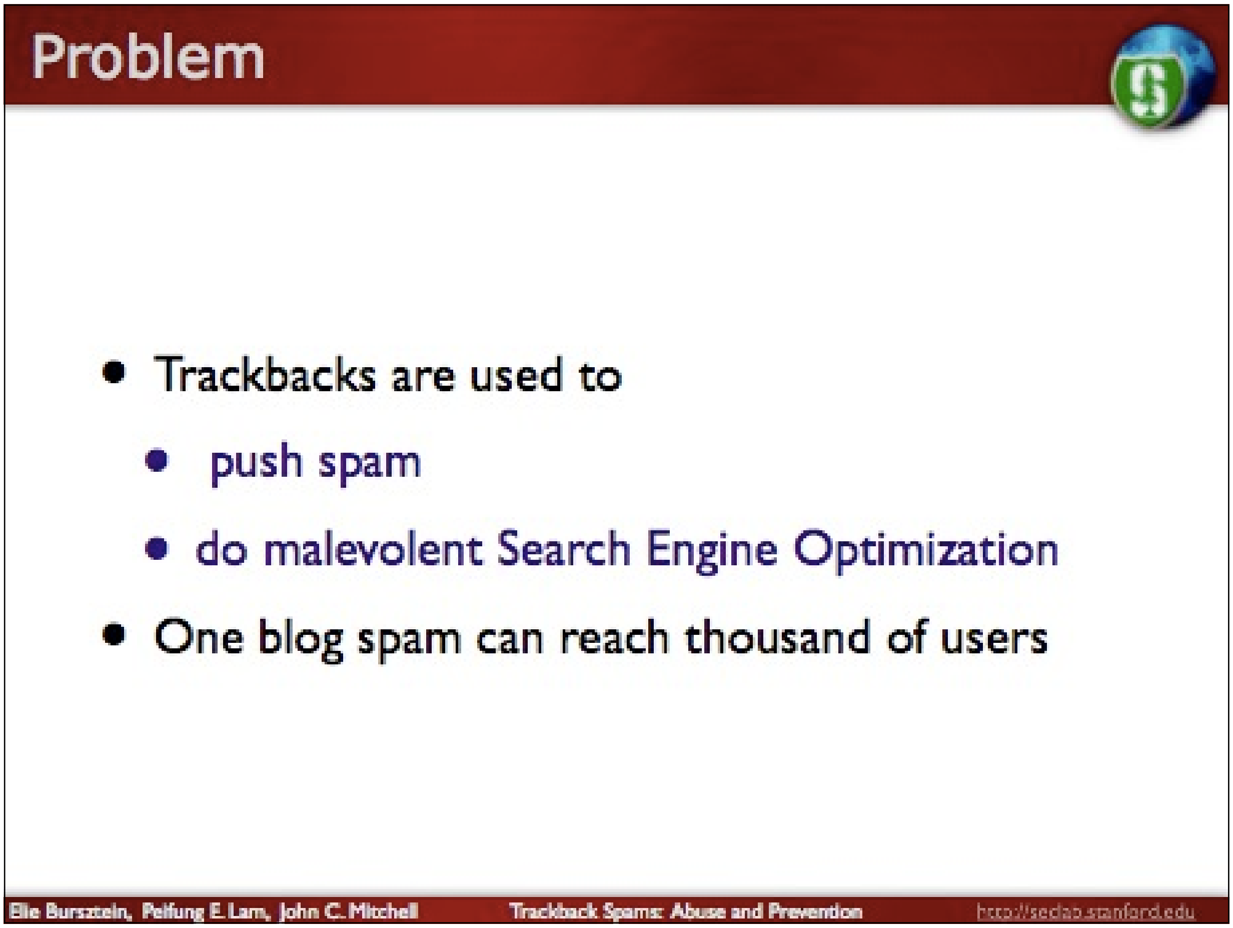
## Trackback Post variables

[**title**] => Title of the referencing blog entry

[**url**] => http://www.mysite.com/page

[**excerpt**] => Post excerpt ....

[**blog_name**] => Mysite blog



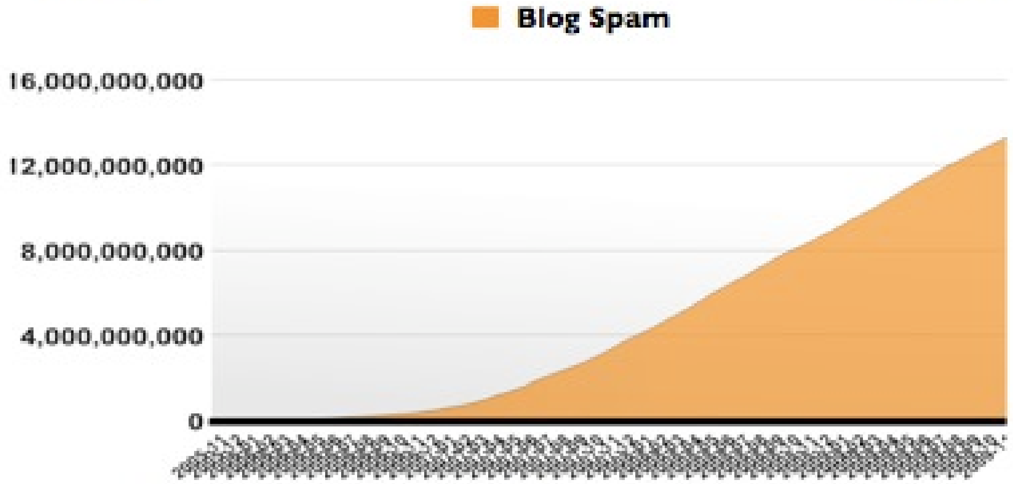
## Problem

- Trackbacks are used to
  - push spam
  - do malevolent Search Engine Optimization
- One blog spam can reach thousand of users

# How big is the problem?

Source: Akismet.com

Totals



**Total spam:** 13,275,940,950

**Total ham:** 2,701,440,026

Today (UTC, 1 hours left)

**Spam today:** 12,141,992

**Ham today:** 2,567,818

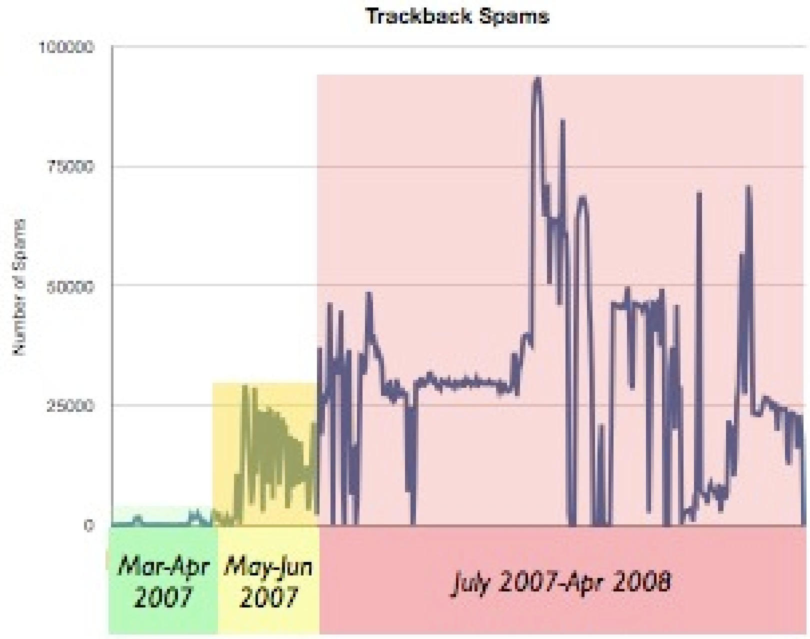
# Honey Blog

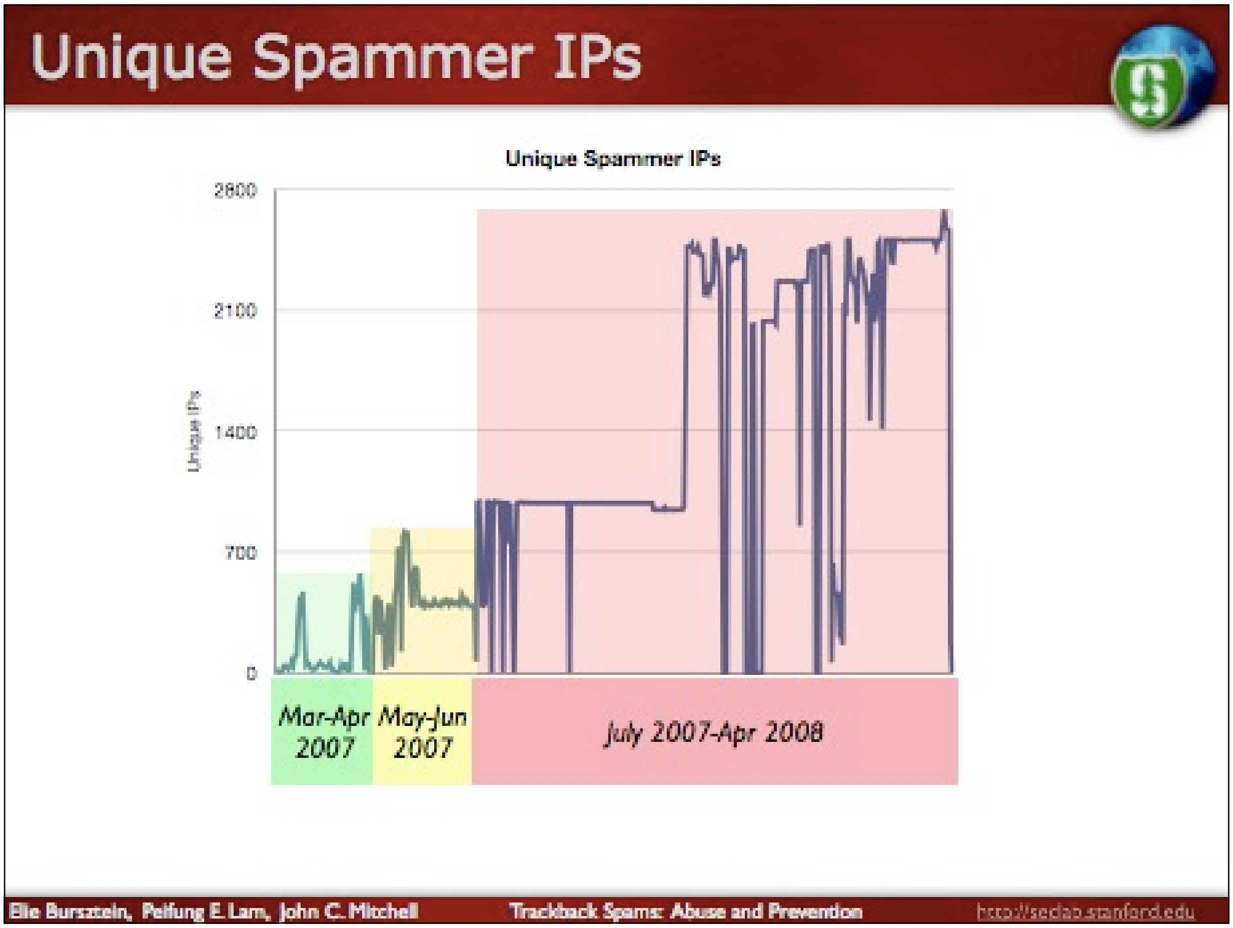
- A blog acting as a potential target for spamming
- Instrumented our blog site and analyzed spams

# Setup

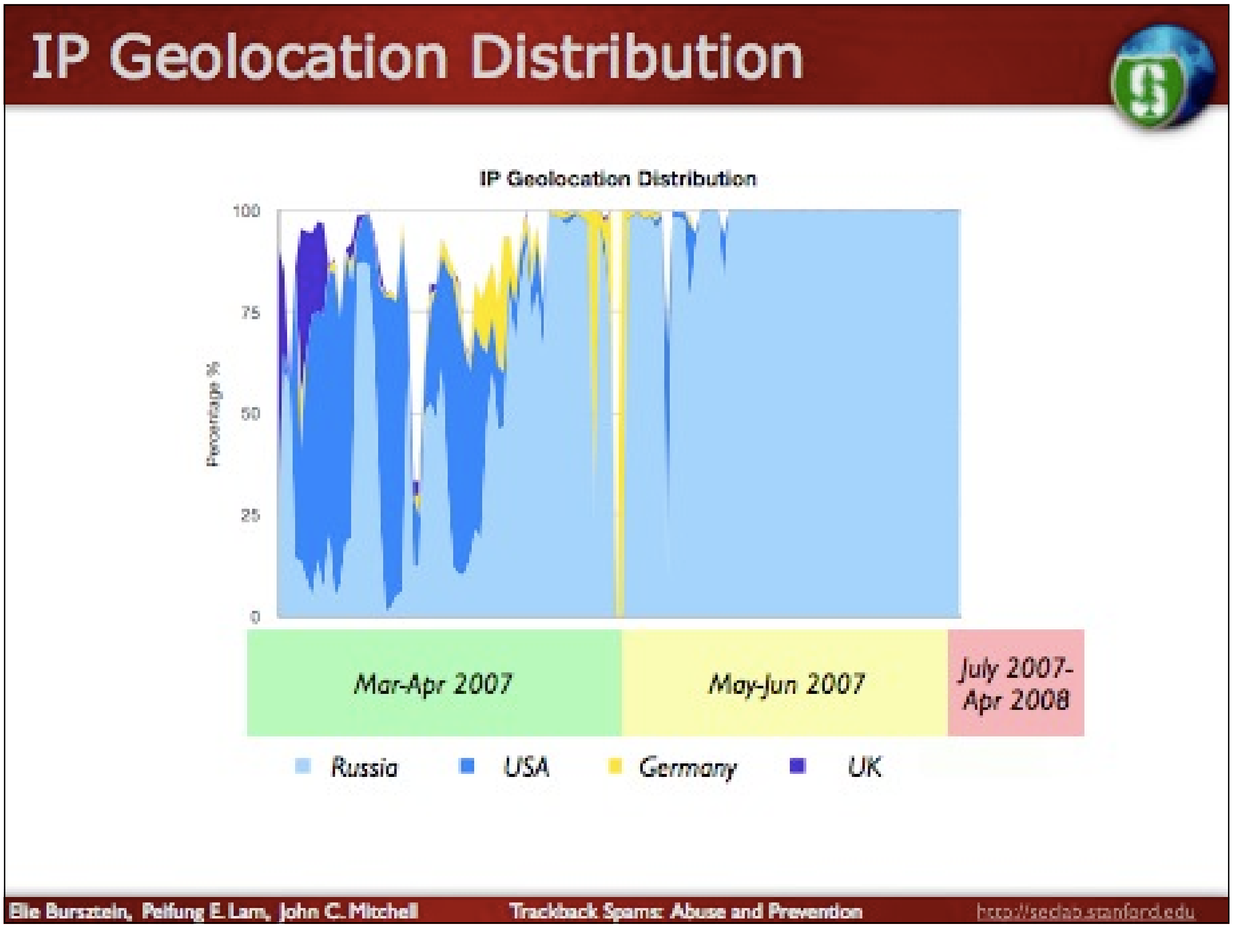
- Hosted a real blog (dotclear) with a modified TrackBack mechanism
- Record TrackBacks
- Passive fingerprinting
- Sample the lure site

# Activity

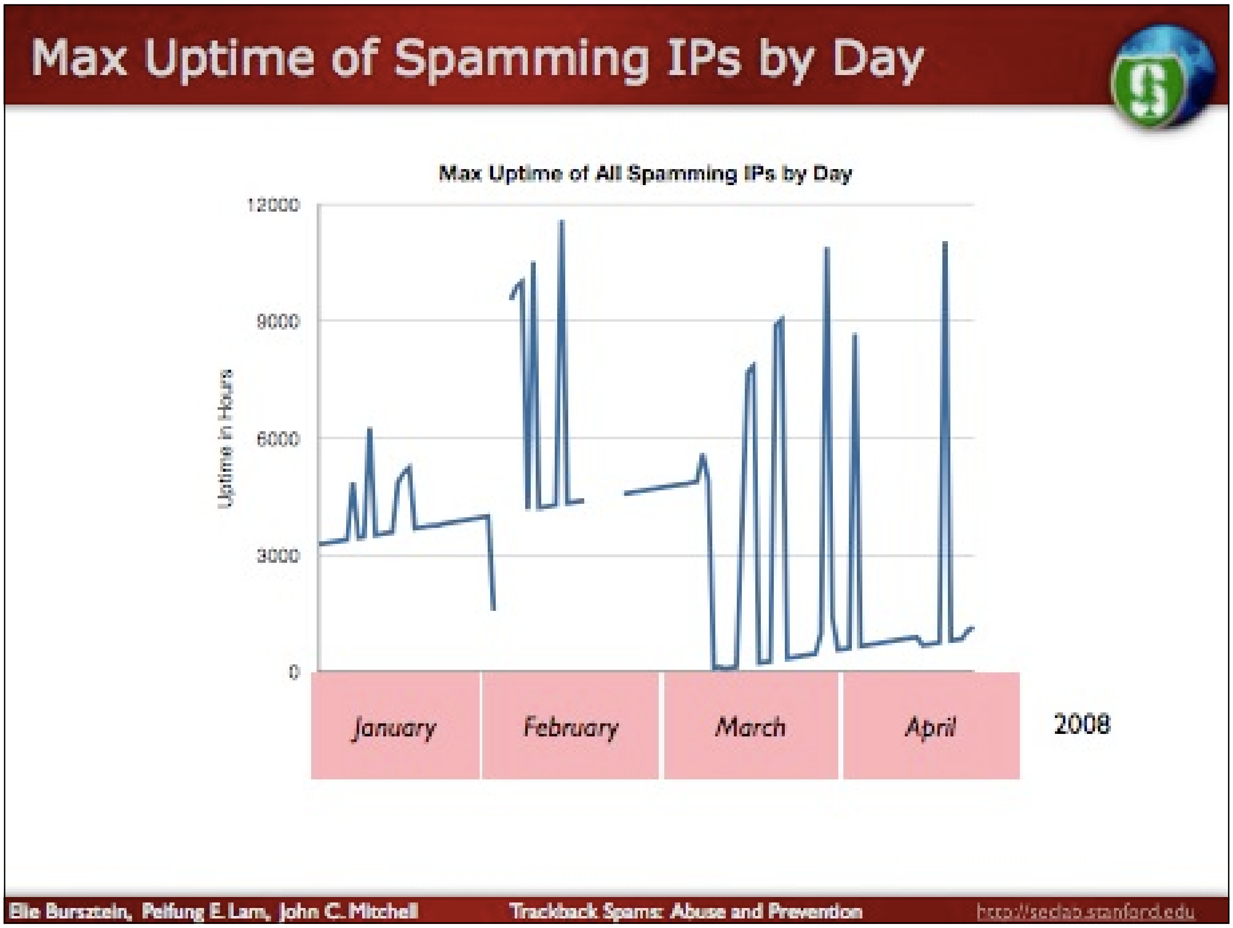




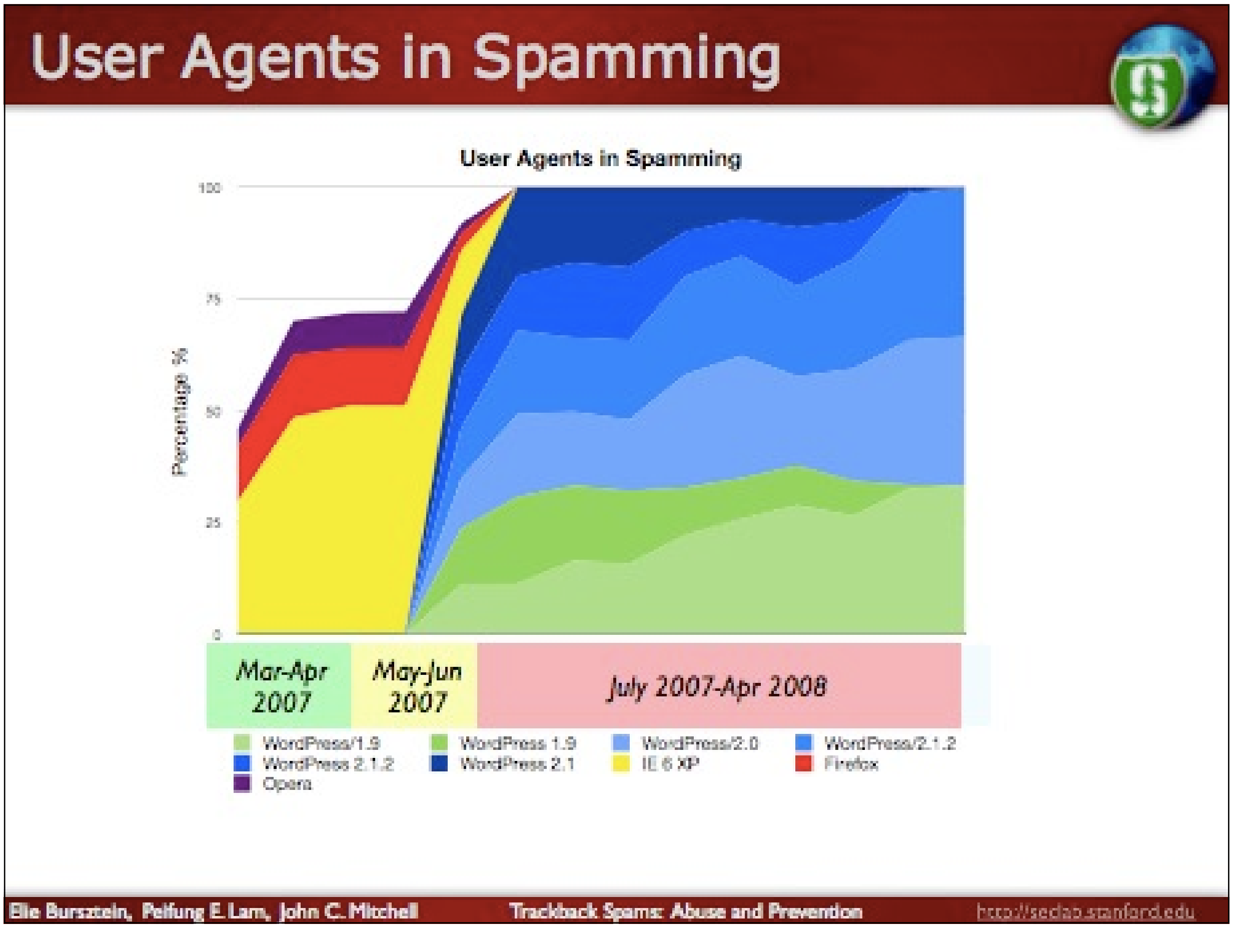
# Unique Spammer IPs



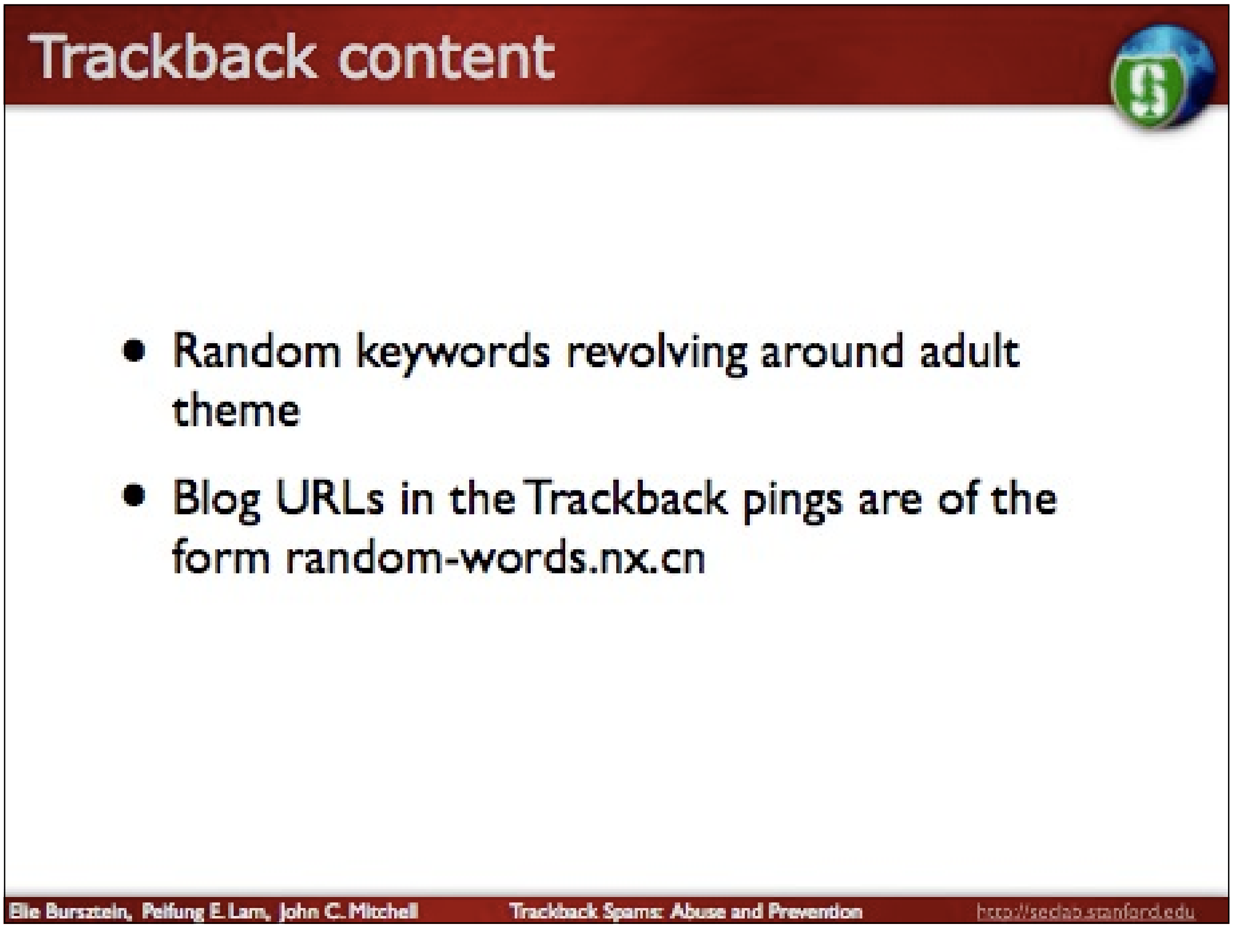
# IP Geolocation Distribution



# Max Uptime of Spamming IPs by Day

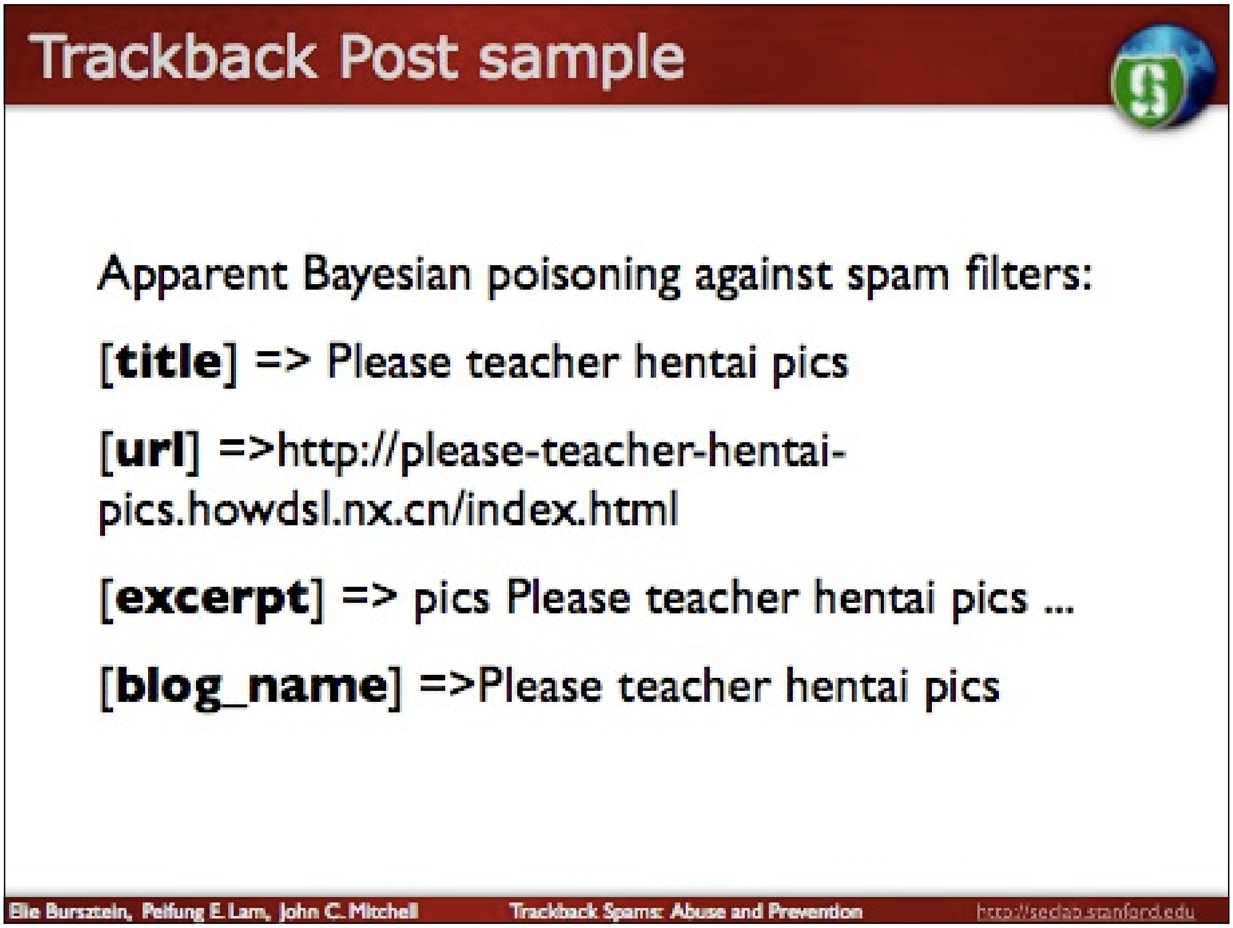


# User Agents in Spamming



## Trackback content

- Random keywords revolving around adult theme
- Blog URLs in the Trackback pings are of the form random-words.nx.cn



## Trackback Post sample

Apparent Bayesian poisoning against spam filters:

[**title**] => Please teacher hentai pics

[**url**] =>http://please-teacher-hentai-pics.howdsl.nx.cn/index.html

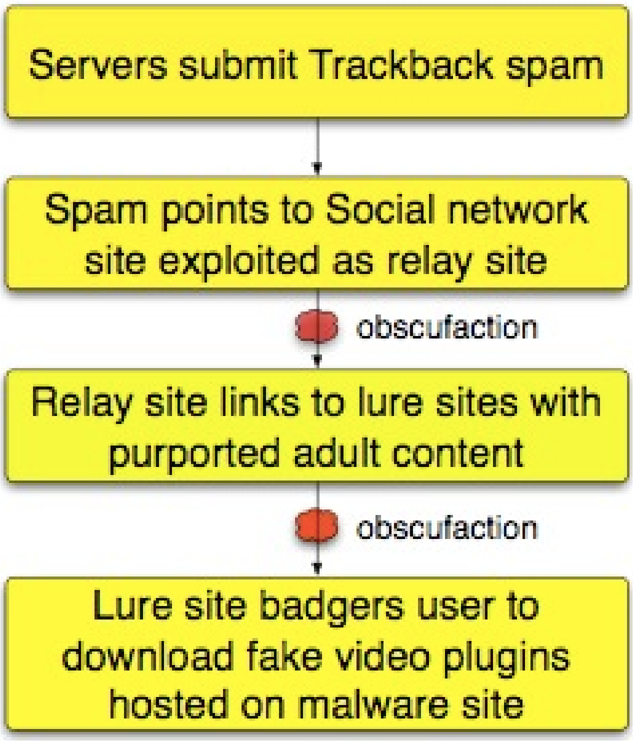
[**excerpt**] => pics Please teacher hentai pics ...

[**blog_name**] =>Please teacher hentai pics

Nude
Sex
porn
pics
free
girls
naked
girl
pictures
cartoon
pussy
Videos
boobs
Vintage
movies
photos
porno
video
teen
clips
stories
gay
sexy
adult
hentai
penis
breast
pants
Teens
Female
Escort
gallery
playboy
women
Big
wife
tits
Asian
Anal

Created using Wordle

## Spam Workflow



Servers submit Trackback spam

Spam points to Social network site exploited as relay site

obscufaction

Relay site links to lure sites with purported adult content

obscufaction

Lure site badgers user to download fake video plugins hosted on malware site

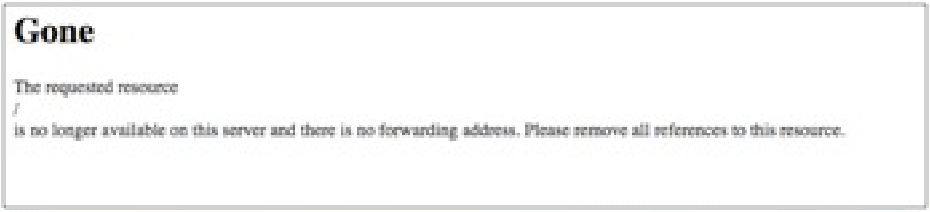
# Relay URL

- Www.nx.cn, a community hosting site at Ningxia province, PRC
- Exploited by attackers as relay
- The hosting site started to use CAPTCHA (some in Chinese) around May, 2008
- We observed a corresponding drop of spam activities using them as relay

Administrative Divisions of the People's Republic of China (PRC)
XINJIANG UYGHUR AUTONOMOUS REGION
Indian claim line
TIBET AUTONOMOUS REGION
Lhasa
Chinese claim line
QINGHAI
Xining
NINGXIA HUI AUTONOMOUS REGION
Hohhot
INNER MONGOLIA AUTONOMOUS REGION
HEILONGJIANG
Harbin
JILIN
Changchun
Shenyang
LIAONING
BEIJING
TIANJIN
HEBEI
Shijiazhuang
Taiyuan
SHANXI
Jinan
SHANDONG
GANSU
Xi'an
SHAANXI
Zhengzhou
HENAN
JIANGSU
Nanjing
SHANGHAI
Hefei
ANHUI
SICHUAN
Chengdu
HUBEI
Wuhan
Hangzhou
ZHEJIANG
Nanchang
Changsha
HUNAN
JIANGXI
Fuzhou
FUJIAN
GUIZHOU
Guiyang
Kunming
YUNNAN
GUANGXI ZHUANG AUTONOMOUS REGION
Nanning
GUANGDONG
Guangzhou
HONG KONG S.A.R.
MACAU S.A.R.
Haikou
HAINAN
TAIWAN controlled by the Republic of China, claimed by the People's Republic of China
Province
Autonomous Region
Municipality
Special Administrative Region

# Behind the relay

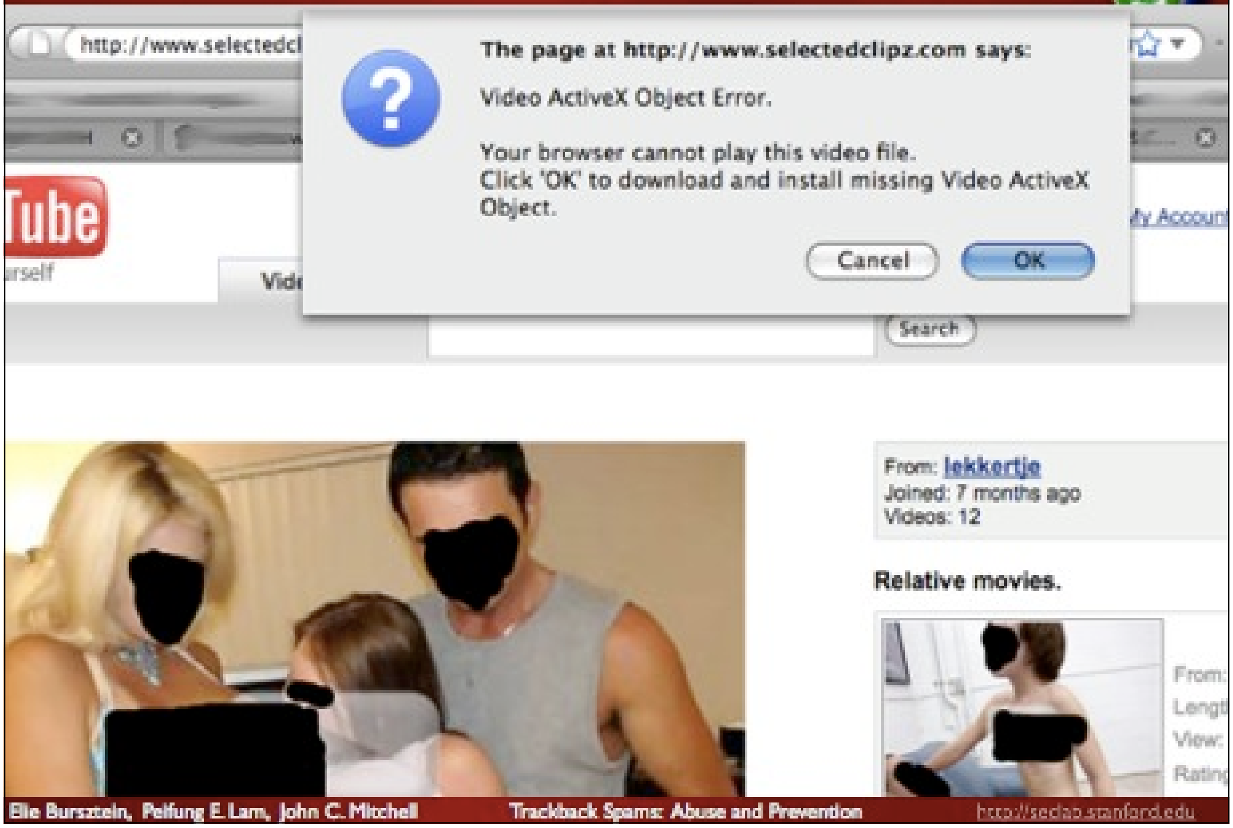
- Lead to various sites
  - selectedclipz.com, gogomovz.com (purported adult site)
  - vidzwares.com (malware distribution site)
- Need an id in the url download.php?id=429



## Gone

The requested resource
/
is no longer available on this server and there is no forwarding address. Please remove all references to this resource.

# The Lure site



The page at http://www.selectedclipz.com says:

Video ActiveX Object Error.

Your browser cannot play this video file.
Click 'OK' to download and install missing Video ActiveX Object.

# Whois

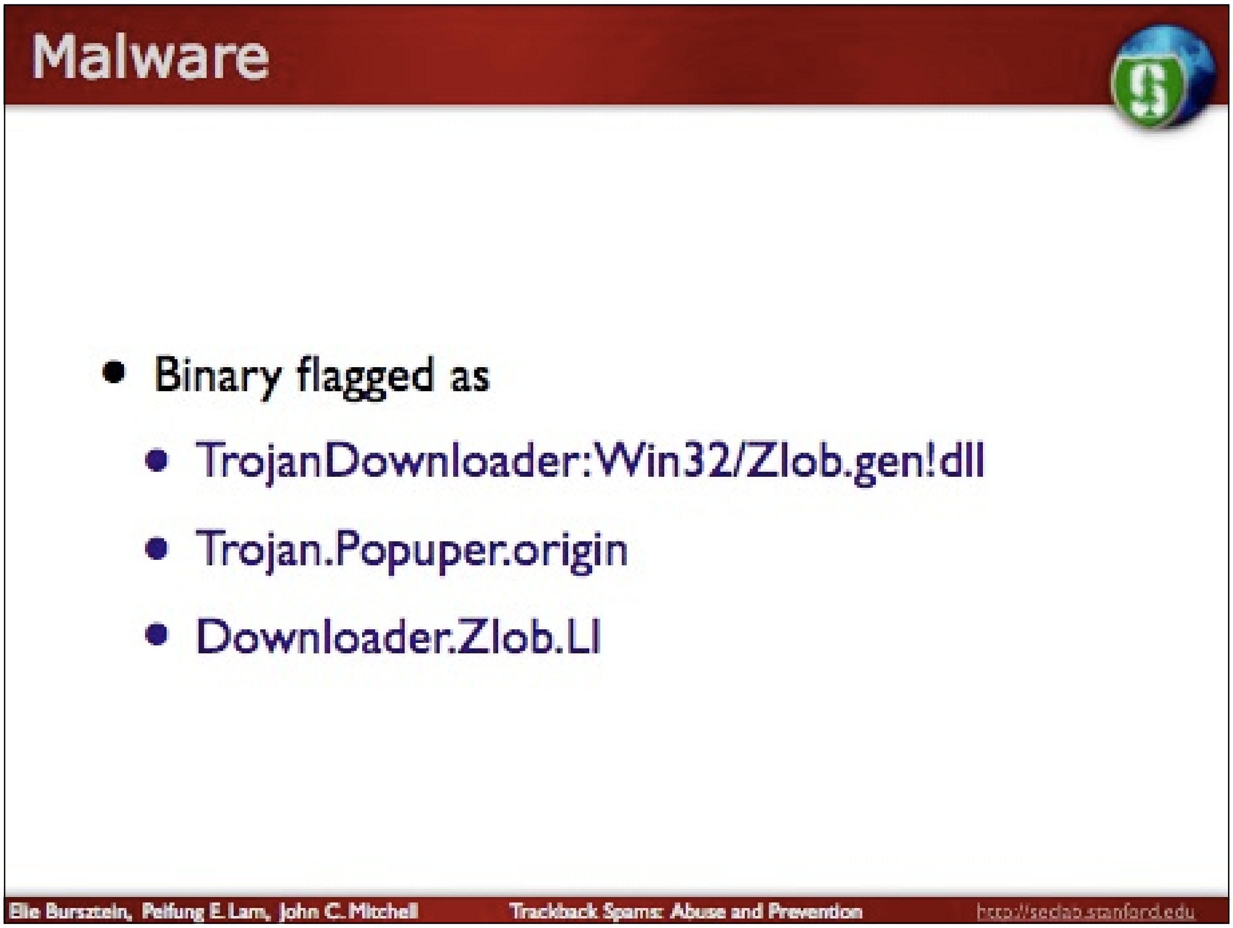
Domain Name: GOGOMOVZ.COM<br>
Registrar: ONLINENIC, INC.<br>
Whois Server: whois.onlinenic.com<br>
Referral URL: http://www.OnlineNIC.com<br>
Name Server: NS1.GOGOMOVZ.COM<br>
Name Server: NS2.GOGOMOVZ.COM.<br>
Updated Date: 22-oct-2008<br>
Creation Date: 22-oct-2008<br>
Expiration Date: 22-oct-2009

Registrant:

...<br>
ul Beketova 3<br>
Nijnii Novgorod,n/a,RUSSIAN FEDERATION 603057

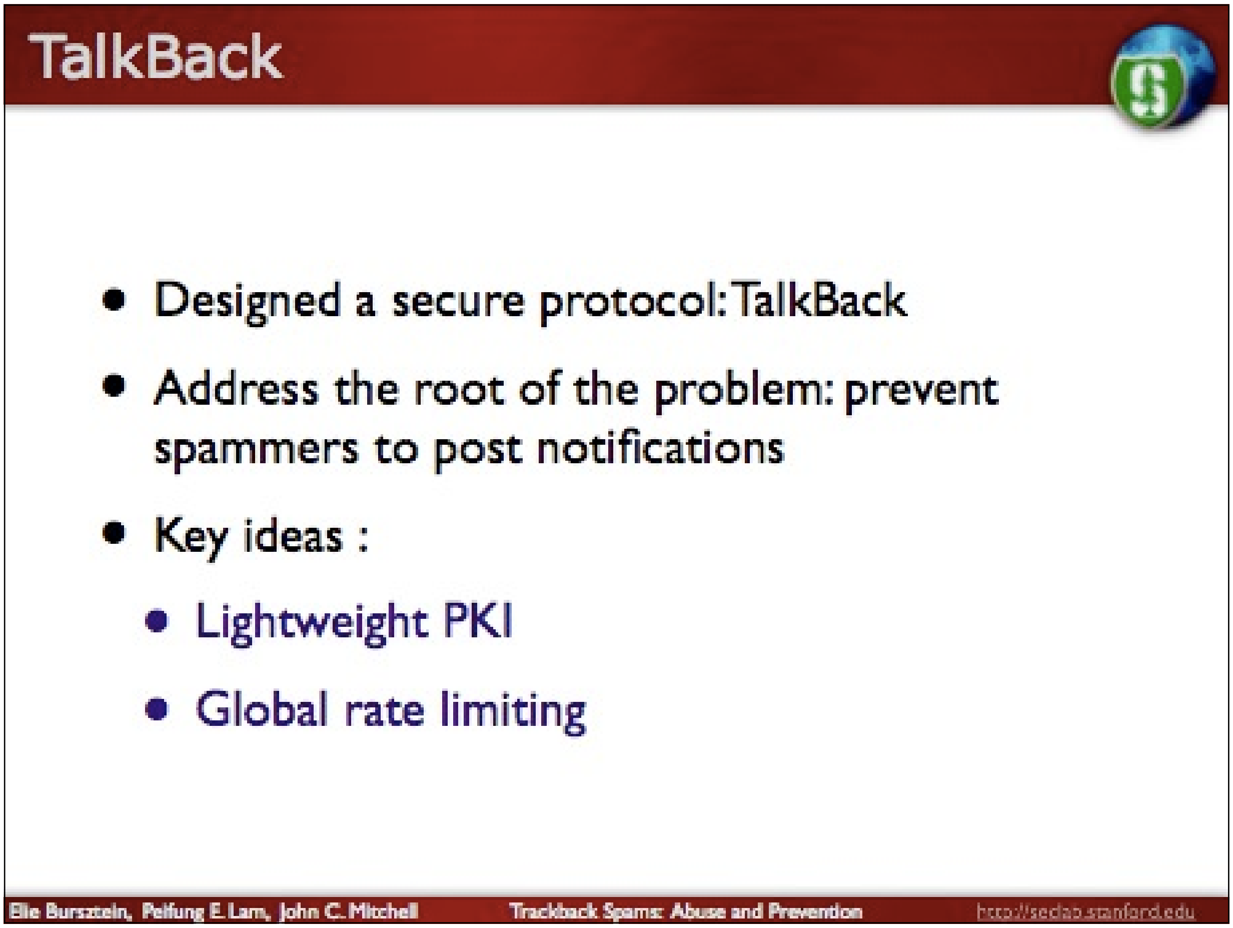
# DNS analysis : related domains

- ns1.clipzsaloon.com
- ns1.clipztube.com
- ns1.freexxxmovz.com
- ns1.itunnelz.com
- ns1.vidzselector.com, and more...



## Malware

- Binary flagged as
  - TrojanDownloader:Win32/Zlob.gen!dll
  - Trojan.Popuper.origin
  - Downloader.Zlob.LI



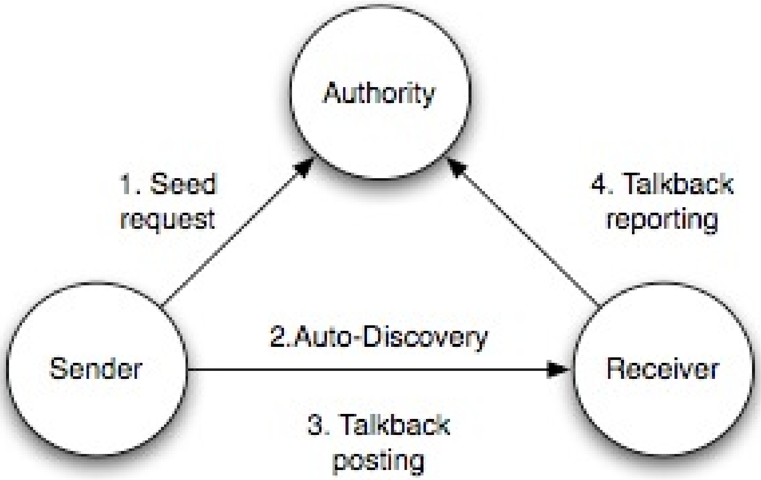
## TalkBack

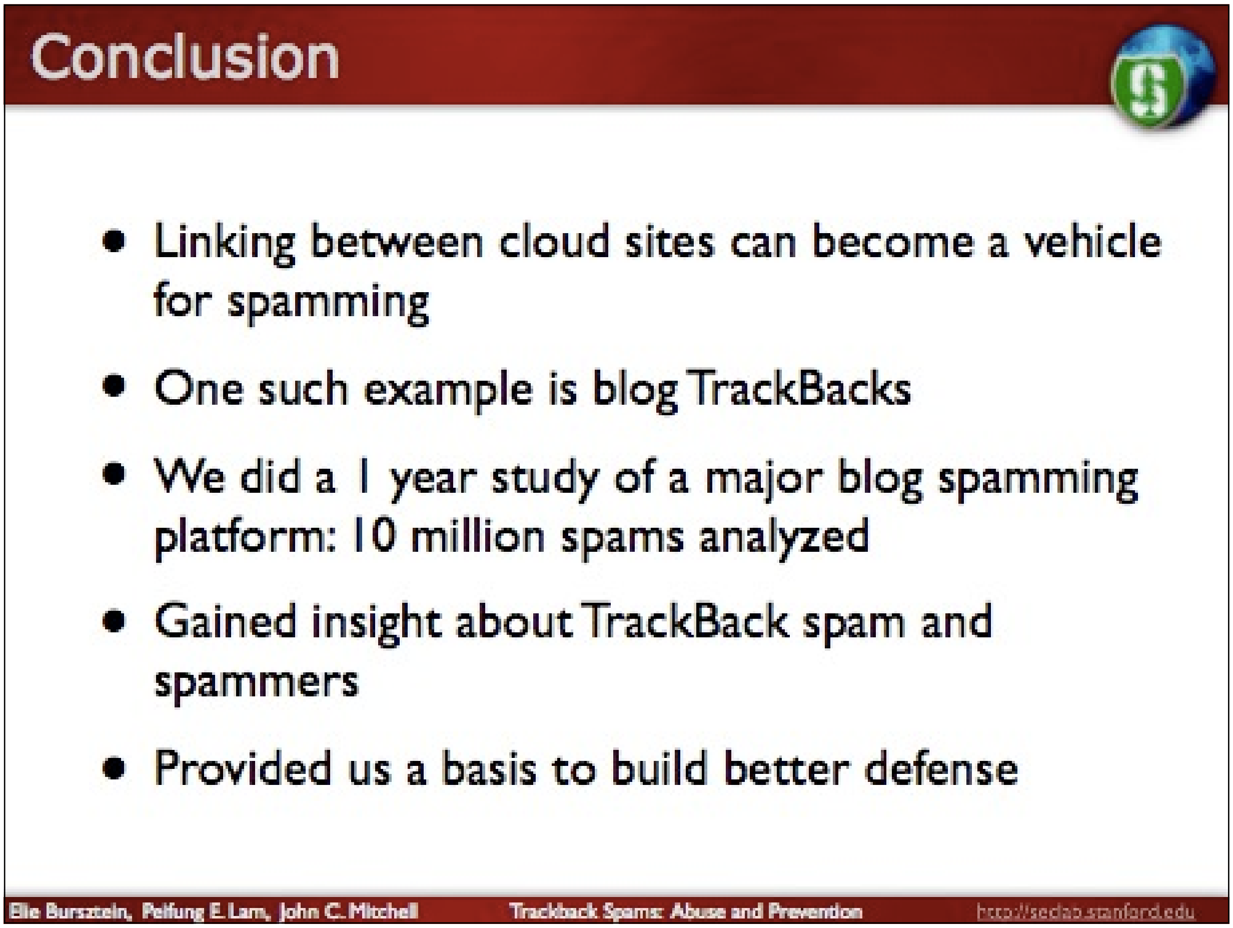
- Designed a secure protocol: TalkBack
- Address the root of the problem: prevent spammers to post notifications
- Key ideas :
  - Lightweight PKI
  - Global rate limiting

# Goals

- Sender authenticity
- Receiver authenticity
- Notification integrity
- Notification irrefutability

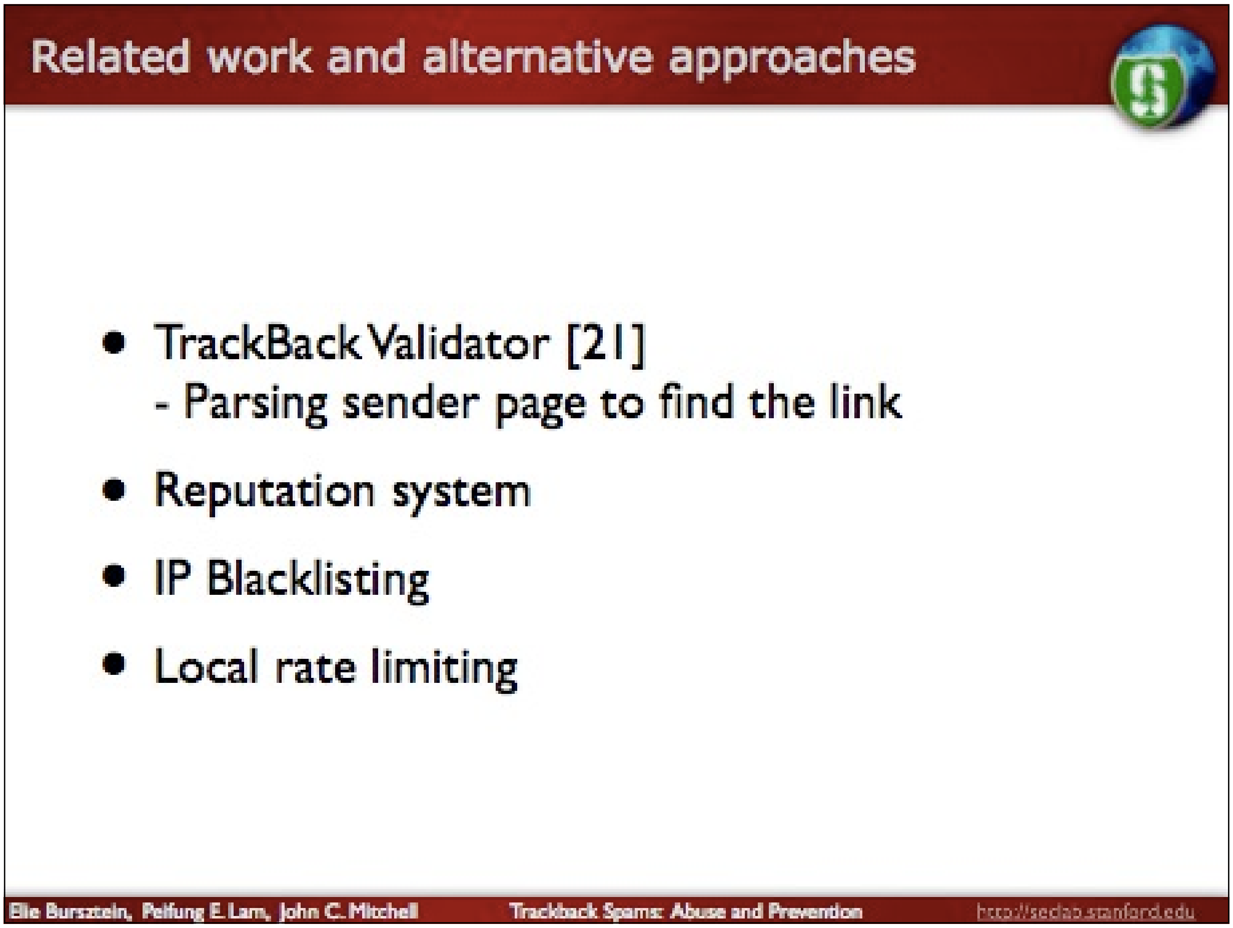
# How it works





## Conclusion

- Linking between cloud sites can become a vehicle for spamming
- One such example is blog TrackBacks
- We did a 1 year study of a major blog spamming platform: 10 million spams analyzed
- Gained insight about TrackBack spam and spammers
- Provided us a basis to build better defense



## Related work and alternative approaches

- TrackBack Validator [21]
  - Parsing sender page to find the link
- Reputation system
- IP Blacklisting
- Local rate limiting

# Questions ?

Thank you!

Stanford Computer Security Lab