Every Second Counts:

Quantifying the Negative Externalities of Cybercrime via Typosquatting

Mohammad Taha Khan[*], Xiang Huo[*], Zhou Li[†] & Chris Kanich[*]

University of Illinois at Chicago[*] & RSA Labs[†]

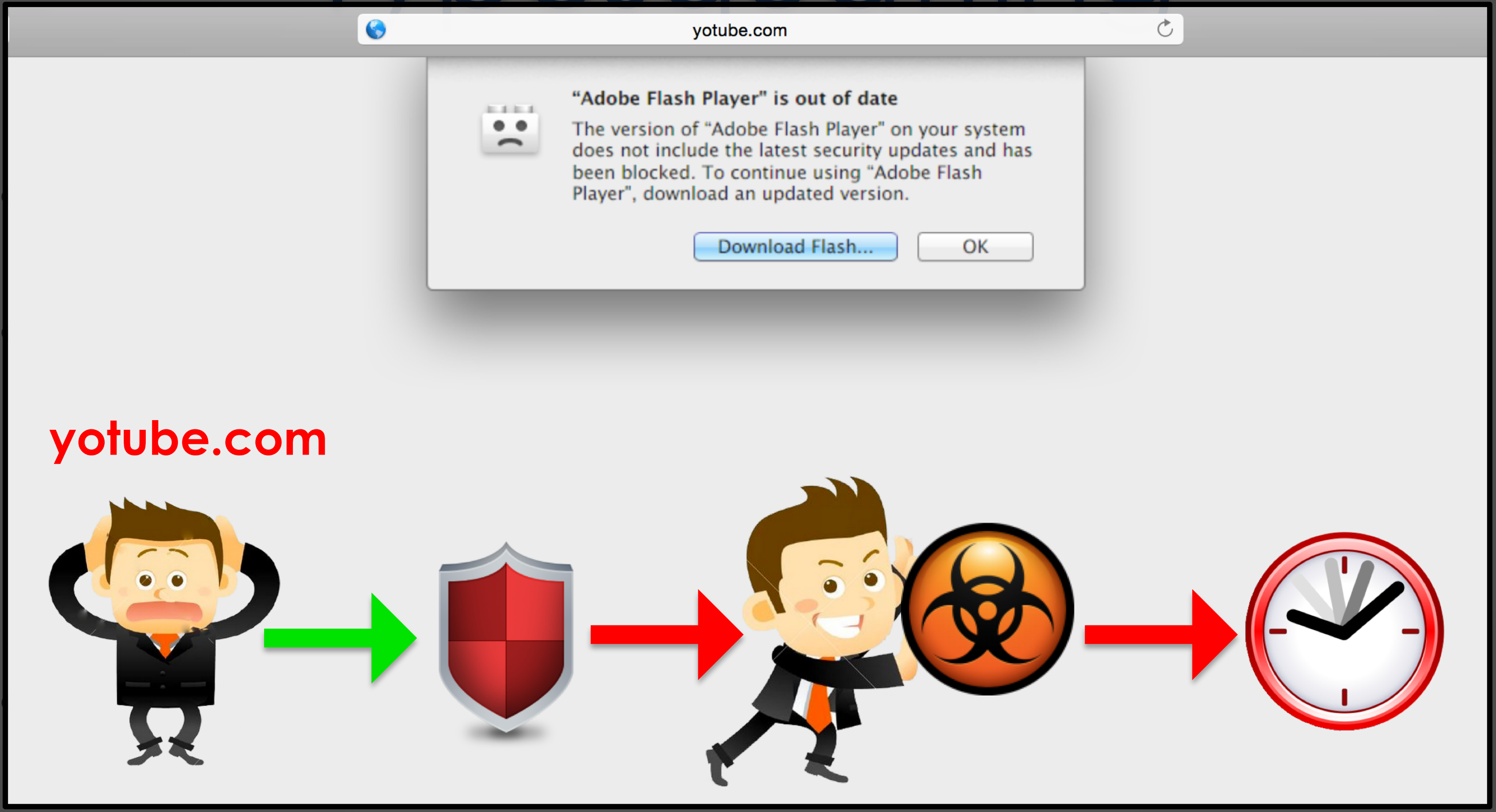
Cybersecurity In General...

- Generally focus on:
  - Detecting malicious programs
  - Finding and fixing bugs and flaws
  - Economic analyses

Why Is This Important?

- The ultimate goal:
  - Minimize the harm caused to users
  - Harm: Monetary, **wasted effort, loss of time**

Typosquatting

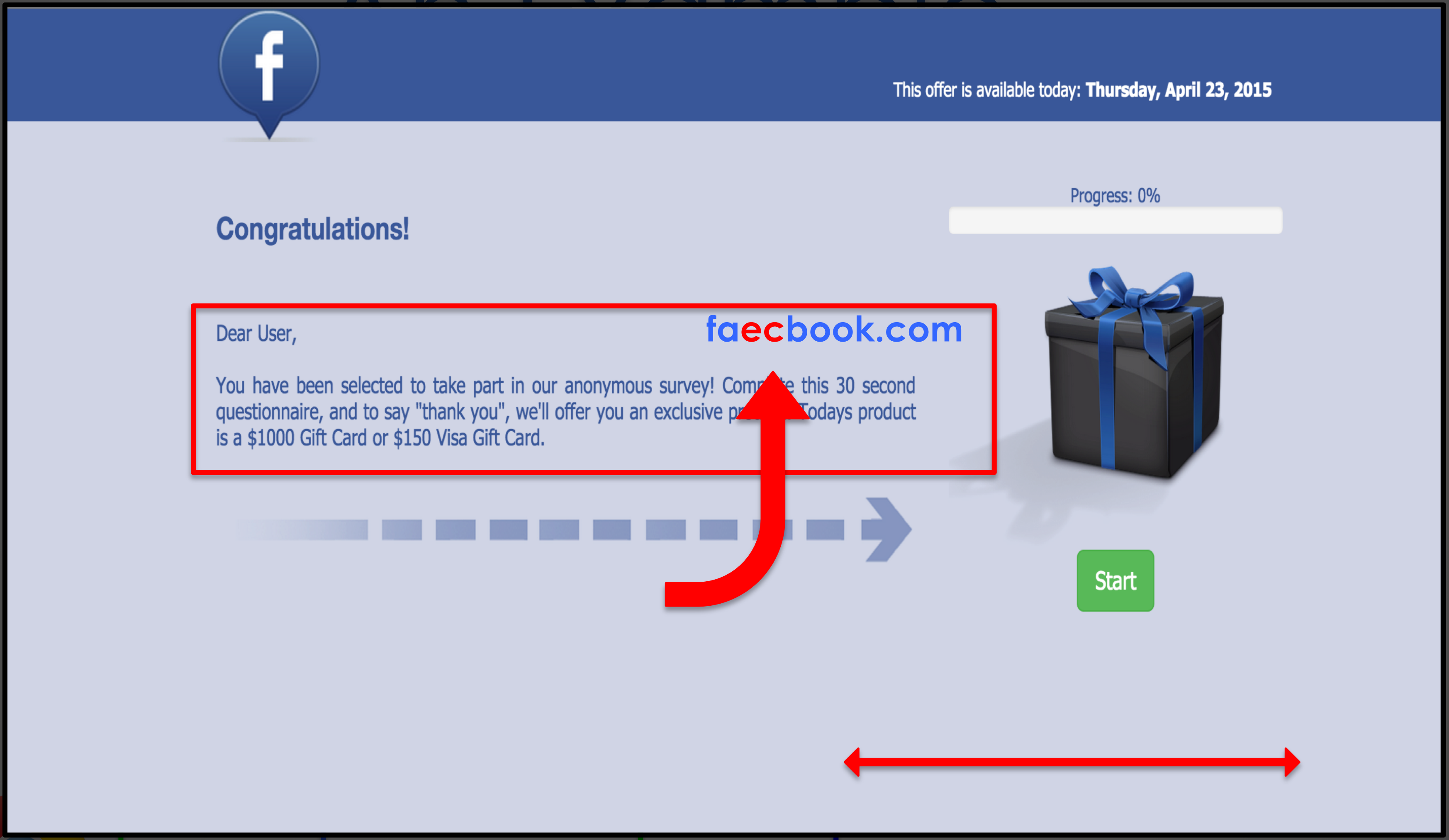


“Adobe Flash Player” is out of date

The version of “Adobe Flash Player” on your system does not include the latest security updates and has been blocked. To continue using “Adobe Flash Player”, download an updated version.

*[J. Szurdi, B. Kocso, G. Cseh, M. Felegyhazi, and C. Kanich, “The Long “Taile” of Typosquatting Domain Names, *USENIX*, 2014.]

An Example



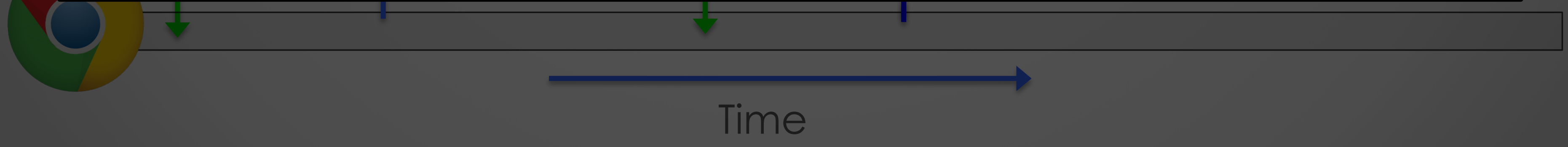
This offer is available today: **Thursday, April 23, 2015**

Progress: 0%

Congratulations!

Dear User,

You have been selected to take part in our anonymous survey! Comp e this 30 second questionnaire, and to say "thank you", we'll offer you an exclusive pr Todays product is a $1000 Gift Card or $150 Visa Gift Card.



Typosquatting

- Evidence that typosquatting is **PERVASIVE**:
  - Large organizations invest into defensive registrations
  - Internet users continue to make typos
- What makes it **FEASIBLE** to study:
  - Observable from a network level
  - Can infer **User intent** from available data

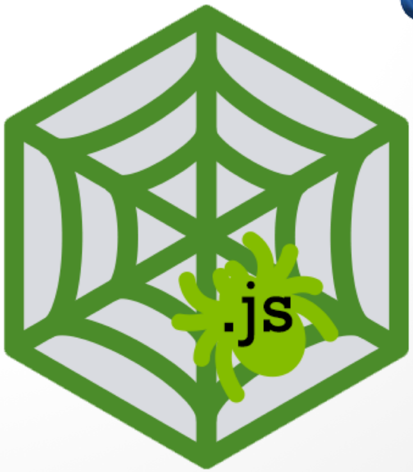
Our Contributions

- Passive detection of typosquatting domains using a conditional probability model
- Present a harm metric in the form of loss of time and users
- Apply this metric to quantify the cybercrime of typosquatting
- Our work uses an **open methodology** with fine grained measurements

Data-Sets

- Passive Sources
  - HTTP data logs
  - DNS logs from recursive resolver
  - Enterprise proxy data

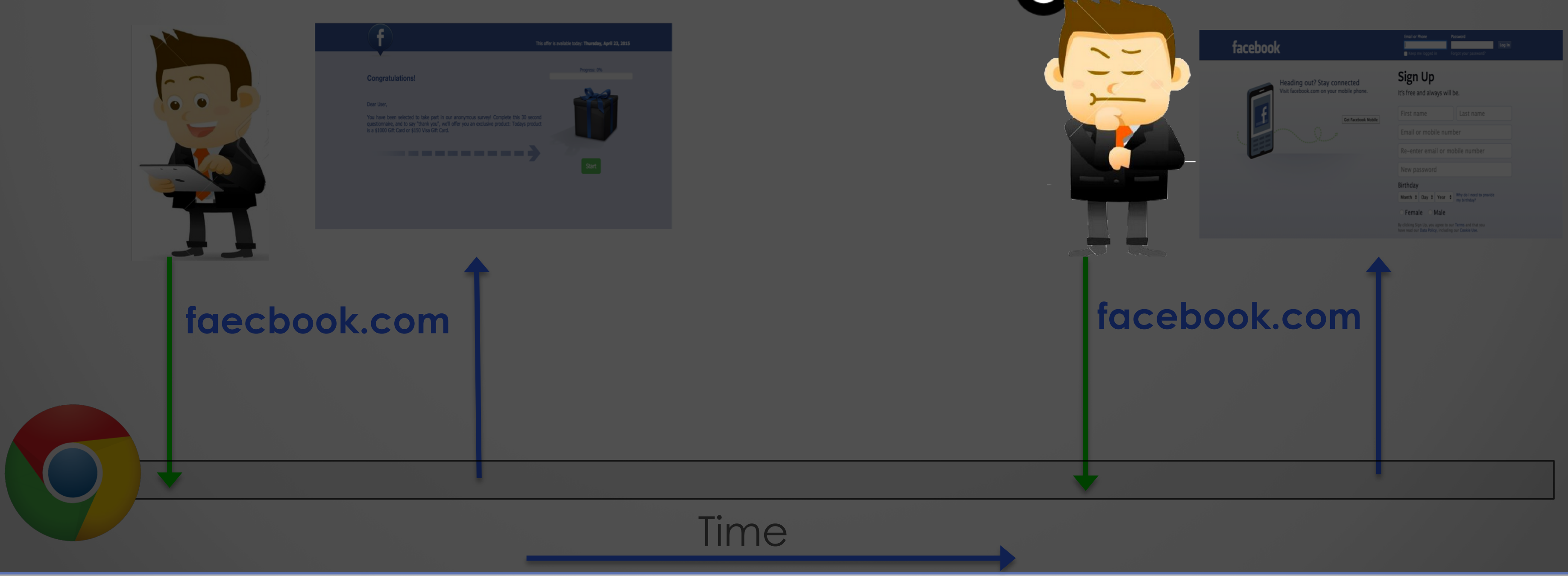
- Active Sources
  - High Fidelity Crawler



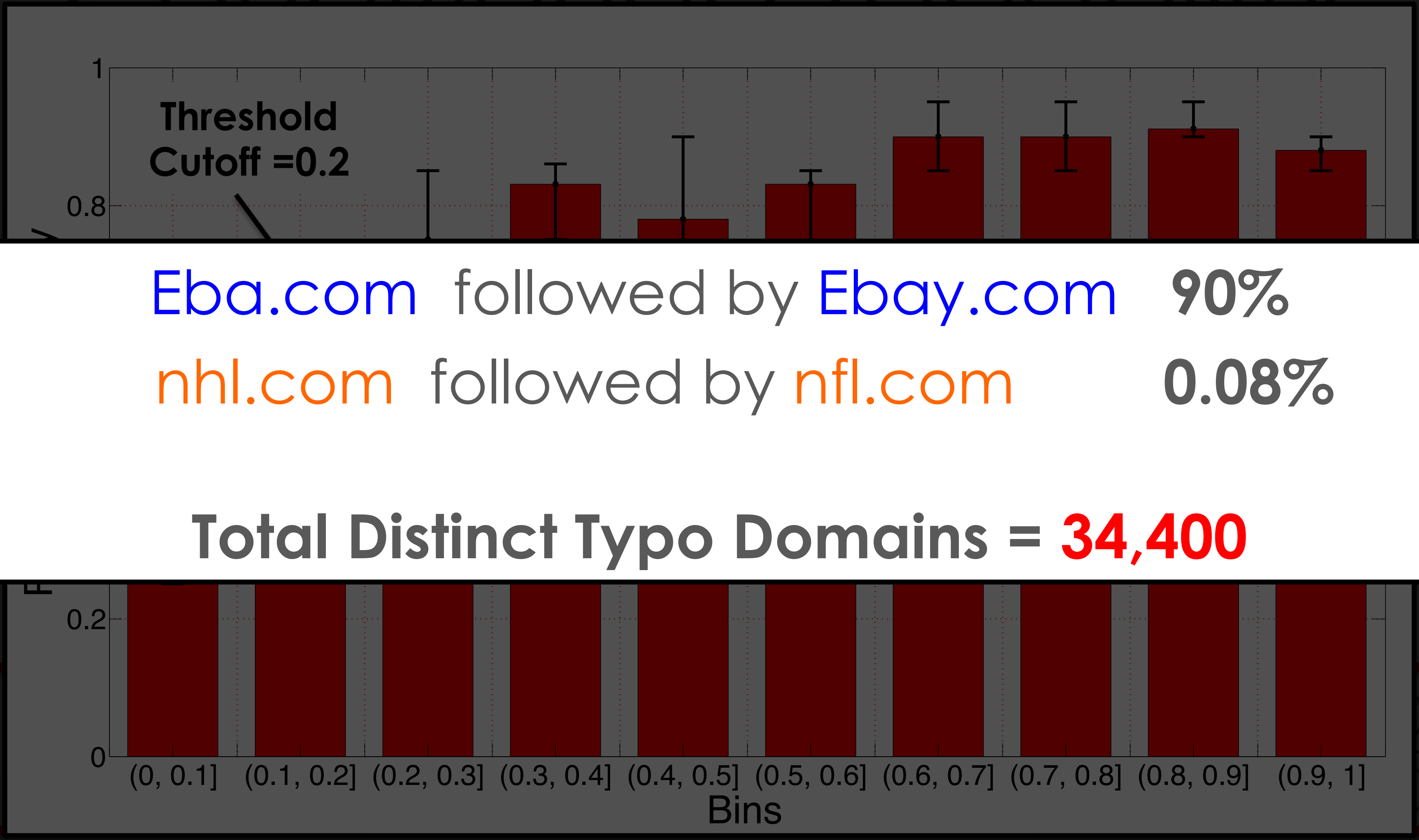
User I nt

- User Intent to visit the e generates similar pairs of domain

- User intent is manifested in various discovery methods



Conditional Probability



Threshold Cutoff =0.2

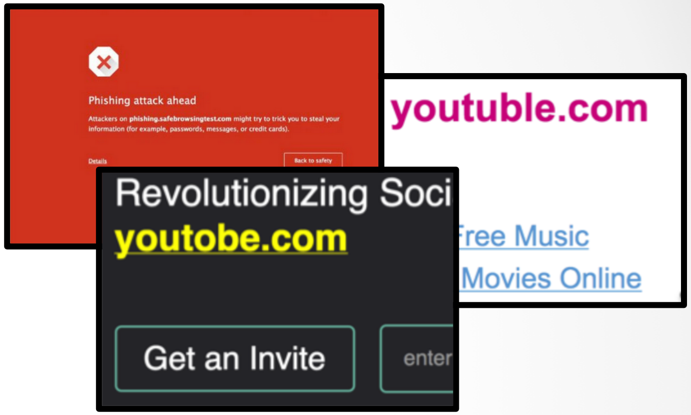
Eba.com followed by Ebay.com 90%

nhl.com followed by nfl.com 0.08%

Total Distinct Typo Domains = 34,400

Typo Characterization

- Adversarial registrations
  - Parked Domains
  - Malicious Websites
  - Other

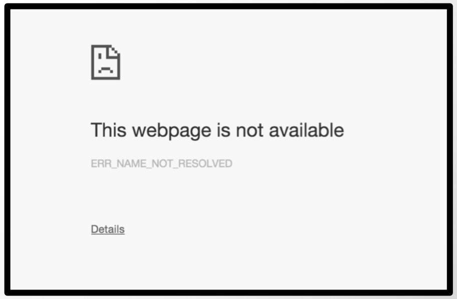


Phishing attack ahead

Attackers on **phishing.safebrowsingtest.com** might try to trick you to steal your information (for example, passwords, messages, or credit cards).

- Cooperative registrations
  - JavaScript and 3xx redirections
  - Defensive registrations

- Unregistered websites
  - NX Domains

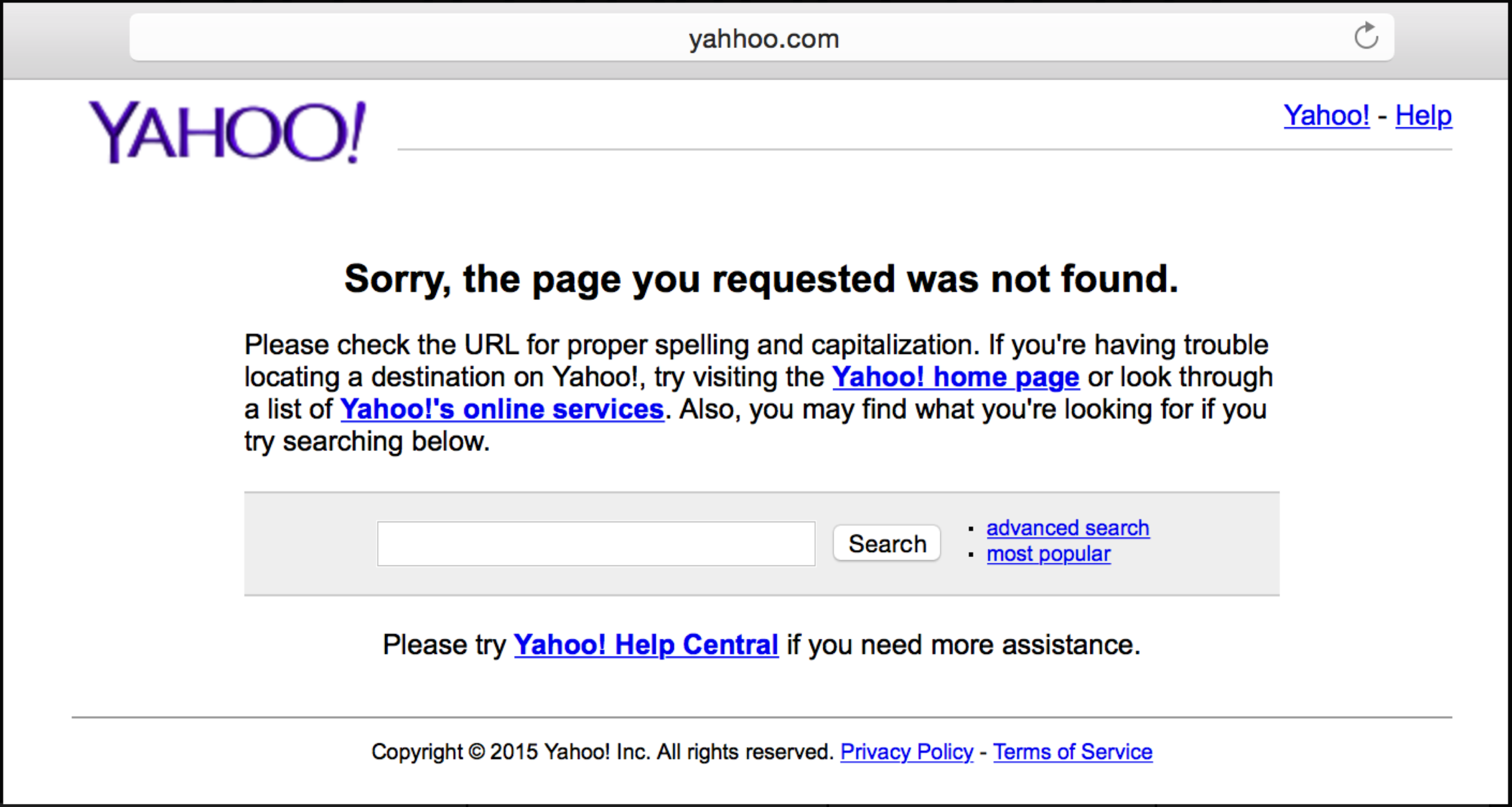


This webpage is not available

ERR_NAME_NOT_RESOLVED

What Next…??

Quantifying Harm



Sorry, the page you requested was not found.

Please check the URL for proper spelling and capitalization. If you're having trouble locating a destination on Yahoo!, try visiting the **Yahoo! home page** or look through a list of **Yahoo!'s online services**. Also, you may find what you're looking for if you try searching below.

Please try **Yahoo! Help Central** if you need more assistance.

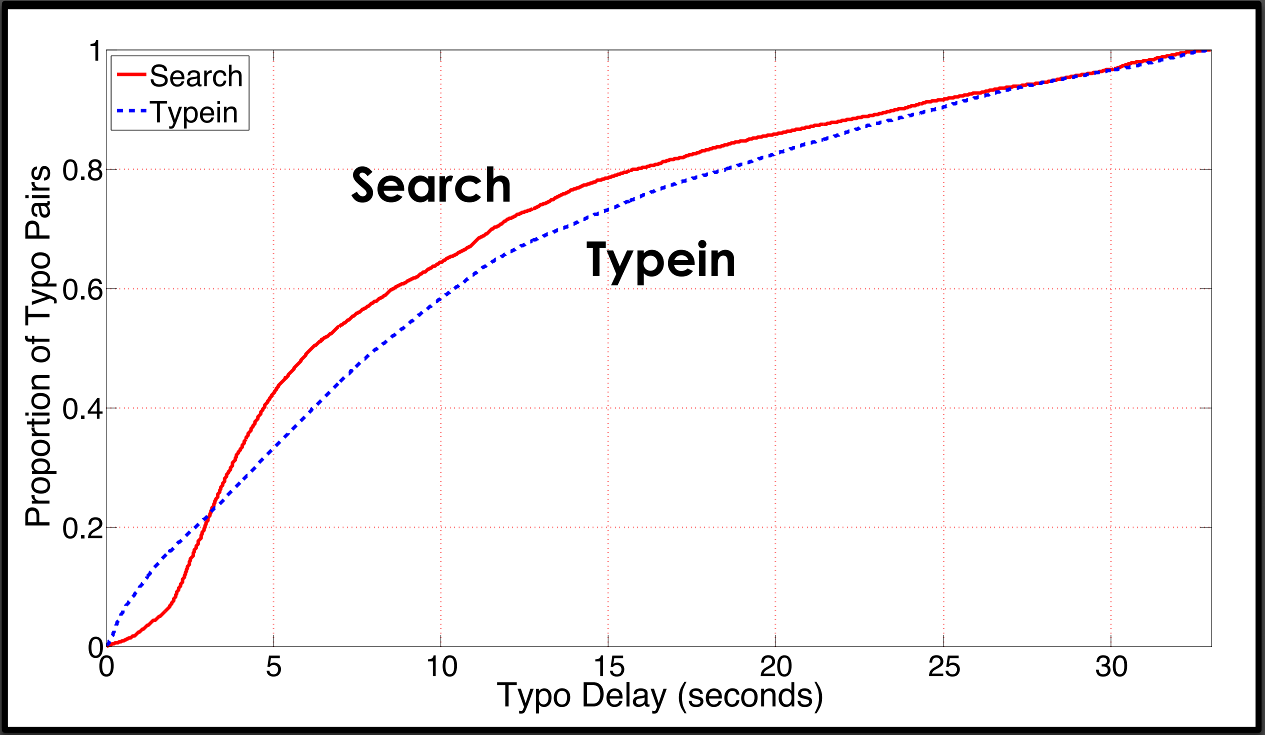
Copyright © 2015 Yahoo! Inc. All rights reserved. Privacy Policy - Terms of Service

	Cooperative	Adversarial	Unregistered
Average Delay (s)	2.87	9.58	10.38
Average User Loss (%)	3.30	16.81	11.53

Search VS Typein Delays

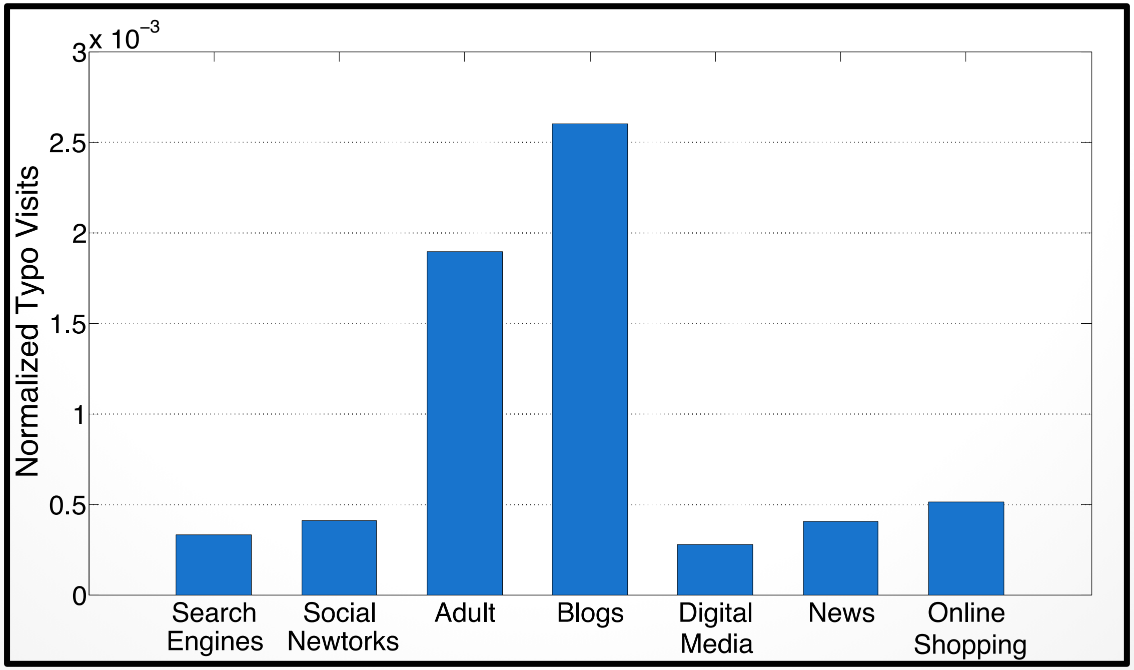
- Different discovery methods show varying delay trends

Unregistered Domains



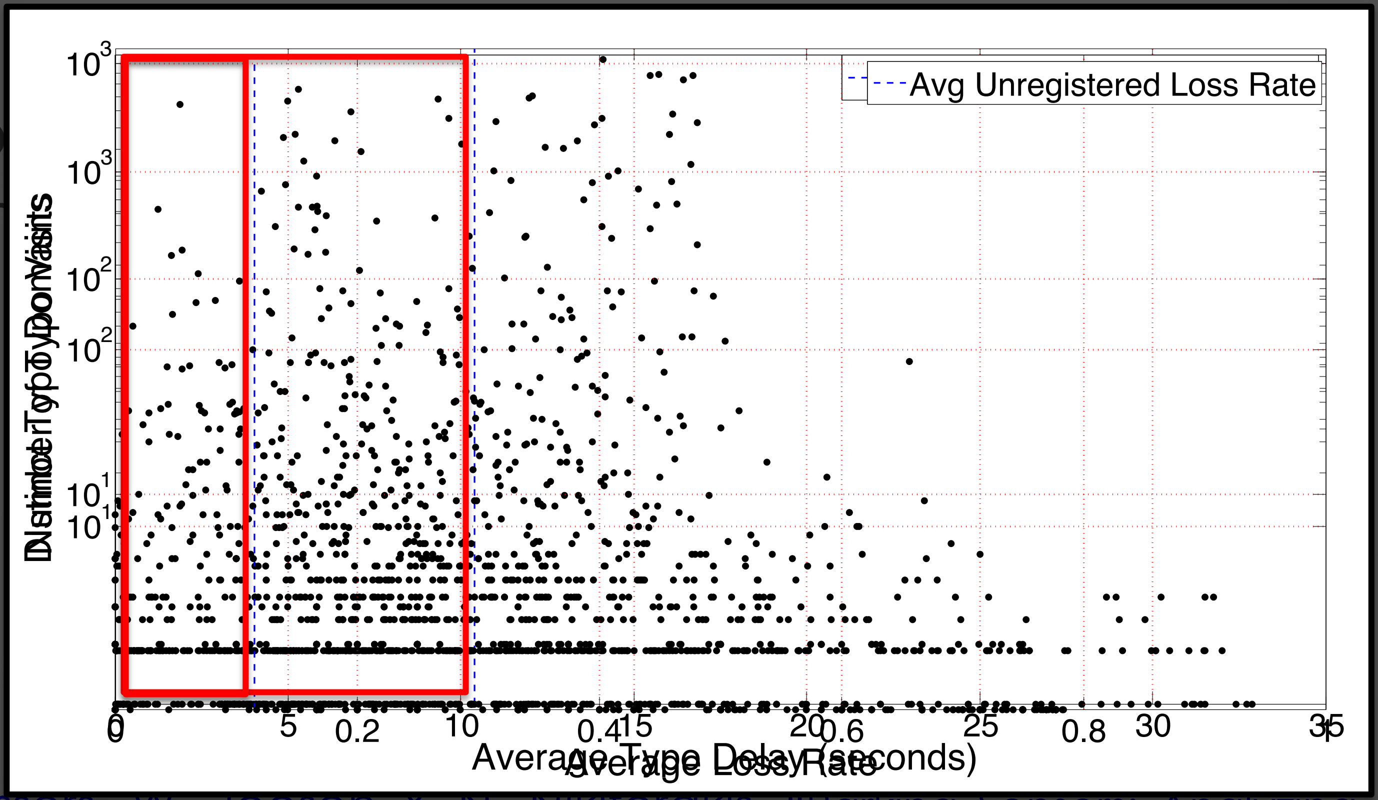
Target Domain Category

- Most Typos exist in the long tail of popularity
- Most distinct typos belonged to Adult and Blogs



Delay & Success Clustering

- Some typo domains help users to get faster to their destination websites
- Top
  dom



*[T. Vissers, W. Joosen & N. Nikiforakis, "Parking Sensors: Analyzing and Detecting Parked Domains". *NDSS 2015*]

Loss of Revenue

- Convert time and user loss into dollars.
- Intended site owner has a negative externality ratio of **18:1** against the typosquatter
- Using per capita income an average user loses **$0.29** to typosquatting per year
- For defenders, the effort ratio is **4.62:1**, far lower than non-violent crime*

*[J. M. Rao and D. H. Reiley, "The Economics of Spam," *The Journal of Economic Perspectives*, pp. 87–110, 2012.]

Conclusions

- Typosquatting is much less societally damaging than other non-violent crimes
- Defensive registrations do help against mistyping but not much against typosquatting
- Special technical or policy interventions are not necessarily required to deal with it

Thank You!

Qeustions?