Youtube.com
Fourteen Years of Typosquatting Research

- 2003: Edelman: first case study on one typosquatter
- 2006: Wang et al.: detection
- 2008: Banerjee et al.: detection
- 2009: Chen et al.: detection
- 2010: Moore and Edelman: monetization
- 2011: Banerjee et al.: detection
- 2014: Szurdi et al.: large scale study
- 2015: Agten et al.: longitudinal study and Khan et al.: quantifying harm to users
- 2017: Miramirkhani et al.: technical support scam
Other Applications Using DNS

• **Email:**

  ![Email example]

• **SSH:**

  ```
  $ ssh admin@secrett.com
  ```

• **FTP:**

  ```
  $ ftp admin@secrett.com
  ```

• **Godai group 2011:** white paper on email typosquatting

• **Vissers et al. 2017:** name server typosquatting
1. Email Typo Mistakes
   • What are the email typo mistakes users can make?

2. In the shoes of typosquatters
   • Do users make email typo mistakes frequently?

3. Typosquatting in the wild
   • Can typosquatters collect emails on a large scale?
   • How much emails typosquatting domains in the wild receive?

4. In the shoes of the victims
   • Do typosquatters actually collect emails?
Email Typo Mistakes
Receiver Typo

Write: all my passwords and credit card numbers

From: [Image]

To: mom@gmail.com

Subject: all my passwords and credit card numbers

Body Text

password123

6011 2222 4444 8888
Reflection Typo

MyUPMC

Date of birth
01/01/1917

Sex
Male  Female

Your date of birth is January 1, 1917, which makes you 100.

Last 4 digits of SSN
1000

Mobile phone
(optional)

Home phone
(optional)

Email
 typo@gmail.com

Log In
When Reflection Typos Are Really Bad

When mistake affects other users!

someone@zohomil.com: we received several
  • job applications
  • with CVs containing personal information

Several job advertisement copy pasted with the same mistyped address
SMTP Typo

SMTP Server

Settings
- Description: My favorite SMTP server
- Server Name: smtpx@gmail.com (with an error)
- Port: 587

Security and Authentication
- Connection security: STARTTLS
- Authentication method: Normal password
- User Name: victim

OK Cancel
In The Shoes of Typosquatters
Collection Ethics

IRB approved
• Took measures beyond IRB requirement

Registering typosquatting domains
• Potential trademark infringement
• On request surrender domains

Collecting personal emails
• Protect personal information
  • Keep on secure server
  • Encrypt emails
• Protect privacy
  • Remove sensitive data
  • Minimize the number of emails viewed
Collection Infrastructure

Registered domains

outlook.com

gmail.com

hotmail.com

smtpverizon.net

DNS “Forwarding”

Virtual Private Servers

SMTP Forwarding

Main Collection Server
Spam Filtering

- Header Based Filtering
- SpamAssassin
- Collaborative Spam Filtering
- Reflection Typo Detection
- Frequency-based filtering

Emails

Filtered emails
Receiver Typo Emails Collected

Infrastructure
Down
SMTP Typo Emails Collected

- Spam
- Real email typos

Log Number of emails received

Date (y/m/d)

16/06/01 16/07/01 16/08/01 16/09/01 16/10/01 16/11/01 16/12/01 17/01/01
Not All Typosquatting Domains Are Equal

Cumulative sum of receiver typo emails

Selected typosquatting domains

75%
## Typosquatting Domain Quality

<table>
<thead>
<tr>
<th>Domain</th>
<th># Emails</th>
<th>Is Fat Finger?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ohtlook.com</td>
<td>1320</td>
<td>TRUE</td>
</tr>
<tr>
<td>outlo0k.com</td>
<td>1170</td>
<td>TRUE</td>
</tr>
<tr>
<td>outmook.com</td>
<td>324</td>
<td>FALSE</td>
</tr>
<tr>
<td>ouulook.com</td>
<td>137</td>
<td>FALSE</td>
</tr>
<tr>
<td>oetlook.com</td>
<td>84</td>
<td>FALSE</td>
</tr>
<tr>
<td>ouvlook.com</td>
<td>25</td>
<td>FALSE</td>
</tr>
<tr>
<td>o/tlook.com</td>
<td>20</td>
<td>TRUE</td>
</tr>
<tr>
<td>ou6look.com</td>
<td>7</td>
<td>TRUE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Domain</th>
<th># Emails</th>
<th>Is Fat Finger?</th>
</tr>
</thead>
<tbody>
<tr>
<td>hovmail.com</td>
<td>1095</td>
<td>FALSE</td>
</tr>
<tr>
<td>ho6mail.com</td>
<td>147</td>
<td>TRUE</td>
</tr>
</tbody>
</table>

### Factors of profitability

- **Popularity** of target domain is the most important
- Keyboard distance
- **Conspicuousness**
Typosquatting In The Wild
Infrastructure Concentration: Registrants

One registrant: 10% of domains

45%
Infrastructure Concentration: Mail Server Records

One Mail Server Record: 14% of domains
Email Typosquatting Eco-system

High SMTP support
- Millions of typosquatting domains
- 2/3 of typo domains can receive emails

Infrastructure serving typosquatting
- Average name servers: 4% typosquatting
- Bad name servers: up to 89% typosquatting

Targeting email protocols
- 41 SMTP typos of Alexa top 10k
  - smtpgmail.com
  - smtphotmail.com
  Both privacy protected and typosquatting
Extrapolation

Model
• Based on our previous observations
• Features: Popularity, conspicuousness and keyboard distance

Extrapolate to
• 1211 typosquatting domains
• Targeting: gmail.com, hotmail.com, outlook.com, comcast.com, verizon.com

Estimate:
• 850,000 emails/year received

One email costs one penny to collect
• Ideal for spear phishing or scam campaigns
In The Shoes of The Victims
Hey guys,

I am sooo excited about our trip to Florida!!! :)
I put together the reservation and payment information. Please check everything is okay!

See you soon!!!

Ella
Hi Joel,

As discussed on the phone, here are the filled out 2016 tax forms for Paluzzo Contracting.
https://docsend.com/view/4dbz4jx

I look forward to hearing from you about possible optimization strategies.

Cheers,
Mike
Large Scale Test

Tested
• 50,000 typosquatting domains

Domains accepting our emails

<table>
<thead>
<tr>
<th>Domain registration type</th>
<th>Percent accepted our emails</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>14 %</td>
</tr>
<tr>
<td>Public registration</td>
<td>4 %</td>
</tr>
<tr>
<td>Private registration</td>
<td>27 %</td>
</tr>
</tbody>
</table>

Sensitive targets
• disvover.com, bankofamericqa.com, nuaghtyamerica.com and comcacst.com

Emails read
• 19 based on our logs
Sensitive Information Test

Tested
• 7269 domains
• previously accepted our email

Emails read
• 15 based on our logs

Sensitive information accessed
• Tax document accessed from Caracas Venezuela
• Shell account access attempt from Poland
Summary

• Users sent us emails with sensitive data
• Typosquatting domains’ profitability depends on
  • Popularity
  • Conspicuousness
  • Keyboard distance
• Typosquatters have infrastructure in place to collect emails
• One email costs one penny to collect
• Exploitation of email typosquatting is not confirmed

jszurdi@andrew.cmu.edu