The Long Taile of Typosquatting Domain Names

Janos Szurdi, Balazs Kocso, Gabor Cseh, Jonathan Spring, Mark Felegyhazi and Chris Kanich
Extent of Typosquatting

GOOGLE.COM

230,000 typos targeting the 3264 Alexa’s top .com domains

4.7 million typos targeting the 488,113 Alexa’s top .com domains

56 million potential typos targeting ALL .com domains

Alexa provides a list of domains ranked by popularity
Outline

• Why Typosquatting?

• How do we find Typosquatting domains?

• Trends in Typosquatting

• Possible defenses
The Target Domain: pncbank.com
Potential Typo Domain: pncbXnk.com
Potential Typo Domain: pncbanX.com
Potential Typo Domain: pncbanX.com
Potential Typo Domain: pncbanX.com

Android Tablet for only $1
+ 3-day trial to NowLucky.com

- 7” screen size
- 4GB data capacity
- WiFi + external 3G
- Front camera
- Earphone jack
- A13 processor

Get Your Android Tablet

Jmase
Bmd
James@secret.com
James@secret.com

I accept NowLucky's terms & conditions

Get My Tablet »
Potential Typo Domain: pncbanX.com
Potential Typo Domain: pmcbank.com
Ecosystem

Target Domain (pncbank.com)

Brand Protection

TypoSquatting Domain

98%

Competitor Brand

Other Domain

Phishing/Malicious

Ads

pmcbank.com
Potential Typo Domains

The Target Domain

pncbank.com

Typo Generation

- Deletion
- Addition
- Substitution
- Transposition

Generated Typo domains

- pncbXnk
- pncbaAnk
- pnZbank
- pnBCank
- pncbank

Potential Typo Domains

Selecting Registered Domains

- pncbnk
- pncban
- pmcbank
True Typo Domains

Potential Typo Domains

Heuristic Rule Set

DNS

Whois

Content

True Typo Domains

pncbnk

pncban

.

.

.

pmcbank

Not Typo Registrations

.

.

.

.

.
Manual Categorization

![Chart showing the proportion of typos in different sample sets.]

- **Sample set**
  - ALEXA-TOP
  - ALEXA-MID
  - ALEXA-TAIL
  - ZONE-SAMPLE

**Proportion of typos (in %)**
- ALEXA-TOP: 80%
- ALEXA-MID: 72%
- ALEXA-TAIL: 70%
- ZONE-SAMPLE: 75%
Accuracy

Based on previous work: Algo. 1 and Algo. 2  Ours: YATT
Trends in registration across time

>= 71% of typo domains remained registered for over one year
# BAD Name Servers

<table>
<thead>
<tr>
<th>Name Server</th>
<th>True Typos</th>
<th>All .com domains</th>
<th>Typo Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>aof.net</td>
<td>5221</td>
<td>6332</td>
<td>82%</td>
</tr>
<tr>
<td>citizenhawk.net</td>
<td>8819</td>
<td>12004</td>
<td>73%</td>
</tr>
<tr>
<td>easily.net</td>
<td>18281</td>
<td>36890</td>
<td>50%</td>
</tr>
<tr>
<td>domainingdepot.com</td>
<td>51854</td>
<td>132864</td>
<td>39%</td>
</tr>
<tr>
<td>next.org</td>
<td>9426</td>
<td>30252</td>
<td>31%</td>
</tr>
<tr>
<td>domainmanager.com</td>
<td>23493</td>
<td>90929</td>
<td>26%</td>
</tr>
</tbody>
</table>
One Name Server: 15% of typos

12 Name Servers: 50% of typos

6 Registrars: 50% of typos

One Registrars: 25% of typos
Tool

Ooops!
The URL you entered might be a typo
Did you want to go to here instead?
www.facebook.com

Or you really wanna go there?
www.faebook.com

Why do you see this page?
The domain you tried to reach is probably a typosquatting domain

To be on the safe side
Takeaways

• Typosquatters invest significant resources

• Typosquatting steadily increases

• Typosquatting could be contained

jszurdi@andrew.cmu.edu