

A Study of Whois Privacy and Proxy Service Abuse

Richard Clayton

richard.clayton AT cl.cam.ac.uk

Tony Mansfield

tony.mansfield AT npl.co.uk



UNIVERSITY OF
CAMBRIDGE

Computer Laboratory



National Physical Laboratory

Overall summary

- When criminals register domain names for use in online criminality they don't provide their names and addresses
- We collected a lot of data to show that this is generally true
- In doing so we found that the way in which contact details are hidden varies somewhat and this gives new insights
- Meantime, people calling for changes to domain "privacy" and "proxy" services "because they are used by criminals" must understand:
 - the impact of such a policy change on other registrants
 - the limitations of such a policy change on criminals

Normal Whois Data

- When a domain name is registered the registrant supplies their name and contact details (street address, perhaps phone & email)
- Other fields give admin/billing/technical/etc. contacts
 - one can often learn registrant phone numbers if the registrant is also admin/billing/etc.
- This data is public
 - and available on the port 43 whois service
 - also sometimes on the web as well
- Whois allows problems to be addressed promptly
 - but some people are shocked by the lack of privacy

Privacy and Proxy Services

- Privacy Service
 - registrant name is provided, but contact details are generic (although sometimes the local part of the email address is specific to the registrant – to allow automated forwarding of email)
- Proxy Service
 - domain is registered in the name of the proxy service and all contact details are generic (although sometimes the local part of the email address is specific to the registrant – to allow automated forwarding of email)
- Note that for “.UK” Whois data may be hidden by individual choice (but not by traders or companies)
 - but .UK isn't one of the domains ICANN looks after

Example Proxy Registration

Domain Name: DOOMZONE.NET

Registrant:

PrivacyProtect.org

Domain Admin (contact@privacyprotect.org)

ID#10760, PO Box 16

Note - All Postal Mails Rejected, visit Privacyprotect.org

Nobby Beach

null,QLD 4218

AU

Tel. +45.36946676

Creation Date: 07-Feb-2012

Expiration Date: 07-Feb-2013



 **POST** Nobby Beach 4218 





15

25

16

26

Summary of Methodology for Study

- Basic approach:
 - obtain various lists of criminal URLs
 - pick out domains being used
 - fetch Whois data for the biz/com/info/net/org domains
 - assess whether registrant is using privacy or proxy service
 - OR look for contact phone number of registrant
- Precise stats for privacy/proxy/no phone number
- Random sample of registrants with phone number
 - phone call made; if answered then one question survey (in registrant's native language)
 - “did you register example.com”
 - if not answered then retried on different days/times

Phone Results

- Phone number had to be “apparently valid” (i.e. have enough digits, not be 99999999 or 00000000, or have an invalid North American area code)
 - BUT could turn out to be invalid when we dialled it
 - OR the number was valid but just rang and rang
 - OR we reached voicemail, or someone answered who could not help us reach the registrant, or registrant wasn't ever available
 - OR phone answered and knowledge of domain denied
 - OR we spoke to the registrant (or someone speaking for a company) and they agreed they had registered the domain

Phone Results

• **NOPHONE** unless “apparently valid”

enough digits, not be 9999999 or 0000000, or have an invalid North American area code)

▪ BUT could turn out to be valid if we dialed it

▪ OR the phone was valid but just a bad string

▪ OR we reached voicemail, or someone answered who could not help us (i.e. the registrant, or registrant wasn't ever available)

▪ OR phone answered but domain denied

▪ OR we spoke to the registrant (or someone speaking for a company) and they agreed they had registered the domain

Treated as failure

Neither success nor failure

Neither success nor failure

Treated as failure

Treated as success

Phishing (the report in a nutshell)

- Phishing (i.e. email enticing to web page...)
- Source data was 32 806 URLs (one week's worth), using 5 105 domains – 57% in biz/com/info/net/org/
- Used specialist knowledge to split these into three groups:
 - compromised machines (i.e. criminal added phishing pages)
 - 2121 domains
 - third parties (free webhosting domains, cloud services, etc.)
 - 263 domains (plus 1 had no Whois available, so ignored)
 - maliciously registered domain names
 - 449 domains (plus 5 had no Whois data available)

Phishing Analysis Results

- Privacy and proxy usage
 - third parties 14% low
 - compromised machines 25% average
 - maliciously registered domains 31% high
- Able to reach registrant by phone
 - third parties 32%
 - compromised machines 24%
 - maliciously registered domains 2%
- No hope of reaching registrant by phone
 - third parties 50%
 - compromised machines 62%
 - maliciously registered domains 92%

Other Types of Malicious Registration

- WP2: Data from aa419.org (Advanced Fee Fraud &c)
 - 46% of registrants using privacy/proxy services
 - 89% impossible, a priori, to contact by phone
- WP3: Unlicensed pharmacies
 - 55% of registrants using privacy/proxy services
 - 92% impossible, a priori, to contact by phone
- WP5: Child sexual abuse image websites
 - 29% of registrants using privacy/proxy services
 - it is believed that 100% are impossible to contact by phone
- So a range of rates of usage of privacy/proxy services, but criminals seldom contactable by phone

Legal and Harmless Categories

Category	Privacy/ proxy usage	impossible to reach by phone	Did reach by phone [*]
Legal pharmacies	9%	24%	24%
Law firms	13%	34%	25%
Executive search consultants	22%	37%	33%
Banks	28%	45%	15%
Alexa top 3500 (being typo-squatted)	19%	47%	29%
Adult websites	44%	55%	6%

* CAVEAT: small samples mean quite large error bounds for this column

The Story So Far...

- Average usage of privacy/proxy services:
 - 20% NORC measurement across all domains
 - 25% our measure of compromised websites
- Criminals use these services more than average
 - ranges from 29% to 55%
 - BUT some harmless activities also above average too
 - banks 28%, adult websites 44%
- Criminals don't reveal contact phone numbers. So consider the *a priori* "impossible to contact" rates
 - ie usage privacy/proxy or bad/missing phone number rates
 - criminal activities: 88% – 92% (perhaps 100%)
 - legal and harmless: 24% – 62%

More Complex Datasets

- WP8: StopBadware (malware related domains)
 - Mainly compromised sites, but some malicious registrations
 - 20% of registrants use privacy/proxy services
 - But 51% not possible to reach by phone
- WP8: SURBL (domains indicating email is spammy)
 - Mainly maliciously registered, but by no means all
 - 44% of registrants use privacy/proxy services
 - but only 59% not possible to reach by phone
 - CAUTION: high error bounds with this dataset because many domains had the same contact phone number
 - ALSO: some evidence of report inflation, i.e. all possible domains listed when multiple domains can be resolved to same location

Typosquatting

- Already mentioned “typosquatted domains” : Alexa 3500 sites where small variants of domain name exist hoping to be visited by sloppy tpyers
- WP4: typoquatting domains
 - privacy/proxy services used by 48% of registrants
 - 11% reached by phone (c.f. adult websites 6%)
 - BUT very high error bounds (small number of people involved)
- Clearly some typosquatters are attempting to avoid being identified, whereas others are more laid back
 - NB this isn't criminal – but civil action is more likely if the brand owner can identify “economies of scale”

Summary of Numerical Results of Study

Work package	Privacy or proxy usage	Not possible to call registrant	Maliciously registered?
Legal pharmacies	8.8%	24.2%	no
Law firms	13.4%	33.6%	no
Executive search consultants	22.4%	36.7%	no
Banks	28.2%	44.6%	no
Typosquatted domains	19.2%	47.1%	no
Phishing: third parties	13.7%	49.6%	no
StopBadware domains	20.4%	51.4%	some
Adult websites	44.2%	55.1%	no
SURBL domains	44.1%	58.5%	mostly
Phishing: compromised sites	24.7%	61.7%	no
Typosquatting	48.2%	67.7%	yes
Advanced Fee Fraud	46.5%	88.9%	yes
Unlicensed pharmacies	54.8%	91.8%	yes
Phishing: malicious registration	31.2%	92.5%	yes

Summary of Findings

- Criminals DO use privacy/proxy services > average
- BUT so do some legal and harmless activities as well
- When criminals don't use privacy/proxy services then they don't provide valid contact numbers – so overall the effect is that at least 9/10 can't be reached
- BUT many lawful and harmless activities fail to provide valid contact numbers either, with anything between a quarter and two third of them being inherently unreachable
- BUT the Whois phone number is not the only way to reach legitimate registrants...

Policy Conundrums

- Study shows (recall the typosquatting, the adult websites and the banks) that the reasons for using privacy and proxy services are many and various...
- Some people believe that privacy / proxy services are so abused that they should be forbidden
 - BUT many legitimate businesses & individuals are using them
 - clearly criminals will just fail to provide valid contact details
- Some people want compulsion to provide valid contact details (and these should be checked)
 - BUT between a quarter and two thirds of existing legitimate domain registrations don't provide valid contact details so hard to get there from here!

*A Study of
Whois Privacy and Proxy Service
Abuse*

<http://www.lightbluetouchpaper.org>



UNIVERSITY OF
CAMBRIDGE

Computer Laboratory



National Physical Laboratory