

# You're Going to Connect to the Wrong Domain

@erbbysam

whoami

@erbbysam

Software Engineer

DC23, DC24 black badge (Badge Challenge, Co9)



The opinions expressed here are my own &  
no one was phished in the creation of this presentation



# Typosquatting

Humans are not perfect, we mistype domain names. A malicious person can register these mistyped domains.

Only 2 of the 61 1-keyboard letter off variants of americanexpress.com are unregistered.

## Example:

qmericanexpress.com

Expires:2018-01-06T00:00:00

Created:2008-01-06T00:00:00

Updated:2017-01-06T00:00:00

Registrar:ABOVE.COMTPTYTLTD.

americanexpress

snweuxsbwzoewaa

wjrtovwmrcltrdd

qksdjdhss dsee

z dfkfbzjdd fdww

xx

zz

# Bitsquatting

A form of typosquatting, but one where the computer gets the domain name wrong by flipping a bit.

eoogole-analytics

01100101 01101111 01101111 01100111 01101100 01100101 00101101 01100001 01101110 01100001 01101100 01111001 01110100 01101001 01100011 01110011  
01100111 01101111 01101111 01100111 01101100 01100101 00101101 01100001 01101110 01100001 01101100 01111001 01110100 01101001 01100011 01110011

google-analytics

## fooogle-analytics.com Expires:2017-09-14T00:00:00 Created:2011-09-14T00:00:00 Updated:2016-08-13T00:00:00 Registrar:MARKMONITORTINC.

google-analytics.com Expires:2017-09-14T00:00:00 Created:2011-09-14T00:00:00 Updated:2016-08-13T00:00:00 Registrar:MARKMONITORTINC.

gnogle-analytics.com Expires:2017-09-14T00:00:00 Created:2011-09-14T00:00:00 Updated:2016-08-13T00:00:00 Registrar:MARKMONITORTINC.

gongle-analytics.com Expires:2017-09-13T00:00:00 Created:2011-09-13T00:00:00 Updated:2016-08-12T00:00:00 Registrar:MARKMONITORTINC.

gokgle-analytics.com Expires:2017-09-14T00:00:00 Created:2011-09-14T00:00:00 Updated:2016-08-13T00:00:00 Registrar:MARKMONITORTINC.

google-analytics.com Expires:2017-09-14T00:00:00 Created:2011-09-14T00:00:00 Updated:2016-08-13T00:00:00 Registrar:MARKMONITORTINC.

googde-analytics.com Expires:2018-03-22T00:00:00 Created:2016-03-22T00:00:00 Updated:2017-04-15T00:00:00 Registrar:ENOM,TINC.

googlg-analytics.com Expires:2016-03-22T00:00:00 Created:2016-03-22T00:00:00 Updated:2017-04-15T00:00:00 Registrar:ENOM,TINC.

googlm-analytics.com Expires:2018-03-22T00:00:00 Created:2016-03-22T00:00:00 Updated:2017-04-15T00:00:00 Registrar:ENOM,TINC.

googleanalytics.com Expires:201-08-05T00:00:00 Created:2013-08-05T00:00:00 Updated:2016-08-09T00:00:00 Registrar:GODADDY.COM,LLC

google-analytics.com Expires:2016-09-14T00:00:00 Created:2011-09-14T00:00:00 Updated:2016-08-13T00:00:00 Registrar:MARKMONITORINC.

# Bitsquatting - Example

- Registered eoogole-analytics.com
- Used Let's Encrypt to get a TLS certificate
- Found a misconfigured server, but within 24 hours saw 2 “hits”

```
174 <script type="text/javascript">
175 var gaJsHost = (("https:" == document.location.protocol) ? "https://ssl." : "http://www.");
176 document.write(unescape("%3Cscript src='" + gaJsHost + "eoogole-analytics.com/ga.js' type='text/javascript'%3E%3C/script%3E"));
177 </script>
178 <script type="text/javascript">
179 try {
180 var pageTracker = _gat._getTracker("          ");
181 pageTracker._trackPageview();
182 } catch(err) {}</script>
183
```

HTTPServerRequest(protocol='http', host='www.eoogole-analytics.com',  
method='GET', uri='/r/\_utm.gif?[removed]', version='HTTP/1.1',  
remote\_ip='[removed]', headers={'Accept-Language': 'ja-JP, en-US;q=0.8',  
'Accept-Encoding': 'gzip, deflate', 'X-Wap-Profile':  
'http://[removed].com/[removed].xml', 'X-Getzip': 'supported', 'Host':  
'www.eoogole-analytics.com', 'User-Agent': '[removed]', 'Accept-Charset': 'utf-8,  
iso-8859-1, utf-16, \*;q=0.7', 'Connection': 'keep-alive', 'X-Requested-With':  
'com.android.browser', 'Referer': 'http://[removed].net/823.html', 'Cache-Control':  
'no-cache', 'Cookie': 'VisitorID=[removed]&Exp=11/12/2018 9:13:02 AM'})

HTTPServerRequest(protocol='https',  
host='www.eoogole-analytics.com', method='GET', uri='/analytics.js',  
version='HTTP/1.1', remote\_ip='[removed]', headers={'Save-Data':  
'on', 'Accept-Language': 'en-US,en;q=0.8', 'Accept-Encoding': 'gzip,  
deflate, sdch, br', 'Host': 'www.eoogole-analytics.com', 'Accept': '\*/\*',  
'User-Agent': 'Mozilla/5.0 (Linux; Android 5.0.2; P01V  
Build/LRX22G) AppleWebKit/537.36 (KHTML, like Gecko)  
Chrome/58.0.3029.83 Safari/537.36', 'Connection': 'keep-alive',  
'Referer': 'http://www.[removed].lk/channel'})



# Bitsquatting - gTLD

.got

01100111 01101111 01110100  
01100111 01101111 01110110

.gov

.bom

01100010 01101111 01101101  
01100011 01101111 01101101

.com



# IDN Homoglyphs

IDN = “Internationalized domain name”, stored as punycode (xn--\*)

Homoglyphs are 2 characters that look the same

Example:

xn--ggle-55da.com = g**oo**gle.com

xn--e1anr4f.com = **T**ime.com

# IDN Homoglyphs - Identification

I wanted to conduct a survey of existing IDN homoglyph domains against popular .com domains...

3 options to gather domain names:

- 1) zone files (hard/impossible to acquire)
- 2) Certificate Transparency
- 3) Third party lists (\$\$\$)

Google “pilot” CT log contains  
~100 million certificates

- 400GB (compressed)
- searchable (crt.sh)
- A great source of data

# IDN Homoglyphs - Identification Continued

Another reason for using Certificate Transparency - if a certificate was registered, the domain was more likely to be used

Let's build a pipeline:

[Google CT Pilot log] → [parse CN, SAN domains] → [filter punycode .com domains (ex. xn--\*.com)]

Cross Reference:

- 1) Pipeline list, rendered as unicode, passed through the python unidecode package (ex. P → P)
- 2) Alexa top 1 million domains

End result:

1,938 CT certificates containing impersonating domains, modified Chromium unit test for punycode display status  
[TODO - insert github link]

# IDN Homoglyphs - Cross Referenced Results

κ, 22, 0x138, "LATIN SMALL LETTER KRA"

96074858, 1509667199, xn--faceboo-jhb.com, facebookκ.com , κ, [facebook.com](https://facebook.com), 3, 1

86142753, 1507679999, xn--autodes-jhb.com, autodesκ.com , κ, [autodesk.com](https://autodesk.com), 697, 1

ł, 5, 0x142, "LATIN SMALL LETTER L WITH STROKE"

94011919, 1524055021, xn--ppe-8ka60c.com, àpple.com , ł, [apple.com](https://apple.com), 69, 1

94724468, 1500291180, xn--sack-01a.com, slack.com , ł, [slack.com](https://slack.com), 205, 1

ı, 100, 0x131, "LATIN SMALL LETTER DOTLESS I"

18331655, 1488327078, xn--reddt-q4a.com, reddıt.com , ı, [reddit.com](https://reddit.com), 7, 1

95900673, 1500493680, xn--t-fka.com, ti.com , ı, [ti.com](https://ti.com), 3235, 1

84518766, 1497998760, xn--gml-kua34j.com, gmàıl.com , ı, [gmail.com](https://gmail.com), 22463, 1

95900424, 1500493860, xn--fat-jua.com, fıat.com , ı, [fiat.com](https://fiat.com), 54102, 1

94504694, 1509148799, xn--curacao-egamng-hgc.com, curacao-egamıng.com , ı, [curacao-egaming.com](https://curacao-egaming.com), 524456, 1

94724500, 1500493920, xn--suzu-kza.com, ısuzu.com , ı, [isuzu.com](https://isuzu.com), 866480, 1

ì, 25, 0xec, "LATIN SMALL LETTER I WITH GRAVE"

95900680, 1500670920, xn--twtr-7raz.com, twittèr.com , ì, [twitter.com](https://twitter.com), 11, 1

85019386, 1507161599, xn--polonex-3ya.com, polonìex.com , ì, [poloniex.com](https://poloniex.com), 1595, 1

83724035, 1497798600, xn--gma-pma40b.com, gmailì.com , ì, [gmail.com](https://gmail.com), 22463, 1

# IDN Homoglyphs - More Cross Referenced Results

2 interesting domains observed bypasses Chromium checks by using only cyrillic characters:

07022746, 1443571199, xn--80aac5cct.com, **taoobao**.com , таобао, taobao.com, 10, 1

10303999, 1461542399, xn--e1anr4f.com, **time**.com , time, time.com, 817, 1

# IDN Homoglyphs - Results Contd

A breakdown of the unicode blocks observed:

89 LATIN

25 CYRILLIC

16 HEBREW

14 GREEK

8 KATAKANA

4 ARABIC

3 HANGUL

2 HIRAGANA

1 RUNIC

1 MALAYALAM

**1 CANADIAN SYLLABICS**

# IDN Homoglyphs - Canadian Aboriginal Syllabics

Firefox & Chromium IDN checks were bypassed(punycode value was not displayed) for the following sample domains. Firefox & Chromium security bugs were reported.

`http://xn--youtue-084a.com/` -- `youtuᐅe.com` -- example domain  
`http://xn--youtbe-z72a.com/` -- `youtᐅbe.com` -- example domain  
`http://xn--uny-8wq.com/` -- `ᐅuny.com` -- example domain  
`http://xn--oor-hxq.com` -- `ᐅoor.com` -- example domain  
`http://xn--ego-73q.com/` -- `ᐅego.com` -- example domain  
`http://xn--fc-lym.com/` -- `fcᐅ.ᐅcom` -- example domain  
`http://xn--ulu-7sr.com/` -- `ᐅulu.com` -- example domain  
`http://invalid.xn--acebook-yp9a.com/` -- `ᐅacebook.com` -- example domain

Chromium - CVE-2017-5076


Firefox - CVE-2017-7764



# IDN Homoglyphs - Policy to the Rescue

While a domain name may render in a browser, you may not be able to register it!

<https://www.verisign.com/assets/idn/idn-canadian-aboriginal.html>

 VERISIGN

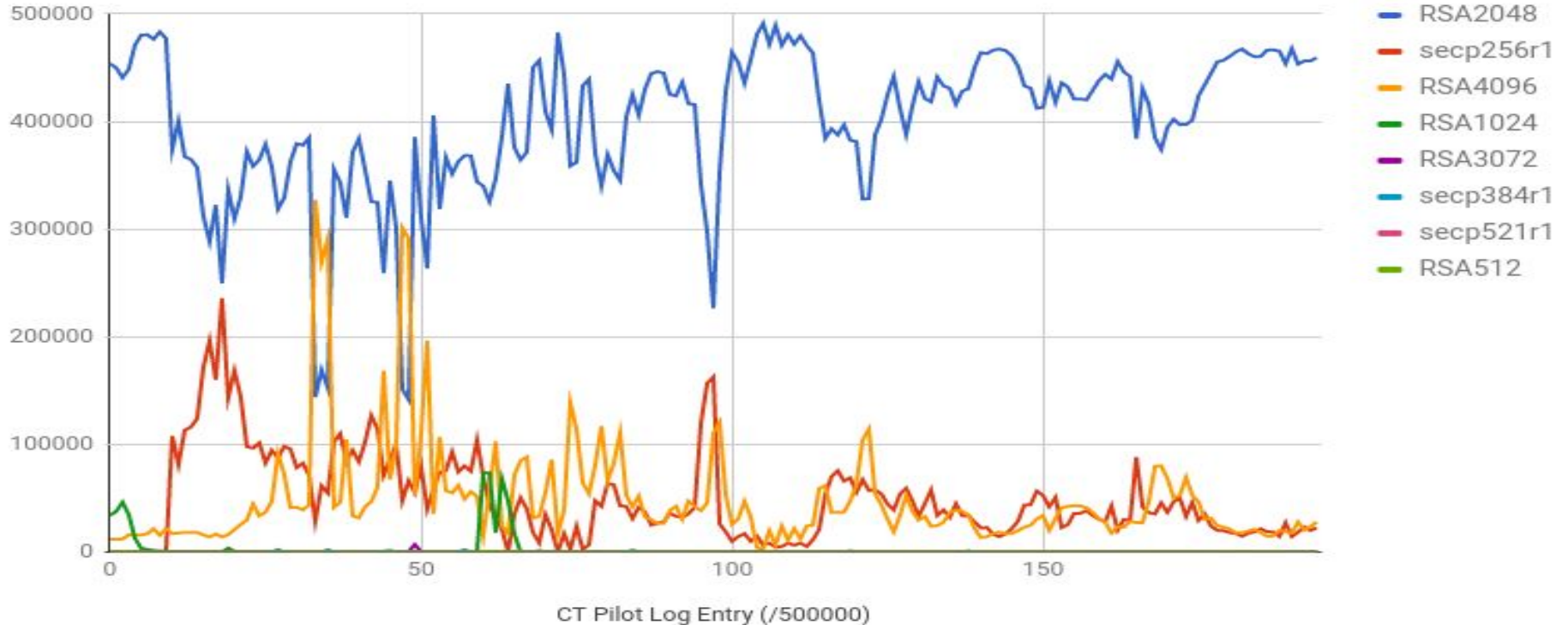
BACK TO VERISIGN, INC. <

Registration Rules  
Canadian Aboriginal

Registry	Verisign Inc.
Script	Canadian Aboriginal
Version	1.0

# Certificate Transparency Fun - Graph of Key Types

**Key Types Observed Over Time**



# Certificate Transparency Fun - Most Common Key Types

RSA2048	78019787
secp256r1	9556582
RSA4096	9447685
RSA1024	484262
RSA3072	45921
secp384r1	39336
RSA512	3026

78019787 RSA2048	10 RSA2050	2 RSA4000	1 RSA3819
9556582 secp256r1	10 RSA2028	2 RSA3957	1 RSA3817
9447685 RSA4096	9 RSA511	2 RSA3925	1 RSA3779
484262 RSA1024	9 RSA3120	2 RSA3892	1 RSA3629
45921 RSA3072	9 RSA3000	2 RSA3210	1 RSA3400
39336 secp384r1	9 RSA2078	2 RSA3092	1 RSA3336
3026 RSA512	8 RSA4906	2 RSA2890	1 RSA3328
2429 RSA8192	8 RSA3087	2 RSA2481	1 RSA3224
1847 RSA2432	8 RSA2304	2 RSA2400	1 RSA3200
418 DSA2048	7 RSA4192	2 RSA2222	1 RSA3163
314 RSA4056	7 RSA2043	2 RSA2182	1 RSA3132
229 RSA1023	7 RSA1280	2 RSA2148	1 RSA3124
226 RSA3248	6 RSA8000	2 RSA2142	1 RSA3103
217 RSA2560	6 RSA4100	2 RSA2136	1 RSA3102
213 RSA2084	6 RSA3073	2 RSA2128	1 RSA3100
195 RSA2047	6 RSA2612	2 RSA2098	1 RSA3098
184 RSA2056	6 RSA2040	2 RSA2087	1 RSA3070
166 RSA2049	5 RSA3584	2 RSA2086	1 RSA3052
153 secp521r1	5 RSA3456	2 RSA2060	1 RSA3049
153 RSA4092	5 RSA3333	2 RSA2042	1 RSA3047
146 RSA3096	5 RSA3192	2 RSA2038	1 RSA3028
131 RSA4048	5 RSA2176	2 RSA2014	1 RSA2999
129 RSA4098	5 RSA1234	2 RSA1924	1 RSA2942
127 RSA16384	4 RSA8092	2 RSA1825	1 RSA2857
124 RSA4086	4 RSA4089	2 RSA16348	1 RSA2685
118 RSA4069	4 RSA4068	2 RSA1204	1 RSA2642
110 RSA1536	4 RSA4024	2 RSA1026	1 RSA2600
72 RSA1048	4 RSA3600	1 RSA9216	1 RSA2580
68 RSA768	4 RSA3128	1 RSA8888	1 RSA2549
65 RSA2058	4 RSA3071	1 RSA8184	1 RSA2344
64 RSA2408	4 RSA3027	1 RSA8182	1 RSA2342
63 RSA2096	4 RSA2066	1 RSA8172	1 RSA2319
52 RSA3024	4 RSA2052	1 RSA7168	1 RSA2291
41 RSA4095	4 RSA2051	1 RSA7094	1 RSA2240
37 RSA3076	4 RSA2045	1 RSA7024	1 RSA2220
30 RSA4046	4 RSA1369	1 RSA5487	1 RSA2190
27 RSA4196	4 RSA1042	1 RSA5192	1 RSA2175
24 RSA8096	4 RSA1034	1 RSA5096	1 RSA2160
23 RSA2064	4 RSA1028	1 RSA5048	1 RSA2146
22 RSA2046	3 RSA4090	1 RSA5001	1 RSA2111
21 RSA5120	3 RSA3702	1 RSA5000	1 RSA2094
21 DSA1024	3 RSA3172	1 RSA500	1 RSA2088
20 RSA2345	3 RSA3080	1 RSA4608	1 RSA2068
20 RSA2024	3 RSA30720	1 RSA4321	1 RSA2059
19 RSA8196	3 RSA3050	1 RSA4198	1 RSA2057
19 RSA4094	3 RSA2536	1 RSA4099	1 RSA2053
19 RSA3768	3 RSA2480	1 RSA4080	1 RSA2044
19 RSA2736	3 RSA2054	1 RSA4076	1 RSA2039
17 RSA2080	2 RSA9192	1 RSA4072	1 RSA2018
16 RSA1025	2 RSA8392	1 RSA4065	1 RSA2010
15 RSA4088	2 RSA8191	1 RSA4013	1 RSA16383
14 RSA6144	2 RSA7680	1 RSA4007	1 RSA16318
14 RSA4097	2 RSA6095	1 RSA4006	1 RSA1548
14 RSA15360	2 RSA5012	1 RSA3983	1 RSA1506
13 RSA3048	2 RSA4611	1 RSA3972	1 RSA13999
12 RSA4028	2 RSA4444	1 RSA3971	1 RSA1027
12 RSA15424	2 RSA4114	1 RSA3931	1 RSA10240
11 RSA2948	2 RSA4084	1 RSA3904	1 RSA1000
11 RSA1212	2 RSA4082	1 RSA3889	1 DSA512
10 RSA2848	2 RSA4042	1 RSA3875	

# Questions?

Contact:

@erbbysam

very@busy.business