Internationalized Domain Name Homograph Attacks

CSE 227: Computer Security - Spring 2017
University of California San Diego

Chen Lai chl588@ucsd.edu
Zhongrong Jian zhjian@ucsd.edu
J. Sidrach jsidrach@ucsd.edu

Abstract

TODO

1 Introduction

TODO - Introduction to the problem - Enumeration of sections this paper talks about

2 Background

TODO - DNS: explanation - IDN: history, explanation - Browsers: display of IDNs, algorithms

3 Related Work

TODO - Brief analysis of previous papers on the same topic

4 Methodology

TODO - Introduction to the methodology - Reference github repository - Subsection: data collection - Reference sources of data - Subsection: data preprocessing - Subsection: clustering, confusables algorithm - Caveats: not mathing /, .com, etc. - Subsection: manual classification, guidelines used

Domains	#	%
Canonical domain names	*458731*	*8.31%*
With IDN homographs	825	6.04%
Without IDN homograpphs	457906	2.27%
International Domain Names	*1045400*	*91.69%*
With canonical homograph	1099	3.68%
Without canonical homograph	1044301	2.74%

Table 1: Overview of the clustering results.

5 Results

TODO - Comment different results obtained - Explain consequences of caveats (low number of matches) - Explain tables - Some other interesting results (topic-related domain hoarding)

TODO: mention: Number of Third Party domains whose Registrant Organization and Email has than one homograph IDN registered: 437 Number of Registrant's Organization and Email that have more than one homograph IDN: 82

6 Ethical Considerations

TODO - Brief explanation why this research is ethical

7 Conclusions

TODO - Conclusions of our work - Possible future work

TODO DELETE [1][1]

Domain	# of IDN homographs
google.com	24
youtube.com	3
facebook.com	9
baidu.com	3
yahoo.com	4
reddit.com	1
qq.com	2
taobao.com	1
live.com	1
vk.com	6

Table 2: Top ten .com domains in the Alexa ranking with IDN homographs.

Status	#	%
Canonical	*88*	*8.31%*
Parking	64	6.04%
Redirect	24	2.27%
Third Party	*971*	*91.69%*
Redirect to Canonical	39	3.68%
Unrelated	29	2.74%
Parking	872	82.34%
Scam	31	2.93%

Table 3: Breakdown of the manually classified homograph IDNs.

Acknowledgements

We would like to thank Louis DeKoven and Stefan Savage for their help and support throughout this project.

References

[1] Internet prototocol, 1981.

Notes

[1]Endnote

Registrant organization	Registrant email	# of homograph IDNs
Domains By Proxy, LLC	–	89
Super Privacy Service c/o Dynadot	privacy@dynadot.com	23
Domain Registries Foundation	–	22
Duong Thien	thiendv@outlook.com	18
Syngenuity Limited	manager@syngenuity.com	12
Helpnet: Brand Development & Sales	help@strongestbrands.com	12
ONUNO L.L.C.	corucas@gmail.com	11
Privacy Protection Service INC d/b/a	contact@privacyprotect.org	10
Hubertus Henz	hu_h5@yahoo.de	9
wuyu	wy65535@126.com	7

Table 4: Top ten registrants with the most homograph IDNs.