Security and Privacy Implications of URL Shortening Services

Alexander Neumann, Johannes Barnickel, Ulrike Meyer

IT Security Research Group
RWTH Aachen University
Germany

May 26th, 2011

Introduction

- Alexander Neumann
  - Computer science student, RWTH Aachen University, Germany
  - Working as Penetration Tester for RedTeam Pentesting GmbH
- Johannes Barnickel
  - Research Advisor
  - Ph.D. Student, IT Security Group, RWTH Aachen University

Definitions

- **Short URL** does not exceed 35 characters
- **Shortened URL** belongs to a shortening service
- Corner cases:
  - Shortened URL but not short:
    `http://urlshorteningservicefortwitter.com/4czyx`
  - Short URL but not shortened:
    `http://www.google.de/search?q=IEEE`

The Problem

- Long URLs tend to wrap (especially in mail)
- Twitter is restricted to 140 characters
- URLs in books should be short
- User tracking and statistics

- "Solution": URL shortening services (USS)

Technical Description

1. User posts long URL to service
   `http://itsec.rwth-aachen.de/research`
2. Service generates short URL
   `http://nvg8.it/8b896c`
3. Other user requests short URL
4. Server responds HTTP 301 or 302
5. Other user is redirected to long URL

Downsides

- Link destination is not transparent
- Reachability of the service is not guaranteed
- USS can silently stop working
  $\Rightarrow$ All shortened URLs defunct
- Request is delayed when connection to USS has high latency
- Service might be hacked and redirect to malware
- Service might be hostile and serve malware to vulnerable browsers
- Secret URLs submitted to USS cannot be deleted
- ...

Facts many users do not know

- URL service accumulates click statistics and creates profiles of users
- Short links might vanish after a few years
- When the USS dies, all shortened URLs are dysfunctional
- When the USS is hacked, links might point to other websites (malware, porn, violence, . . .)
- Short URLs can be enumerated
  $\Rightarrow$ All submitted URLs are public(!)

Research Objectives

- Analyze risks for clients, servers, and privacy of users
- Empirical studies:
  1. Determine popular USS on Twitter
  2. Analyze the use of USS in spam
  3. Test for malicious services
  4. Analyze user tracking abilities
  5. Enumerate shortened URLs
  6. Submit honeypot URLs to USS
  7. Test latency and availability of popular USS

Preliminaries: List of USS

Sources:

- Firefox add-on *ShortenURL*
- List of USS on `longurl.org` (URL expansion service)
- Lists of USS on several blogs
- Hostnames of URLs in Spam E-Mails

Leads to list of 610 USS (distinct host names), includes 527 general purpose USS.

0lv.ru
0rz.tw
1link.in
1url.com
23o.net
2big.at
2.gp
2.ly
2su.de
2tu.us
2ya.com
2ze.us
301.to
301url.com
307.to
3.ly
4ms.me
4sq.com
4url.cc
5.gp
6url.com
7.ly
9mp.com
a2n.eu
a.307.to
aa.cx
abbrr.com
abcurl.net
adf.ly
adjix.com
ad.vu
afx.cc
a.gd
a.gg
aisr.us
all.fuseurl.com
alturl.com
amzn.to
a.nf
clickmeter.com
clickthru.ca
cli.gs
clk.my
cl.lk
cl.ly
clop.in
coge.la
c-o.in
conta.cc
cort.as
cot.ag
crks.me
crum.pl
ctvr.us
curio.us
cut.im
cutt.us
cuturls.com
daa.li
dai.ly
deadsmall.com
decenturl.com
df9.net
dfl8.me
digbig.com
digg.com
digipills.com
disq.us
dld.bz
dlvr.it
dn.vc
doiop.com
do.my
dopen.us
durl.me
durl.us
dwarfurl.com
dy.fi
fwib.net
get.sfu.ca
get-
shorty.com
get-
url.com
gizmo.do
gkurl.us
gl.am
go2cut.com
go2.me
go.9nl.com
good.ly
goo.gl
go.qb.by
goshrink.com
gourl.gr
go.usa.gov
gowat.ch
g.ro.lt
gurl.es
gzurl.com
hao.jp
hex.io
hhvx.com
hiderefer.com
hmm.ph
ho.io
hop.im
hosturl.com
hotredirect.com
href.in
hsblinks.com
htxt.it
hub.tm
huff.to
hulu.com
hurl.it
hurl.me
kurl.nu
kurzurl.net
k.vu
l9k.net
lanjut.in
lat.ms
liip.to
liltext.com
lin.cr
lin.io
linkbee.com
linkbun.ch
linkee.com
linkl.ru
linkslice.com
link.toolbot.com
linxfix.com
liteurl.net
liurl.cn
lnk.by
lnkd.in
lnk.gd
lnk.in
lnk.ly
lnk.ms
lnk.nu
lnk.sk
lnkurl.com
ln-s.net
ln-s.ru
lookleap.com
low.cc
l.pr
lru.jp
lt.tl
lurl.no
macte.ch
makeashorterlink.com
makeitbrief.com
nn.nf
notifyurl.com
notlong.com
not.my
n.pr
nsfw.in
nutshellurl.com
nvg8.it
nxy.in
nyti.ms
oboeyasui.com
oc1.us
odun.net
omf.gd
om.ly
omoikane.net
on.cnn.com
on.mktw.net
ooqx.com
orz.se
ow.ly
o-x.fr
parv.us
paulding.net
pd.am
pendek.in
pic.gd
piko.me
ping.fm
pipes.yahoo.com
piurl.com
pli.gs
plumurl.com
plurl.me
p.ly
pnt.me
politi.co
pk.pe.kr
post.ly
retwt.me
rickroll.it
r.im
ri.ms
riz.gd
rmse.ru
rnk.me
rnm.me
rt.nu
rubyurl.com
ru.ly
rurl.org
rww.tw
s4c.in
s7y.us
safe.mn
sai.ly
sdut.us
services.digg.com
sfu.ca
shar.es
sharetabs.com
shim.net
shink.de
shmyl.com
shorl.com
shortenurl.com
shorterlink.com
shorterlink.co.uk
short.ie
shortio.com
shortlinks.co.uk
shortn.me
short.to
shorturl.com
shorturl.de
shoturl.us
shout.to
show.my
snipurl.com
snkr.me
snurl.com
sokrati.ru
song.ly
sp2.ro
spedr.com
sqrl.it
srnk.net
srs.li
starturl.com
sturly.com
su.pr
surl.co.uk
surl.hu
ta.gd
tbd.ly
t.co
tgr.me
tgr.ph
thesurl.com
thinfi.com
thnlnk.com
tighturl.com
timesurl.at
tiniuri.com
tini.us
tinurl.mobi
tiny123.com
tinyarro.ws
tiny.by
tiny.cc
tinylink.com
tinylink.in
tiny.ly
tiny.pl
tinypl.us
tinyuri.ca
twhub.com
twip.us
twirl.at
twitclicks.com
twitterpan.com
twitterurl.net
twitterurl.org
twittu.ms
twiturl.de
twtr.us
twurl.cc
twurl.nl
u28.de
u6e.de
u76.org
ub0.cc
uforgot.me
uik.in
uiop.me
ulimit.com
ulu.lu
u.mavrev.com
unfake.it
u.nu
updating.me
ur1.ca
urizy.com
url360.me
url4.eu
url9.com
urlac.com
url.ag
urlao.com
url.az
urlbee.com
urlbit.us
urlborg.com
urlbrief.com
urlcorta.es
urlz.at
urlzen.com
usat.ly
use.my
u.to
vb.ly
vdirect.com
v.gd
vgn.am
vi.ly
vl.am
voizle.com
voomr.com
vtc.es
w3t.org
w55.de
wapo.st
wapurl.co.uk
webalias.com
weturl.com
w.hurl.ws
wipi.es
wp.me
xaddr.com
xeeurl.com
xil.in
xlurl.net
xr.com
xrl.in
xrl.us
x.se
xsm.us
xs.to
xurl.es
xurl.jp
x.vu
xxsurl.de
y.ahoo.it
yatuc.com

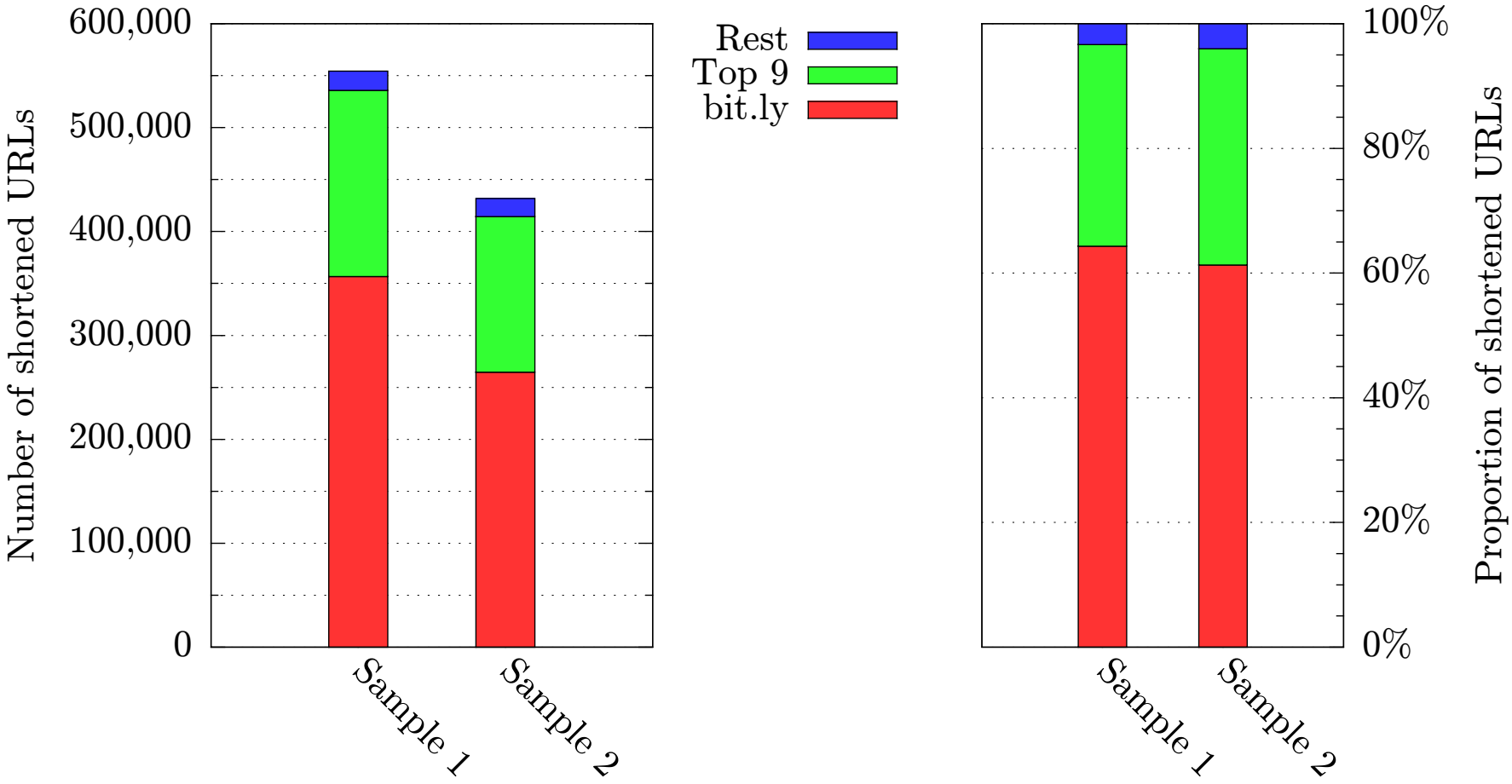
Results

Determine Popular USS on Twitter

- Base: Two samples of Twitter messages (24 hours each, max 10 %)
- Results:
  - 7.5 million / 8.7 million messages
  - 1.2 million / 1.1 million URLs
  - 553,320 / 431,636 shortened URLs
- Top ten services (cover 96% of all USS on Twitter):

1. bit.ly / j.mp
2. t.co
3. tinyurl.com
4. goo.gl
5. ow.ly
6. dlvr.it
7. is.gd
8. migre.me
9. dld.bz
10. lnk.ms

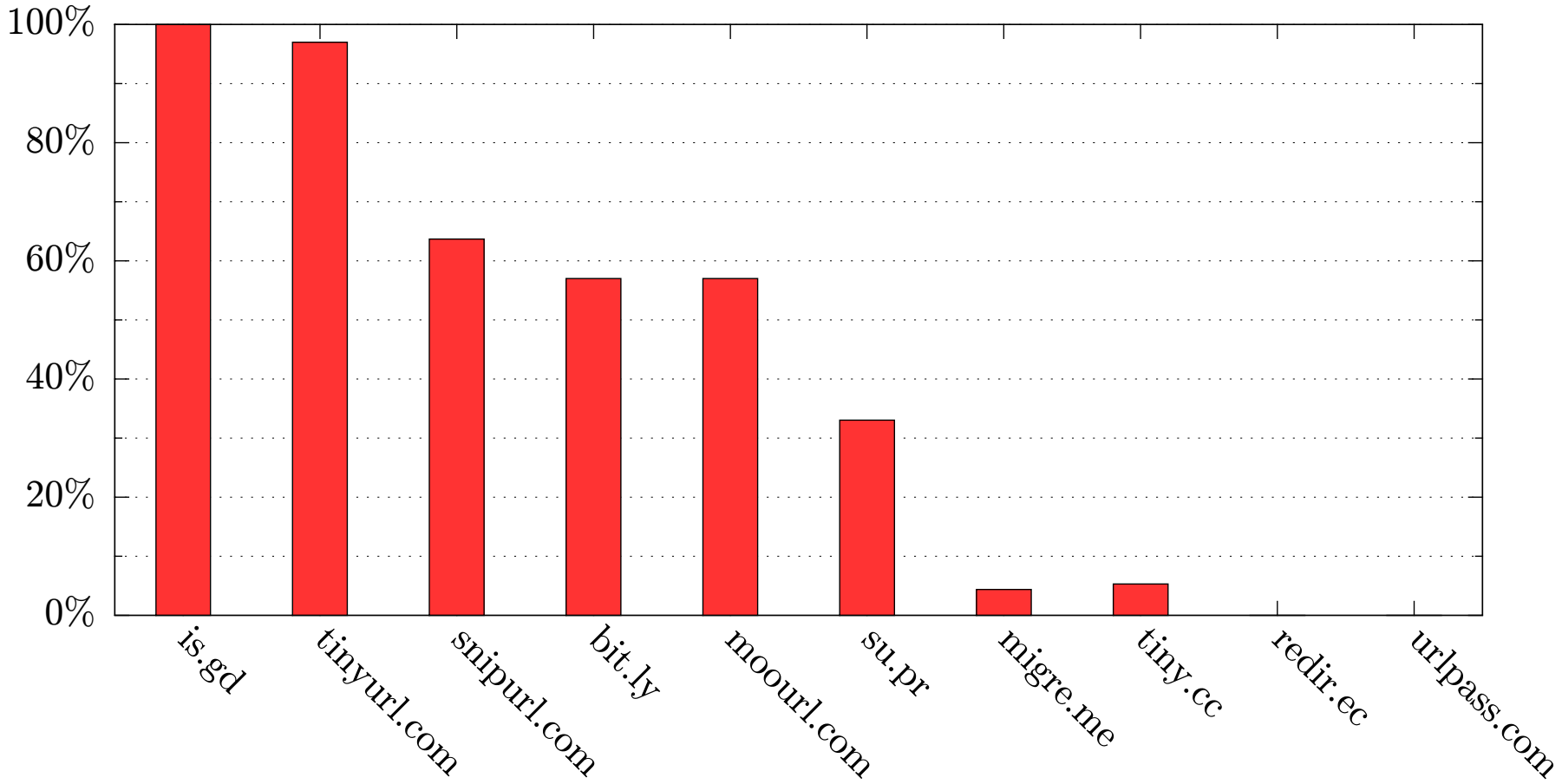
Top ten services



USS in Spam

- Spam e-mails from SCHNUCKI project
- 7.9 million e-mails collected since 2003
- 12.8 million URLs, 0.3 % shortened URLs
- Query shortened URLs and analyze response code
- Calculate spam detection rate for relevant services

Spam detection rate of USS



Malicious USS

Attack scenario: Setup a fast and attractive USS, serve all requests normally, but after a while start sending vulnerable browsers to malware sites.

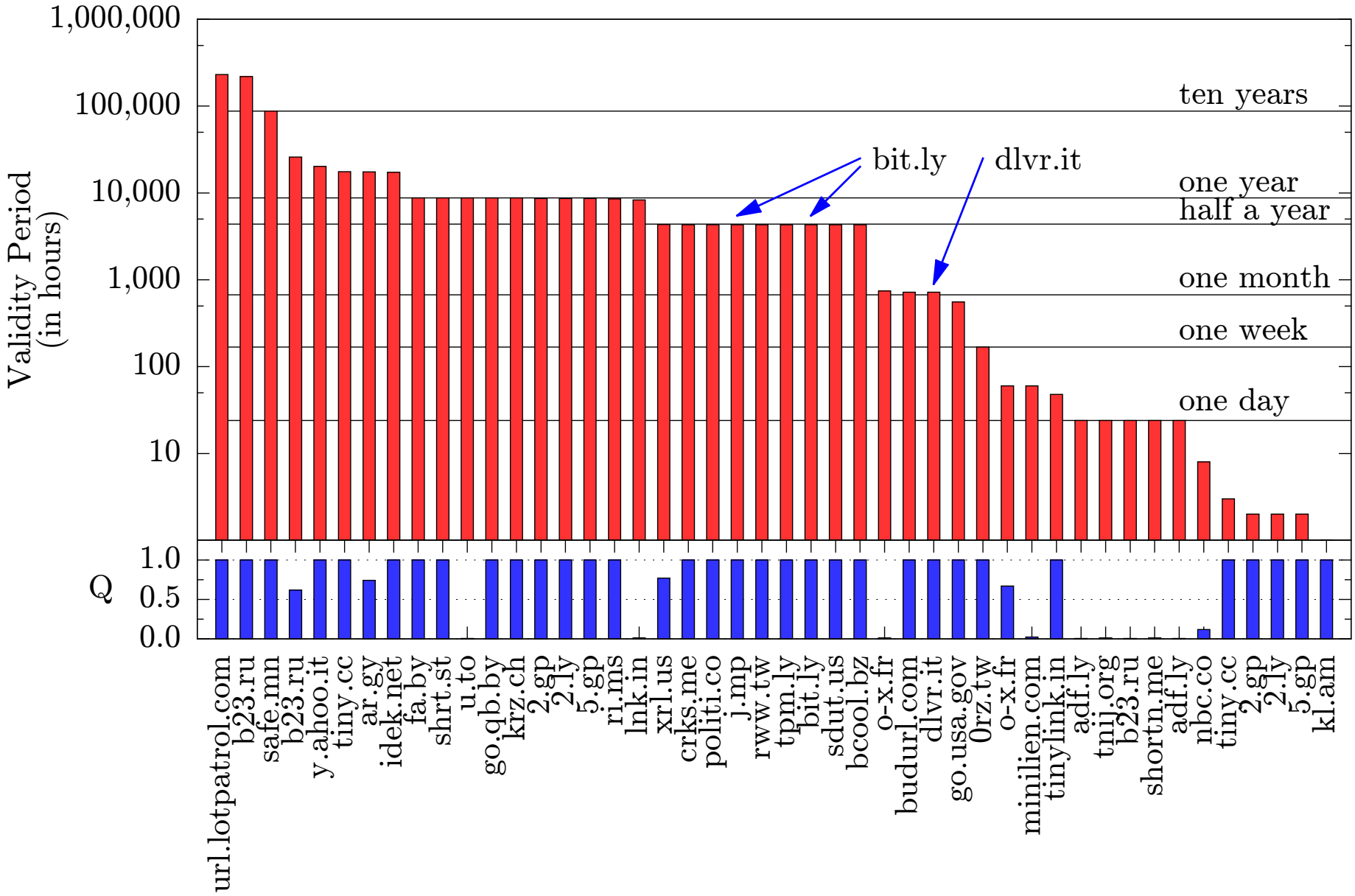
Analysis:

- Query shortened URLs seen on Twitter for 187 different USS with 83 different User-Agent strings
- Analyze HTTP response Location header
- Result: No malicious behaviour found
- But: One service handles browsers different

User-Tracking USS

- Data from previous experiment
- Analyze HTTP response Set-Cookie header
- Results:
  - 65 USS set cookies
  - 38 USS set persistent cookies
  - 28 USS set persistent cookies with validity period 6 months or more
- Define Q as quotient: $\frac{\#\text{all unique values}}{\#\text{all values}}$ received for the cookie

Validity period of cookies



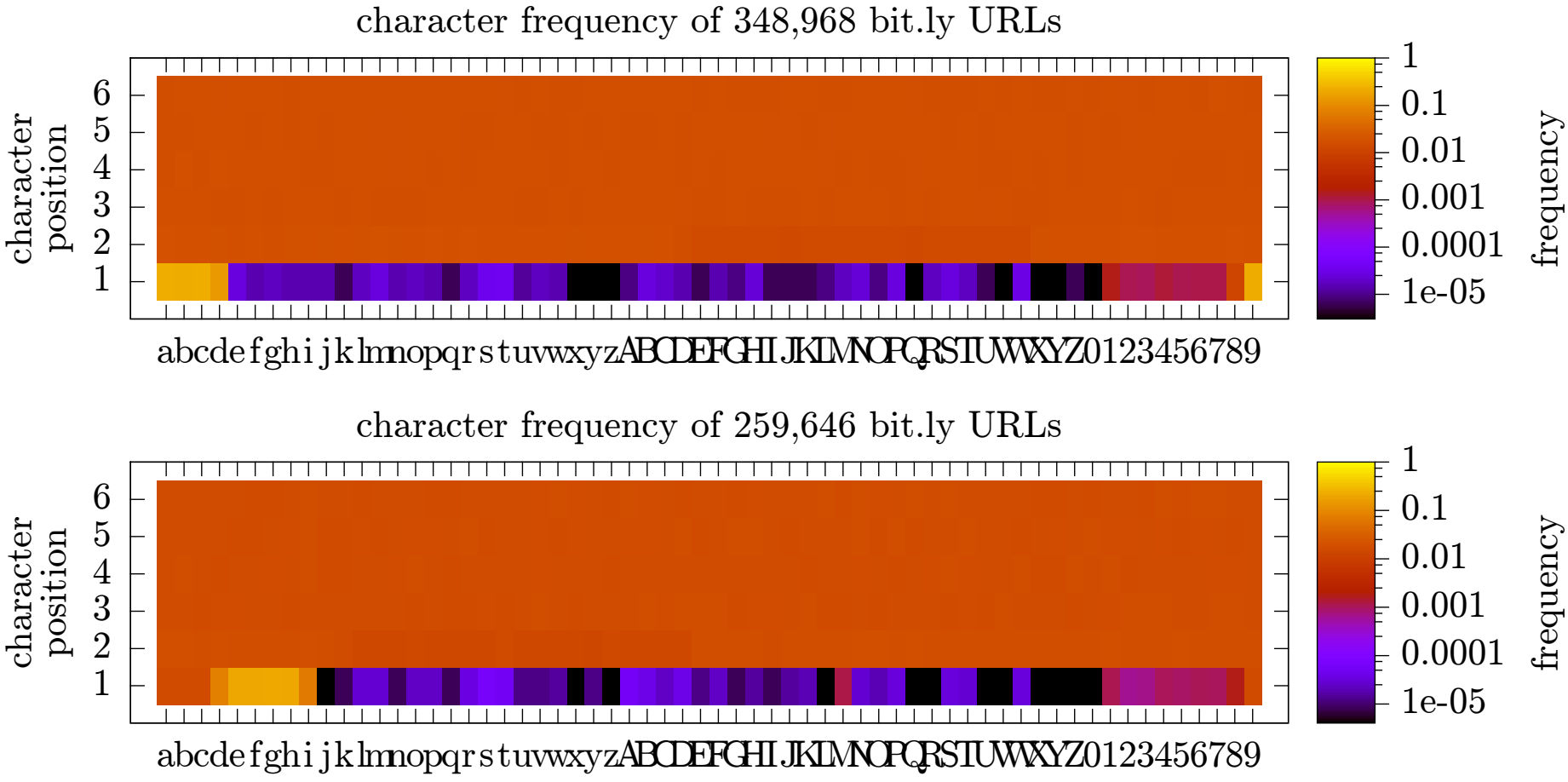
Enumerating shortened URLs

For the top ten USS:

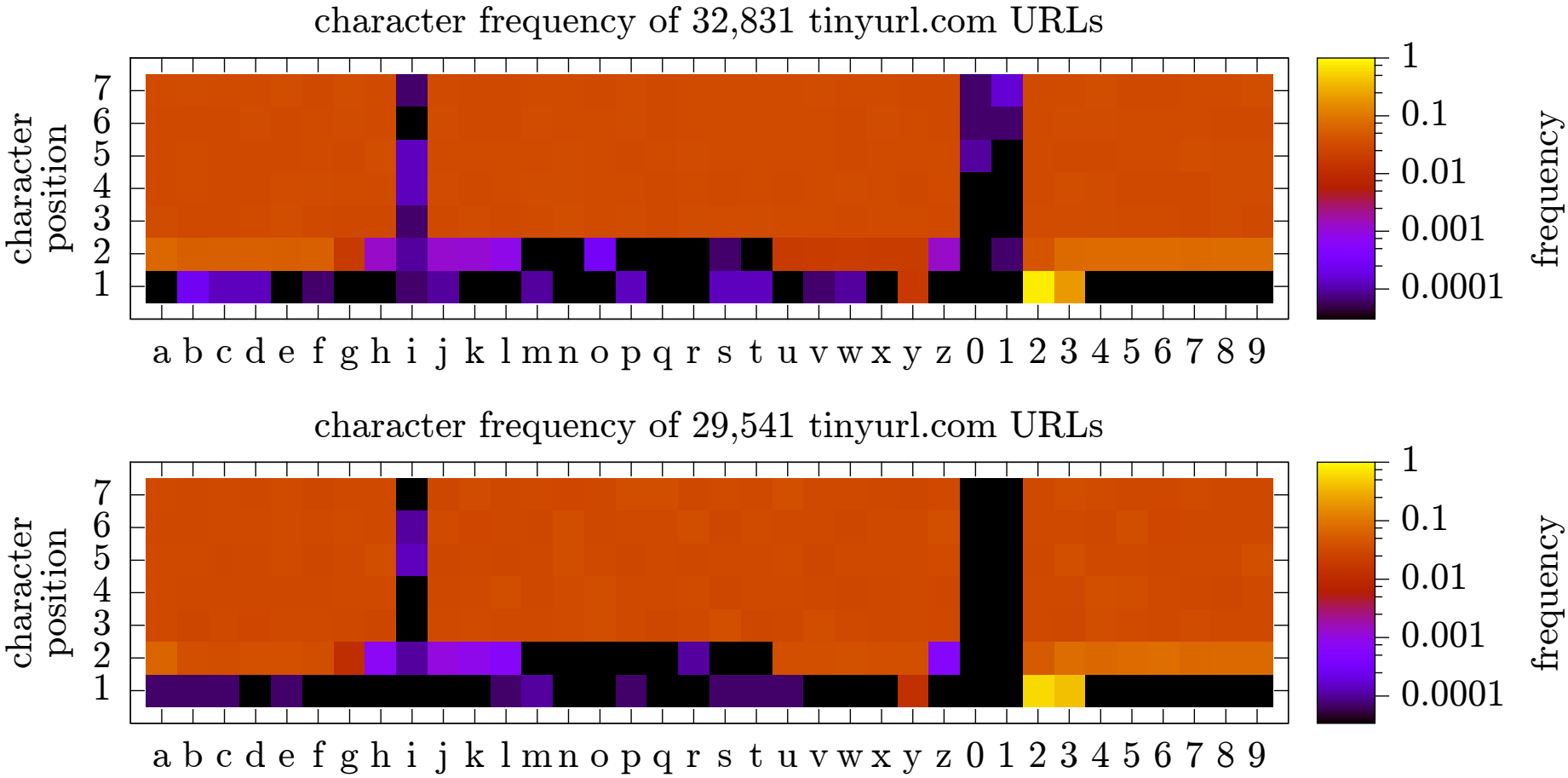
- Analyze structure of shortened URLs in both Twitter samples
  $\Rightarrow$ character frequency analysis using heatmaps
- Select range, ca. 230k URLs per USS
- Enumerate all URLs in range
- Inspect results by hand, search for secret URLs

- Observation: Only `goo.gl` imposed restrictions

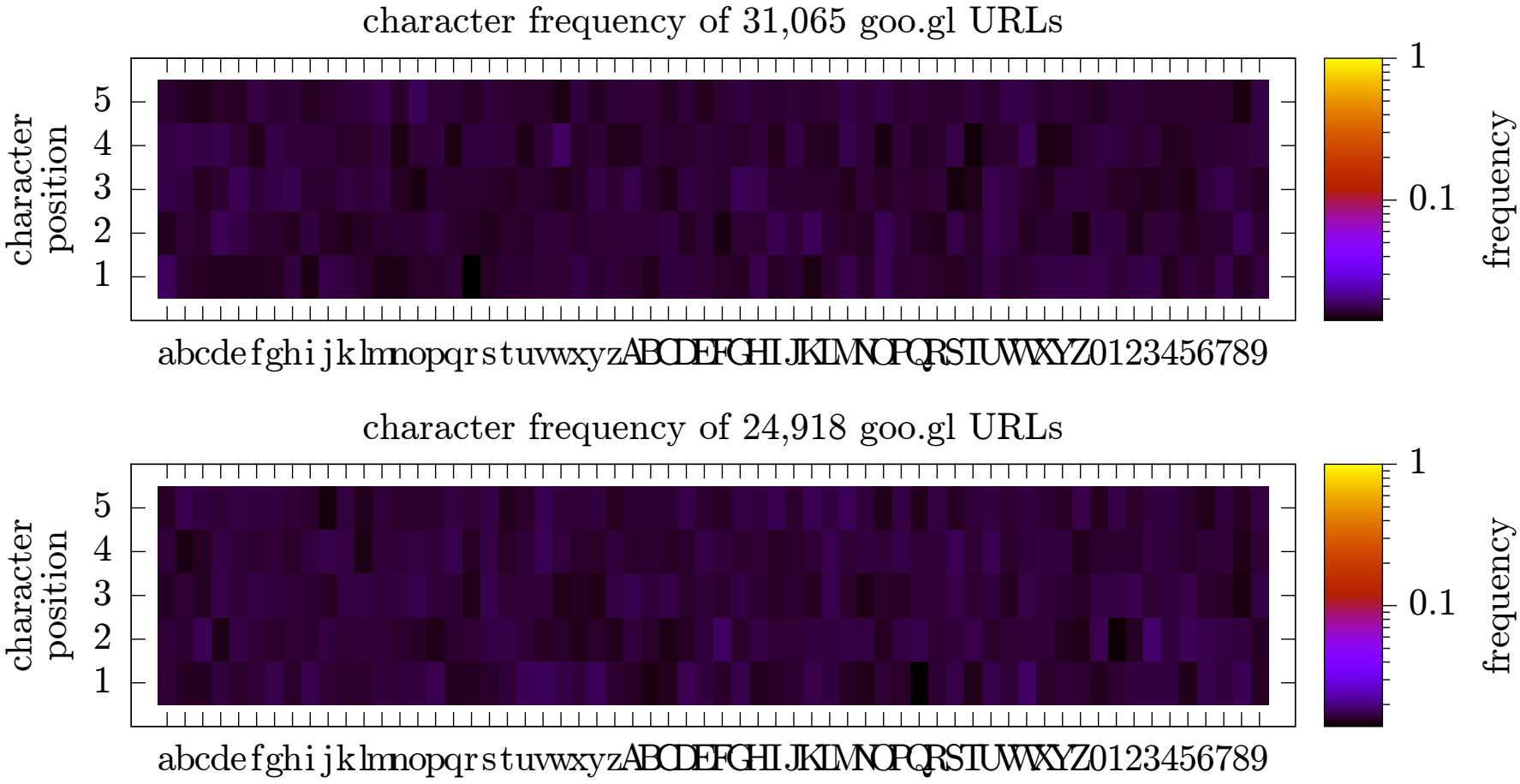
Character frequency analysis: bit.ly



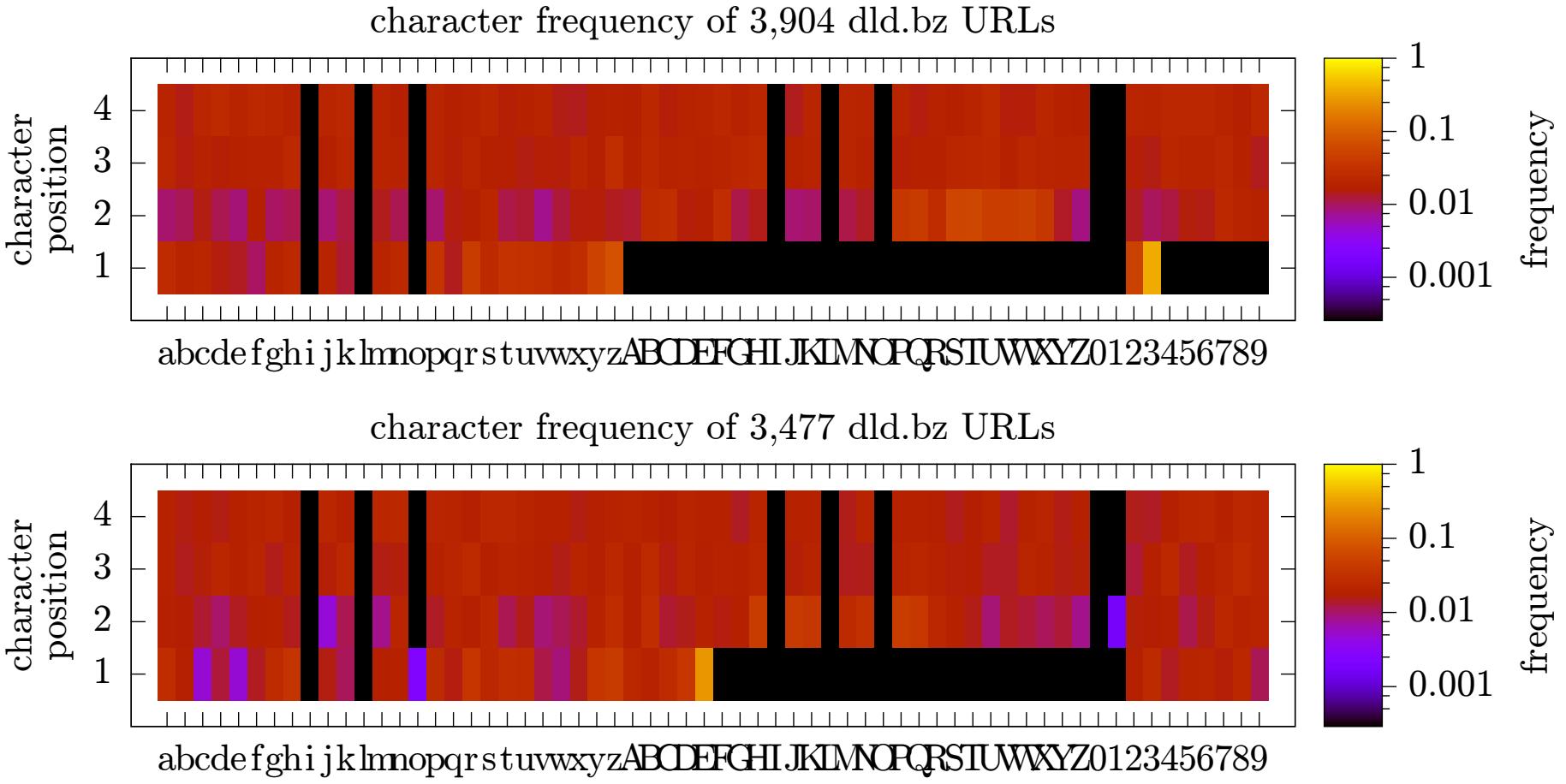
Character frequency analysis: tinyurl.com



Character frequency analysis: goo.gl



Character frequency analysis: dld.bz



Secret URLs found by enumerating

- Archives of private photos
- Several CVs
- Treasurer's report for a company
- List of names and numbers of a kindergarten in Lindlar
- ...

Submitting secret URLs to USS

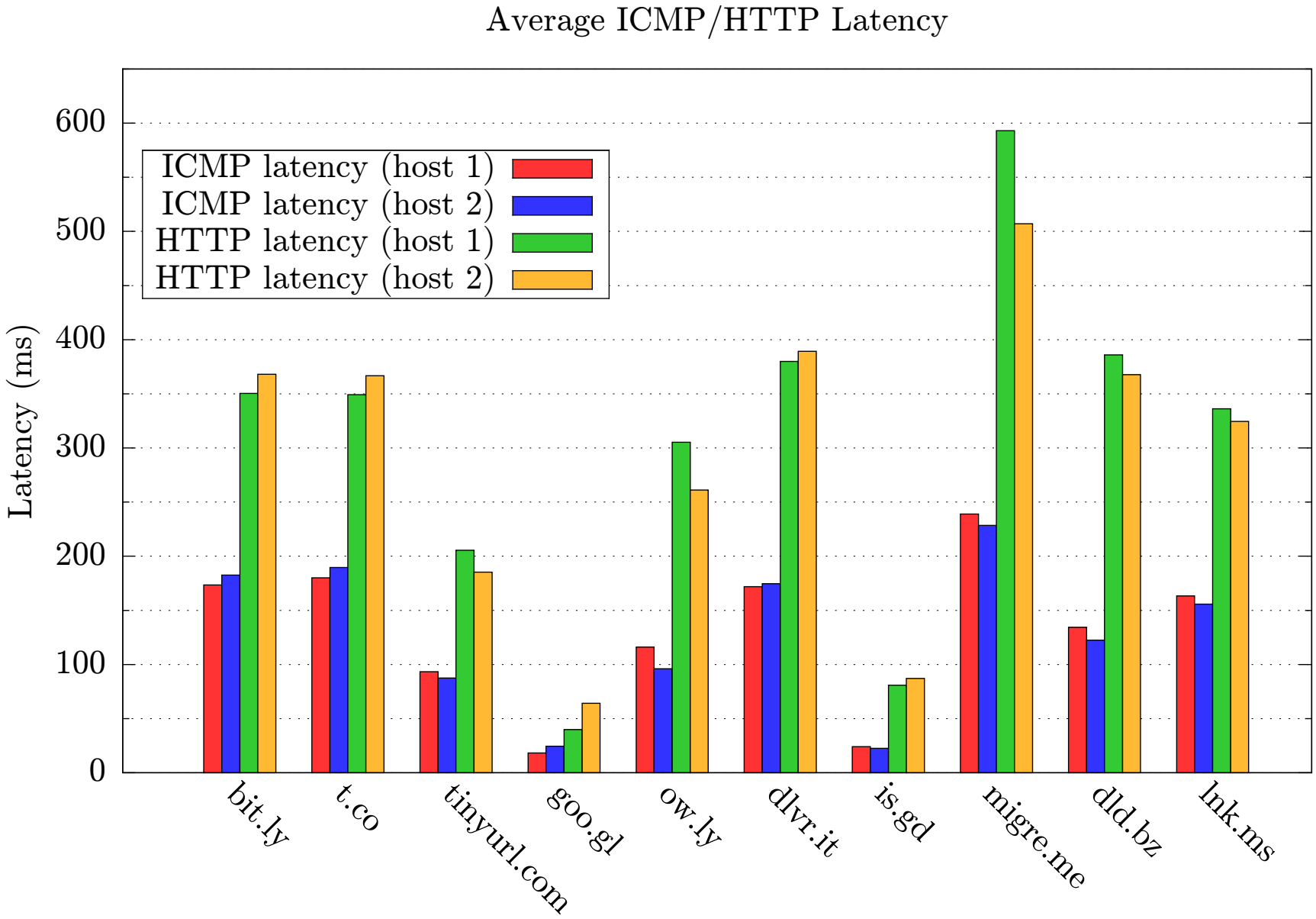
- Question: Are secret URLs submitted to USS leaked?
- Set up honeypot web server
- Generate unique URLs for each service
- Suspicious and harmless URLs
- Examples:
  - `http://fd0.me/secret/a0df29ac/bb42ce8b`
  - `http://www.fd0.me/blog/archive/2011/01/14/index.php?article=69e325eb#a5a6c61c`
- Submit to 255 USS
- Watch for requests

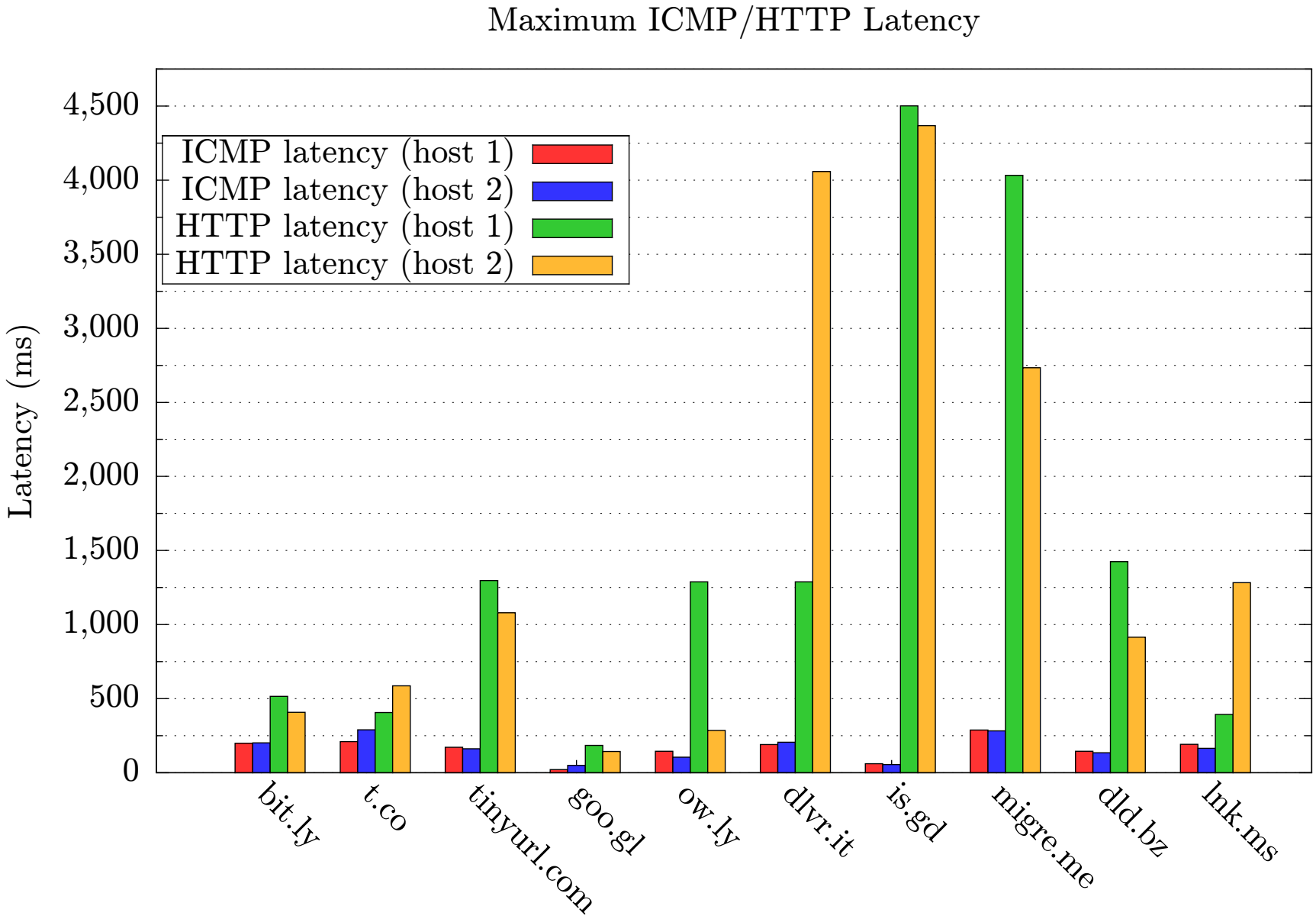
Results (after four weeks)

- Honeypot is found by Google, Yahoo and Baidu
- 15 URLs requested by Google
- 13 URLs requested by Yahoo
- 2 URLs requested by Baidu
- 13 URLs manually checked by USS administrators
  (9 transmitted the admin URL in the HTTP referrer)
- Four administrators contacted us, are interested in the research
- $\Rightarrow$ Never submit private URLs to shortening services!

Latency and availability measurements

- Latency and availability measured with Smokeping
- From two different servers (in Germany)
- ICMP and HTTP latency measured for ten services
- Results:
  - Most services have good average HTTP latency
  - Some services have a very bad worst-case HTTP latency
  - `goo.gl` USS wins


Average ICMP/HTTP Latency
ICMP latency (host 1)
ICMP latency (host 2)
HTTP latency (host 1)
HTTP latency (host 2)
Latency (ms)
0
100
200
300
400
500
600
bit.ly
t.co
tinyurl.com
goo.gl
ow.ly
dlvr.it
is.gd
migre.me
dld.bz
lnk.ms


Maximum ICMP/HTTP Latency
Latency (ms)
4,500
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
ICMP latency (host 1)
ICMP latency (host 2)
HTTP latency (host 1)
HTTP latency (host 2)
bit.ly
t.co
tinyurl.com
goo.gl
ow.ly
dlvr.it
is.gd
migre.me
dld.bz
lnk.ms

Conclusion

- USS have risks
- These risks are very real
- USS leak URLs to search engines
- Do not submit private URLs to USS
- USS are used in spam e-mails
- Several services set long-running cookies and can track the user
- Shortened URLs are not completely random
- `goo.gl` USS dominates all others in every discipline

Questions?

Thank you for your attention.