What Can Figure Learn ack 67861, win 5963, length 1 from an IP? Simran Patil and Nikita Borisov^{818, win 65535, options [mss 1460, nop} University of Illinois at Urbana-Champaign



In the beginning...



GET /~nikitab/ HTTP/1.1
Host: geocities.com

HTTP/1.1 200 OK

...

...

<blink>this page is under construction</blink>





ANRW'19





Methodology



rDNS



Public Suffix List (PSL) match: server1.facebook.com =~ facebook.com

Domains and IPs



IP Anonymity Set



IP Anonymity Sets



S. Patil & N. Borisov, "What Can You Learn from an IP?"

Site-unique IPs



ANRW'19

Site-unique IPs



Page Load Fingerprints

👌 queen-elizabeth-montreal	www.fairmont.com	document	23.64.109.196:443	
js fairmont-prod.js	cloud.moovweb.net	js	192.33.31.70:443	
ss font.css	www.fairmont.com	css	23.64.109.196:443	
ss fairmont.min.css	🔒 d1r80pgesjju2u.cloud	CSS	99.84.112.4:443	
ss tripadvisor.min.css	🔒 d1r80pgesjju2u.cloud	css	99.84.112.4:443	
35 3colbigbanner.min.css	www.fairmont.com	CSS	23.64.109.196:443	site???
js s_code.js	🔒 d1r80pgesjju2u.cloud	js	99.84.112.4:443	
js accorhotelsconnect.js	secure.accorhotels.ws	js	193.200.231.133:443	
js WebResource.axd	www.fairmont.com	js	23.64.109.196:443	
📄 593e534a-7cf8-4956-b5a9-4c8da07c	www.fairmont.com	jpg	23.64.109.196:443	7
	1			

23.64.109.196 192.33.31.70 98.84.112.4 193.200.231.133

Site IP sets



What about CDNs?

- Many CDNs could use same IP address for all sites but don't
 - Ported IP space
 - Connections w/o SNI
- In our data set 200K domains are hosted by CloudFlare, using 91K IPs
 - Including 3% of the sites with a site-unique front page IP
- Randomizing or normalizing IP addresses could help

Conclusions

- DNS privacy offers limited protection
 - For web browsing
 - Against an adversary with a good prior list of sites
- In our Alexa 1M crawl dataset
 - 48% of all IPs map to a single domain
 - 68% of all IPs map to a single site
 - 43% of all sites contain a site-unique IP
 - 95% of sites have a unique IP set
- Changes to web hosting infrastructure could help
 - Normalize or randomize CDN IP addresses