Investigating Advertisers' Domain-changing Behaviors and Their Impacts on Ad-blocker Filter Lists

Su-Chin Lin*, Kai-Hsiang Chou*, Yen Chen, Hsu-Chun Hsiao,
Darion Cassel, Lujo Bauer, Limin Jia

The Web Conference (WWW) 2022

國立臺灣大學
National Taiwan University

Carnegie Mellon University

NSLAB
National Taiwan University
Network Security Lab

CyLab Carnegie Mellon University
Security and Privacy Institute

* Both authors contributed equally to this research.

How Ad Blocker Works

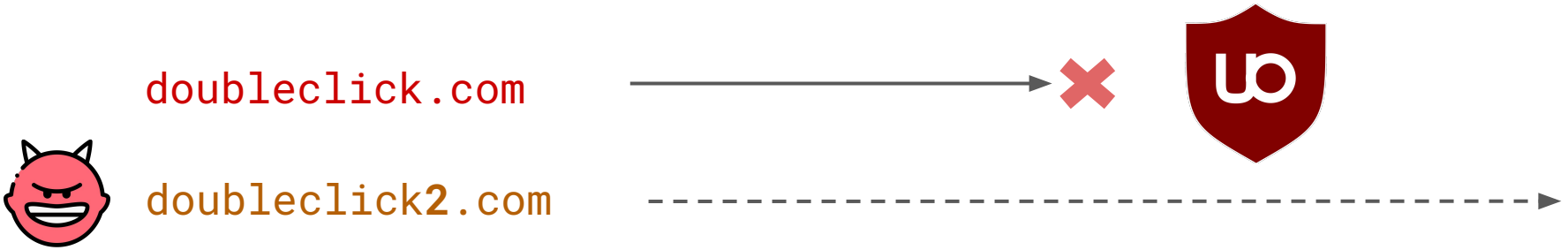
- Ad blockers rely on static filer lists to block ad domains

How Advertisers Circumvent Ad Blockers

- Anti ad-blocker
- Bad user experience

How Advertisers Circumvent Ad Blockers

- Previous research [11,15,36,56] observed that advertisers change domains to circumvent ad blockers
- But no systematic investigation on the impact



We define doubleclick**2**.com as a *replica ad domain (RAD domain)*

A Real-world Case: *Clickadu's* RAD Domains

RAD domain	Appearing date	Blocked date
sghehllds.com	2019-06-21	2019-07-25
drjgjngf.com	2019-06-22	2019-07-15
qumagee.com	2019-07-16	2019-10-04
xineday.com	2019-08-18	2019-09-05
mrzikj.com	2019-08-26	2019-10-09
tibacta.com	2019-09-01	2019-09-07

Motivation

- No prior study has systematically investigated the RAD domain influence
- We would like to know the prevalent and privacy impact of RAD domains
- Analyzing RAD domains can potentially help filter-list maintainers to react to domain-chaning behaviors

Research Questions and HighLights

1. What are the common patterns of RAD domains?
    - We propose methods for discovering RAD domains
    - We present a taxonomy of common doman-changing patterns

Research Questions and HighLights

2. How prevalent are RAD domains?
    - RAD domains appear on 10.24% of the 50K websites we crawled

Research Questions and HighLights

3. What is the privacy impact of RAD domains?
    - RAD domains can additionally extend the timespan of the ad domain for 558 days

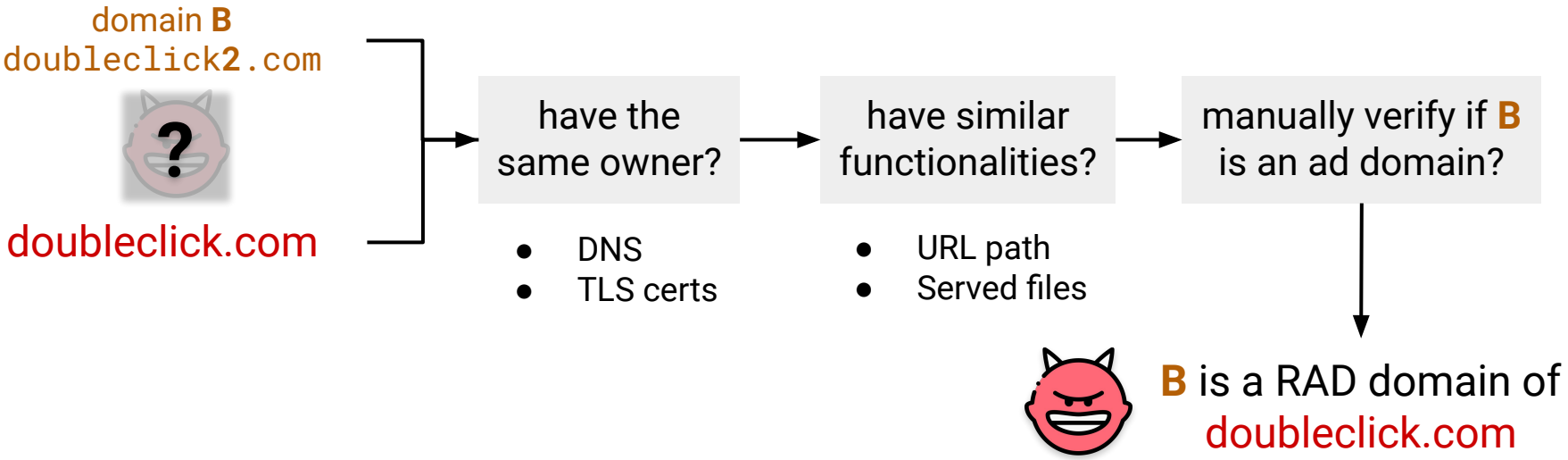
Methodology

- How do we identify RAD domains?
- Is a domain **B** doubleclick2.com an RAD domain of doubleclick.com?

domain **B**
doubleclick**2**.com

doubleclick.com

Methodology



Results - How Prevalent Are RAD Domains?

- We identified 1,748 RAD domains, and 652 of them were not blocked as of Feb. 2021
- We discoverd 10.24% of the 50K websites we crawed sent at least one request to RAD domains

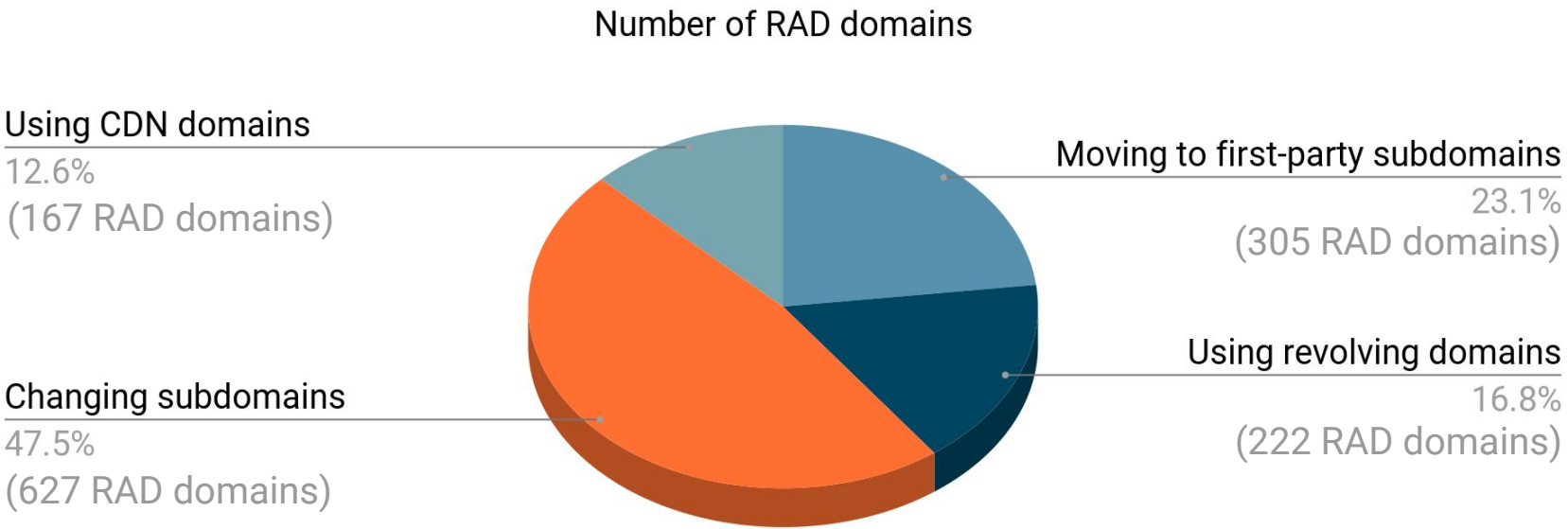
RAD domains are prevalent and can substantially harm user’s privacy

Results - What Is the Privacy Impact?

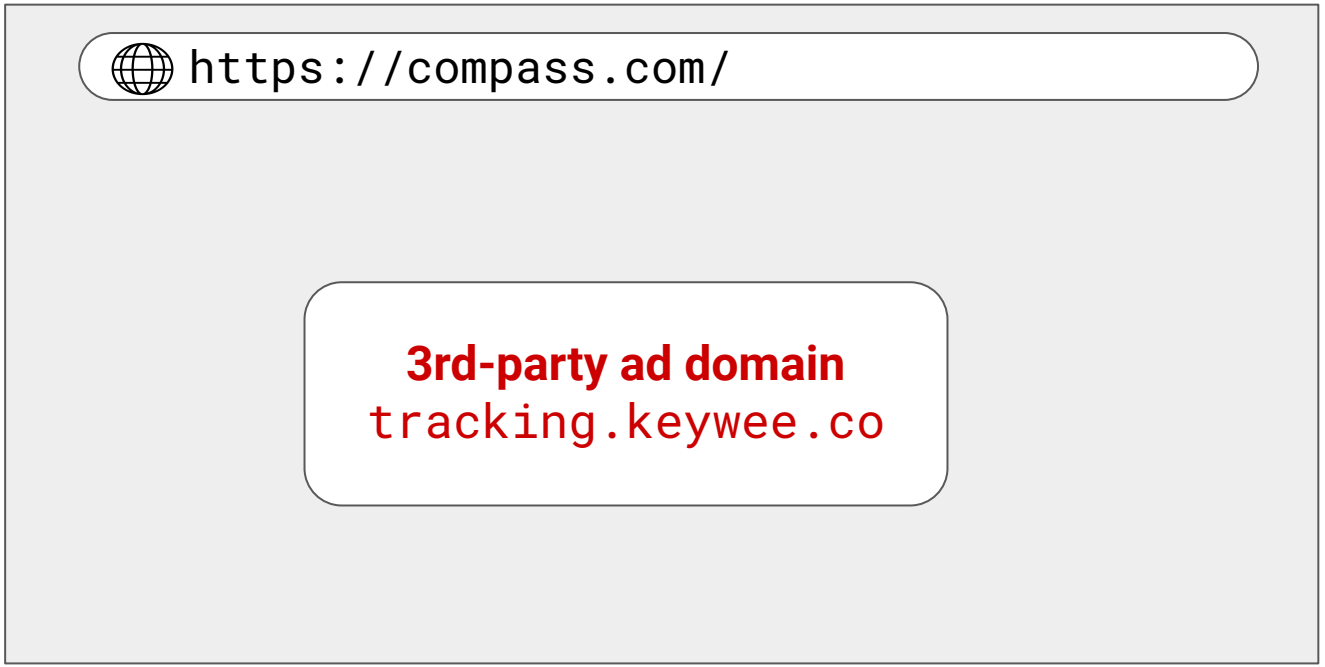
- 415 (23.7%) of RAD domains exhibit privacy-intrusive behaviors by DuckDuckGo's Tracker Radar
- RAD domains can additionally extend the timespan of the ad domain for 558 days

RAD domains are an effective approach to circumvent ad blockers

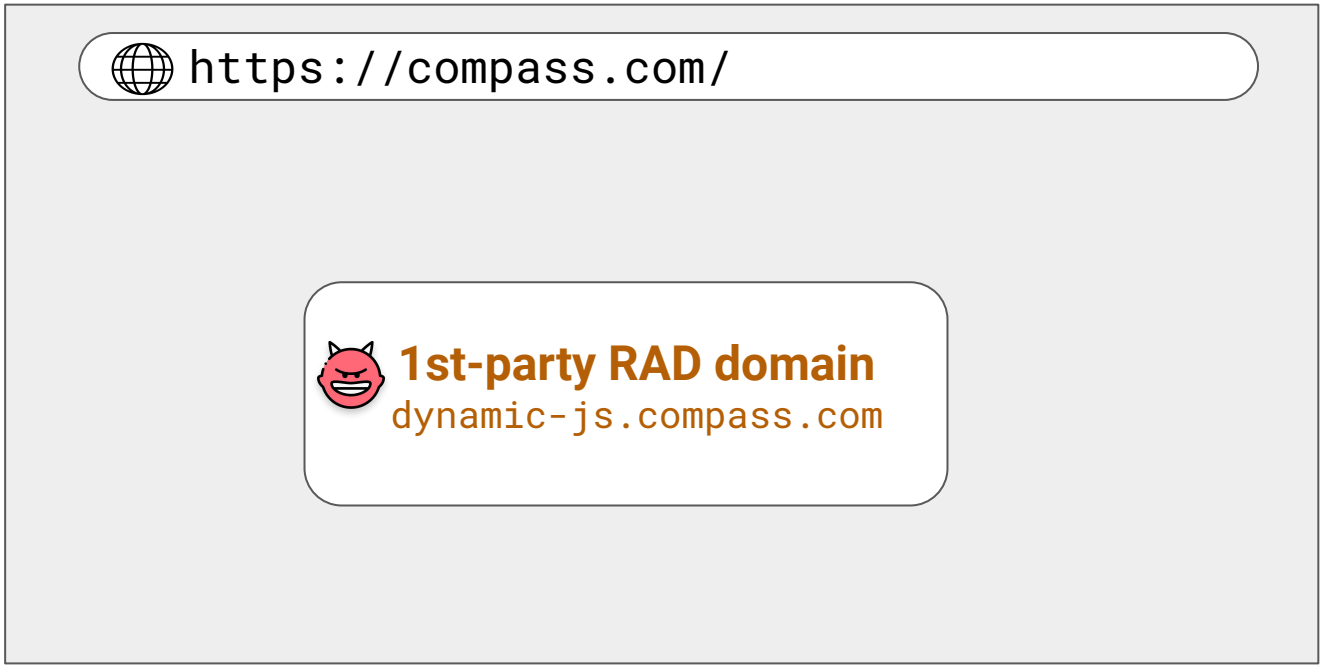
Results - What Are the Common Patterns?



Results - Moving to First-party Subdomains



Results - Moving to First-party Subdomains



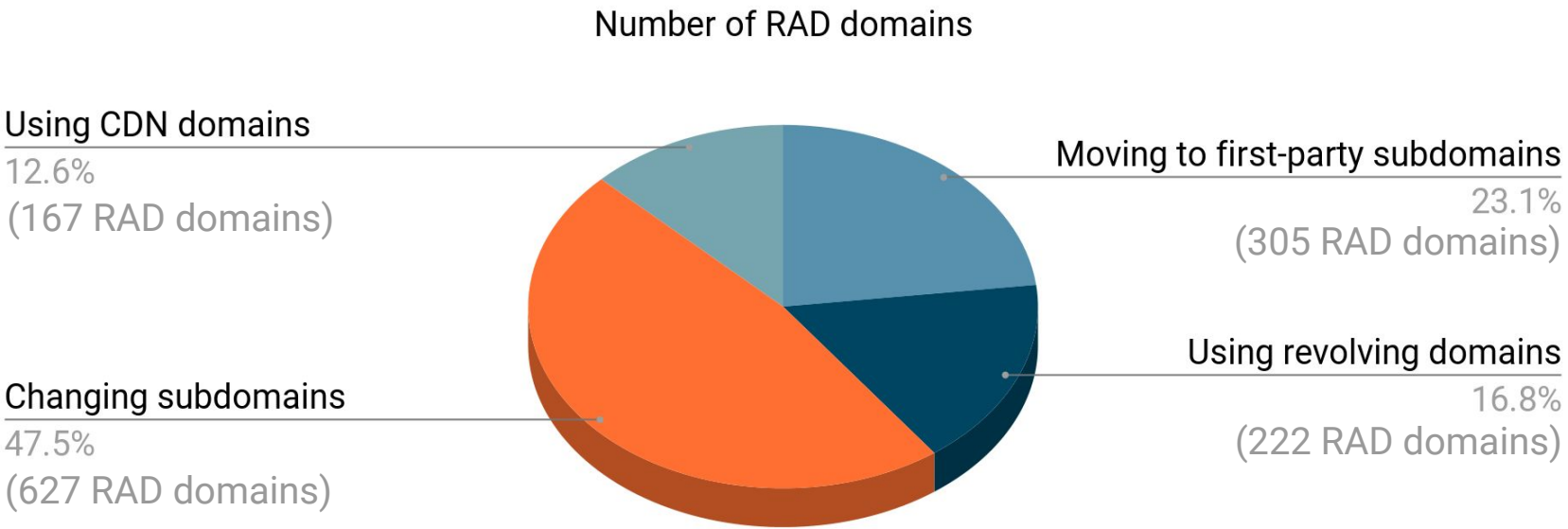
Results - Moving to First-party Subdomains

- Filter list policy is lenient regarding first-party ad domains
- It blurs the trust boundary between 1st and 3rd-party websites
- Advertisers encourages publishers to delegate ads using first-party proxy subdomains to bypass ad blockers

Results - Moving to First-party Subdomains

- Recommendation
  - **Content publisher**: should be transparent about the cooperation with advertisers
  - **Filter-list maintainers**: should consider the actual information flow on first-party domains

Results - What Are the Common Patterns?



Discussion - Possible Reasons For Using RAD Domains

- Customer isolation
- Localization
- Infrastructure changes
- Ad-blocker evasion

Future Work

- Report RAD domains to filter list community
- Improve our methodology to identify RAD domains
- Facilitate filter-list maintainer to react to RAD domains

Conclusion

- We propose methods for discovering RAD domains
- We present a taxonomy of RAD domain patterns
- RAD domains appear on 10.24% of the websites we crawled
- RAD domains extend the timespan of the ad domain for 558 days


Su-Chin Lin


r07922067@ntu.edu.tw


https://github.com/csienslab/RAD-domain-analysis/

Icon Credit

[1] Ad icons created by Freepik - Flaticon

[2] Javascript icons created by Flat Icons - Flaticon

[3] Image icons created by Good Ware - Flaticon

[4] Devil icons created by Freepik - Flaticon

[5] Mail icons created by Freepik - Flaticon

[6] User icons created by Freepik - Flaticon