TrackerSift: Untangling Mixed Tracking and Functional Web Resources
Abdul Haddi Amjad¹, Danial Saleem², Muhammad Ali Gulzar¹, Zubair Shafiq³, Fareed Zaffar²
¹Virginia Tech, ²LUMS, ³University of California, Davis

Introduction:
Two Key Stakeholders:
- Advertisers/Trackers
- Content Blocker

Circumvention Techniques used by Advertisers/Trackers:
1. Change the Network Location
2. Mixing-up Resources
   - Doubleclick.net
   - Unblock-ads.net

On the website nbcnews.com, we see this JavaScript:
https://nodeassets.nbcnews.com/_next/static/chunks/pages/app-3b8be4d8843e5e65.js - aka app.js
Initiates requests to taboola.com - popular ad-exchange
Initiates requests to get live news and login for the website.

Motivation:
Desired solution should block tracking request to taboola.com
Current options for Content Blockers:
1. Block network requests to taboola.com
   - Many previous related work shows that advertisers/trackers frequently change network endpoints to evade content blockers
   - Content delivery networks (cdn) are used to serve both tracking and functionality
2. Block app.js
   - This choice compromises the functionality and causes breakage in navigation.

TrackerSift:
We propose TrackerSift, an approach to untangle mixed resources at multiple levels of granularity in a hierarchical fashion:
- It starts from the coarse granularity of domains which is used in tools like disconnect.
- For mixed domains, it analyzes the hostname information in the fully qualified domain names like pi-hole.
- For mixed hostnames, it analyzes initiator script URLs like oerset
- For mixed scripts, it further analyzes script methods to classify tracking and functional resources.
- To the best of our knowledge, TrackerSift is the first approach that can analyze and detect tracking resources beyond the script-level granularity.

Implementation:

Results:

Future Work:
The 2% of the requests that can not be separated at the method level granularity, we believe that removing the point of confluence i.e., the method that participates in only stack traces of tracking requests will break the chain of events that invoke tracking behavior.