ADINT: Using Targeted Advertising for Information Gathering

Paul Vines
~2017

Outline

1. How the Advertising Ecosystem Works
    a. What they know about you
    b. How an ad gets served
2. ADINT: Using Advertising for Information Gathering
    a. Concept
    b. Case Study
    c. Survey
    d. Potential Uses
    e. Defenses and The Future

How the Advertising Ecosystem Works

Effective Ads = $$$

- More information → greater precision
- Greater precision targeting → greater ad value
- More information → greater value

Tracking Every-Thing in Every-Way

Some Things that are tracked:

➔ Pages visited
➔ Online and offline purchases
➔ Accounts made
➔ Location
➔ Emails

Some Ways of tracking:

➔ Tracking cookies
➔ Browser fingerprinting
➔ Services accounts
➔ Shopping club cards
➔ Email accounts

Information Sharing

- ➔ Entities constantly buy and sell information about you.

- ➔ Only one shared identifier is needed to merge two datasets

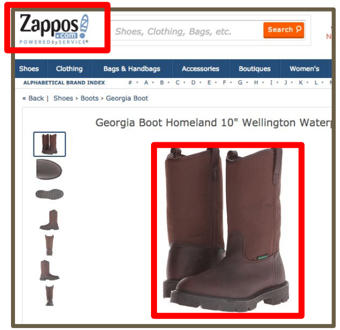
Youtube

Email	Cookie	Interests
joe@email	AF32X93	football, alabama
jane@gmail	CFG344A	finances, printing
mark@email	K9339SA	dance, cooking
sue@email	AEEF334	prisons, baking

Costco		
Email	**Name**	**Address**
joe@email	Joe Savo	123 11th St, MA
jane@gmail	Jane Carovo	897 3rd St, WA
mark@email	Mark Soso	343 9th St, VA
sue@email	Susanne Keel	222 4th St, ND

Amazon		
Cookie	**Purchases**	
AF32X93	charcoal	bud light
CFG344A	ink	yacht
K9339SA	ballet shoes	62" TV
AEEF334	binoculars	hacksaws

How a Targeted Web Ad Happens


Zappos .com
POWERED by SERVICE®
Shoes, Clothing, Bags, etc.
Search
Shoes
Clothing
Bags & Handbags
Accessories
Boutiques
Women's
ALPHABETICAL BRAND INDEX
# • A • B • C • D • E • F • G • H • I • J • K • L •
« Back | Shoes › Boots › Georgia Boot
Georgia Boot Homeland 10" Wellington Water

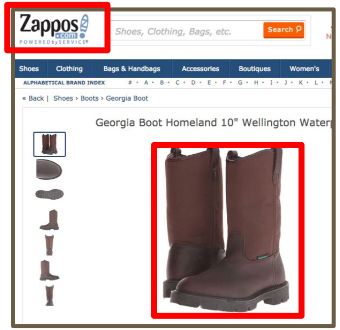
You just installed Chrome

1. **Went to zappos.com and browsed some shoes**

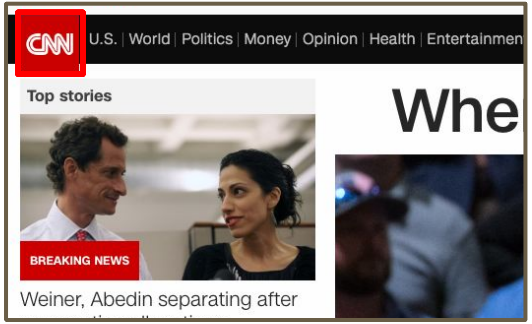
How a Targeted Web Ad Happens

You just installed Chrome

1. Went to zappos.com and browsed some shoes
2. Went to cnn.com


Zappos.com
POWERED by SERVICE®
Shoes, Clothing, Bags, etc.
Search
Shoes
Clothing
Bags & Handbags
Accessories
Boutiques
Women's
ALPHABETICAL BRAND INDEX
# • A • B • C • D • E • F • G • H • I • J • K • L •
« Back | Shoes › Boots › Georgia Boot

Georgia Boot Homeland 10" Wellington Waterp


CNN
U.S. | World | Politics | Money | Opinion | Health | Entertainmen
Top stories
BREAKING NEWS
Whe

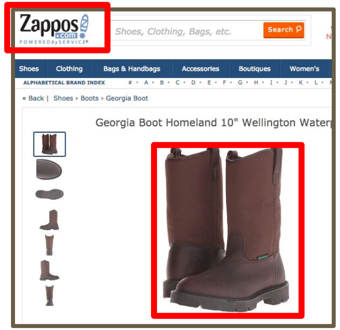
Weiner, Abedin separating after

How a Targeted Web Ad Happens

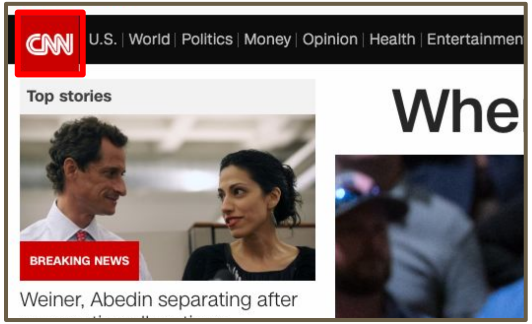
You just installed Chrome

1. Went to zappos.com and browsed some shoes
2. Went to cnn.com
3. On the very first story you go to, you see a picture of the shoes you were just looking at!

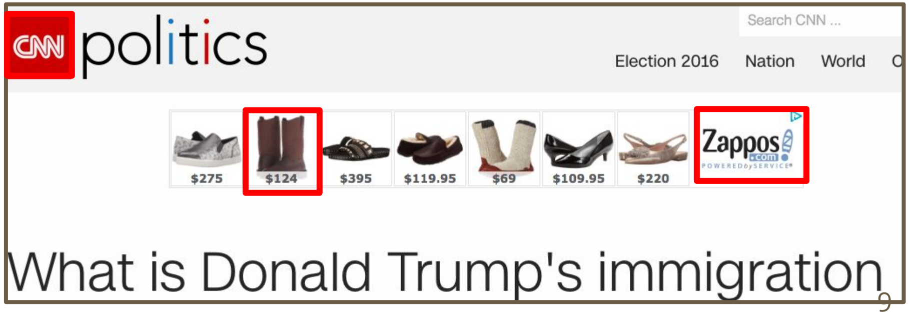
What happened?



Georgia Boot Homeland 10" Wellington Water

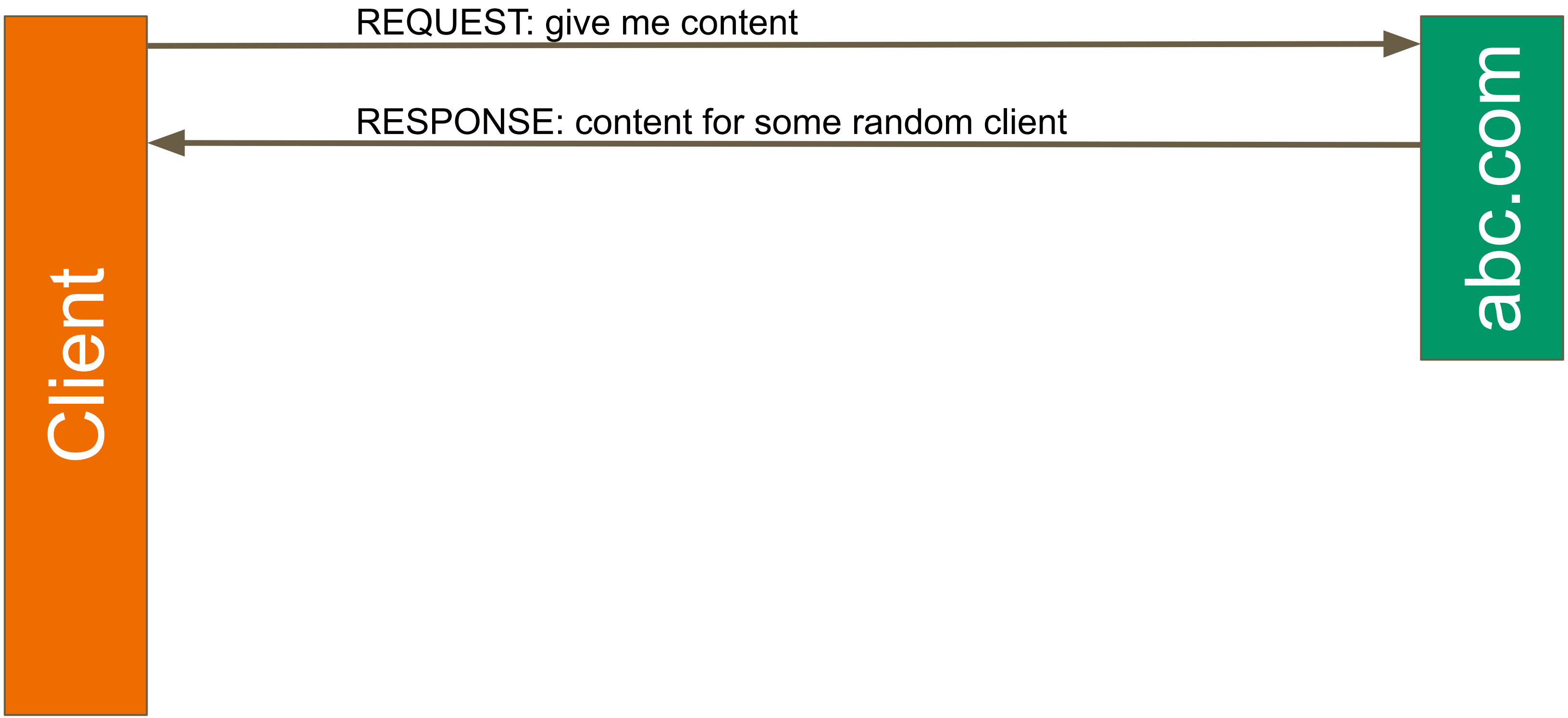

CNN
U.S. | World | Politics | Money | Opinion | Health | Entertainmen
Top stories
BREAKING NEWS
Whe

Weiner, Abedin separating after



What is Donald Trump's immigration

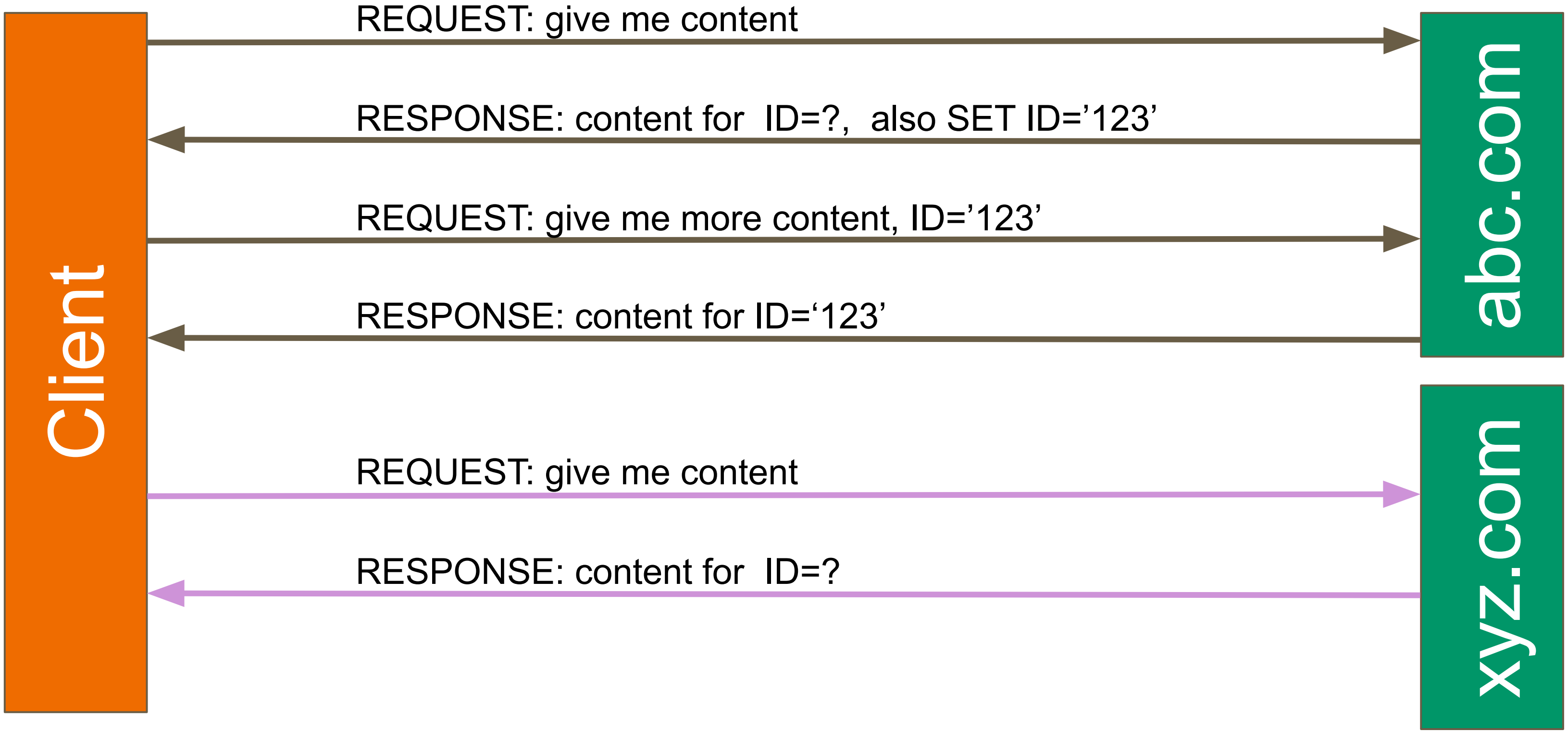
HTTP Requests



REQUEST: give me content

RESPONSE: content for some random client

Cookies



REQUEST: give me content

RESPONSE: content for ID=?, also SET ID=’123’

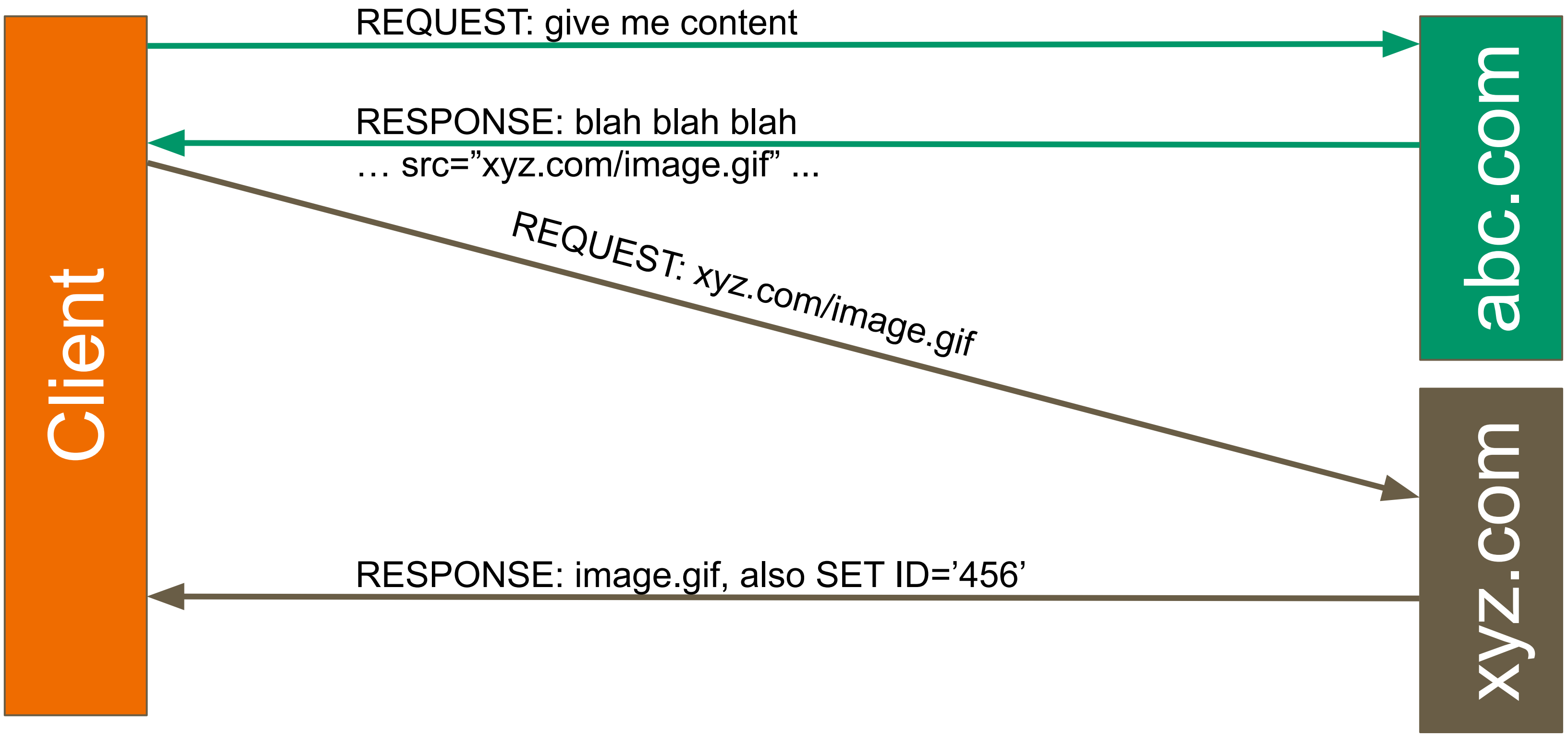
REQUEST: give me more content, ID=’123’

RESPONSE: content for ID=‘123’

REQUEST: give me content

RESPONSE: content for ID=?

Third-Parties



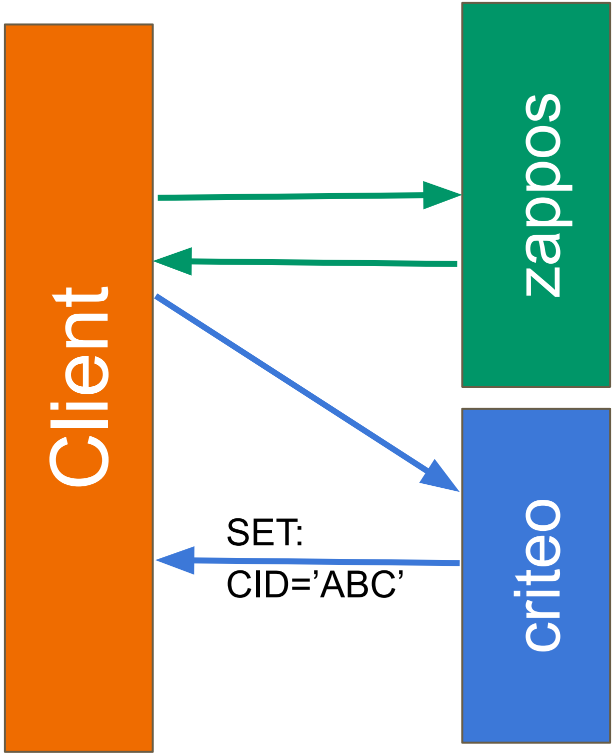
REQUEST: give me content

RESPONSE: blah blah blah
… src=”xyz.com/image.gif” ...

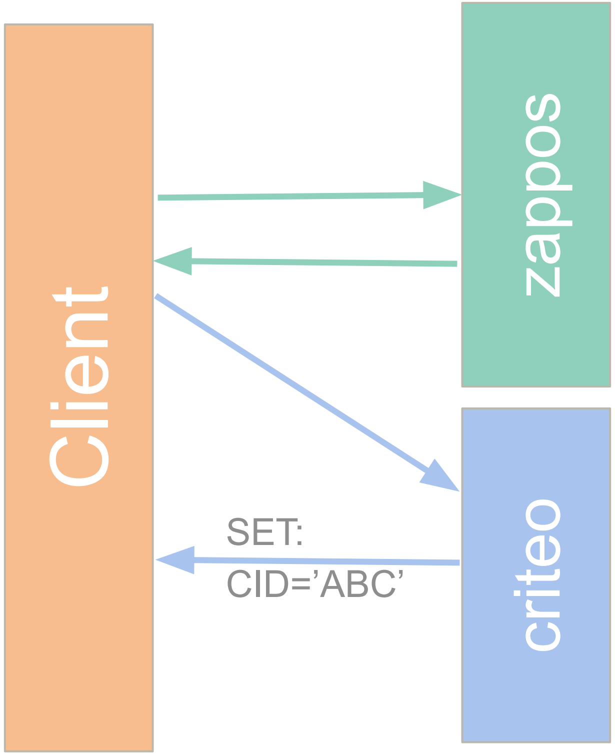
REQUEST: xyz.com/image.gif

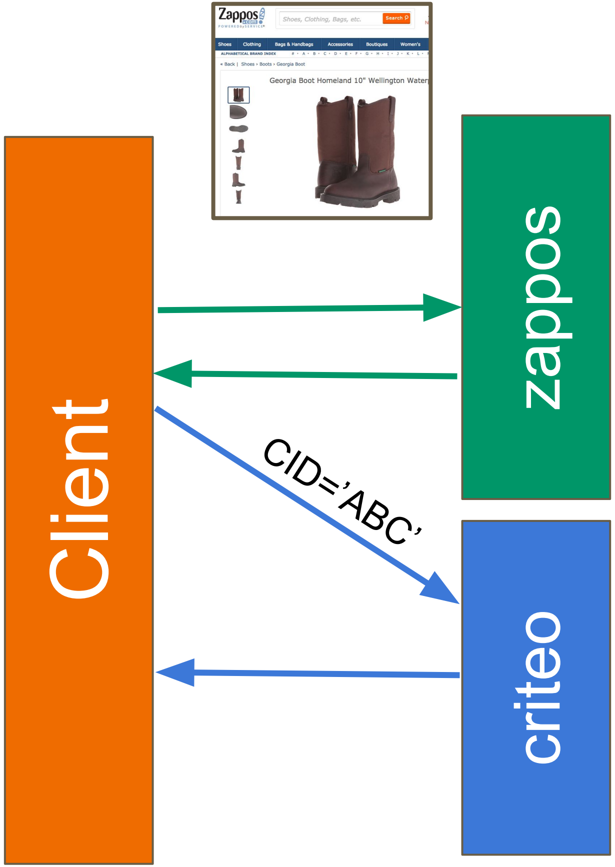
RESPONSE: image.gif, also SET ID=’456’

Visiting Zappos



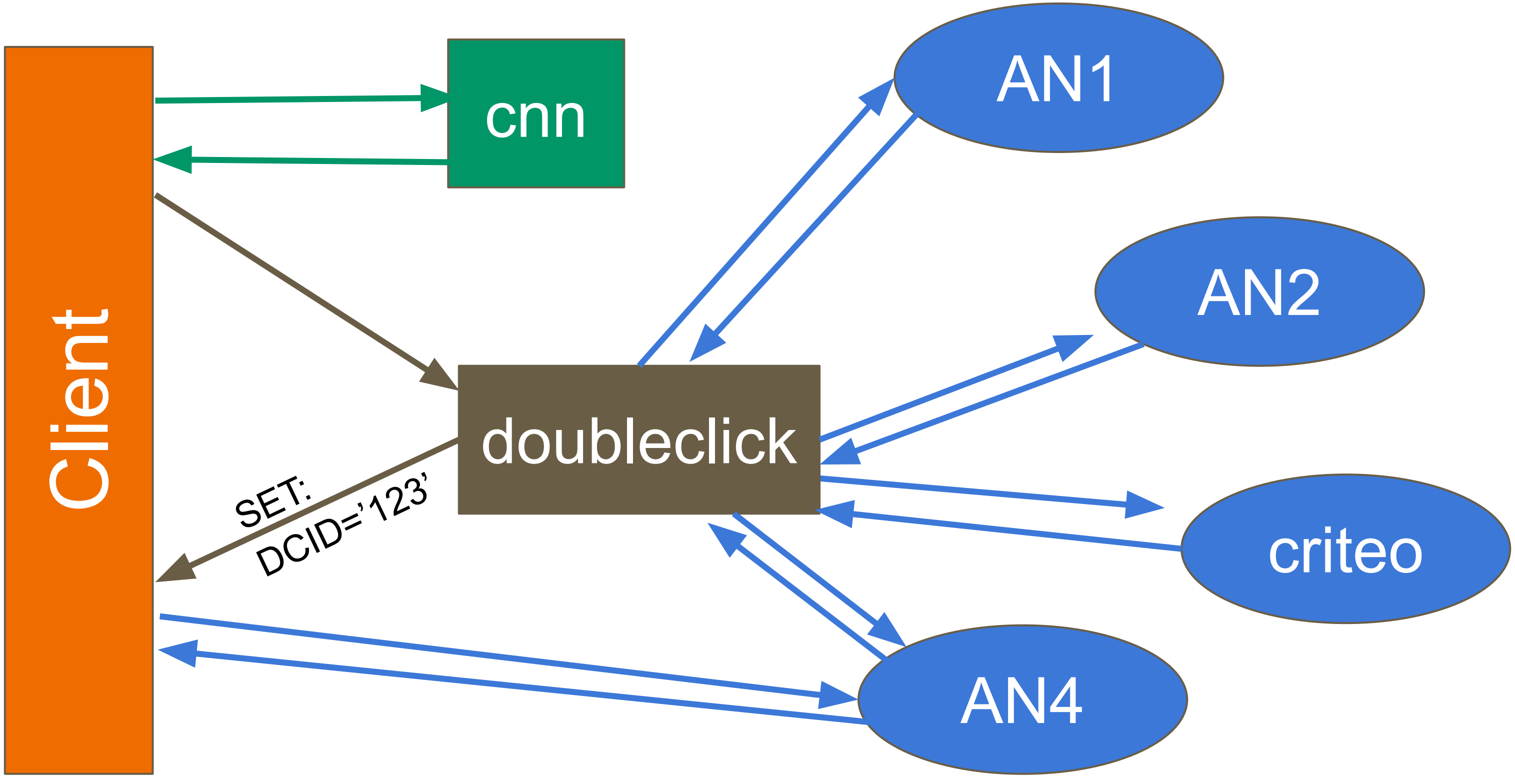
Visiting Zappos, More



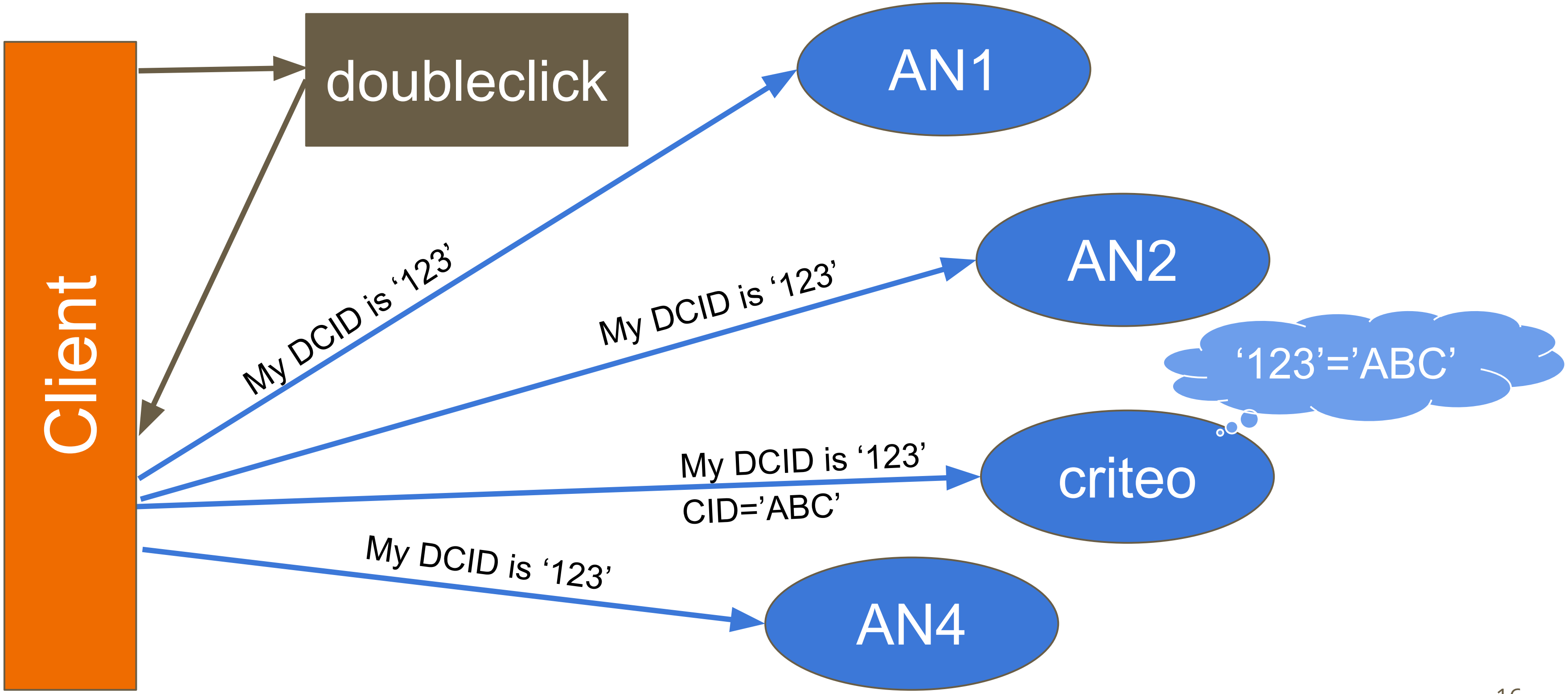


Georgia Boot Homeland 10" Wellington Water

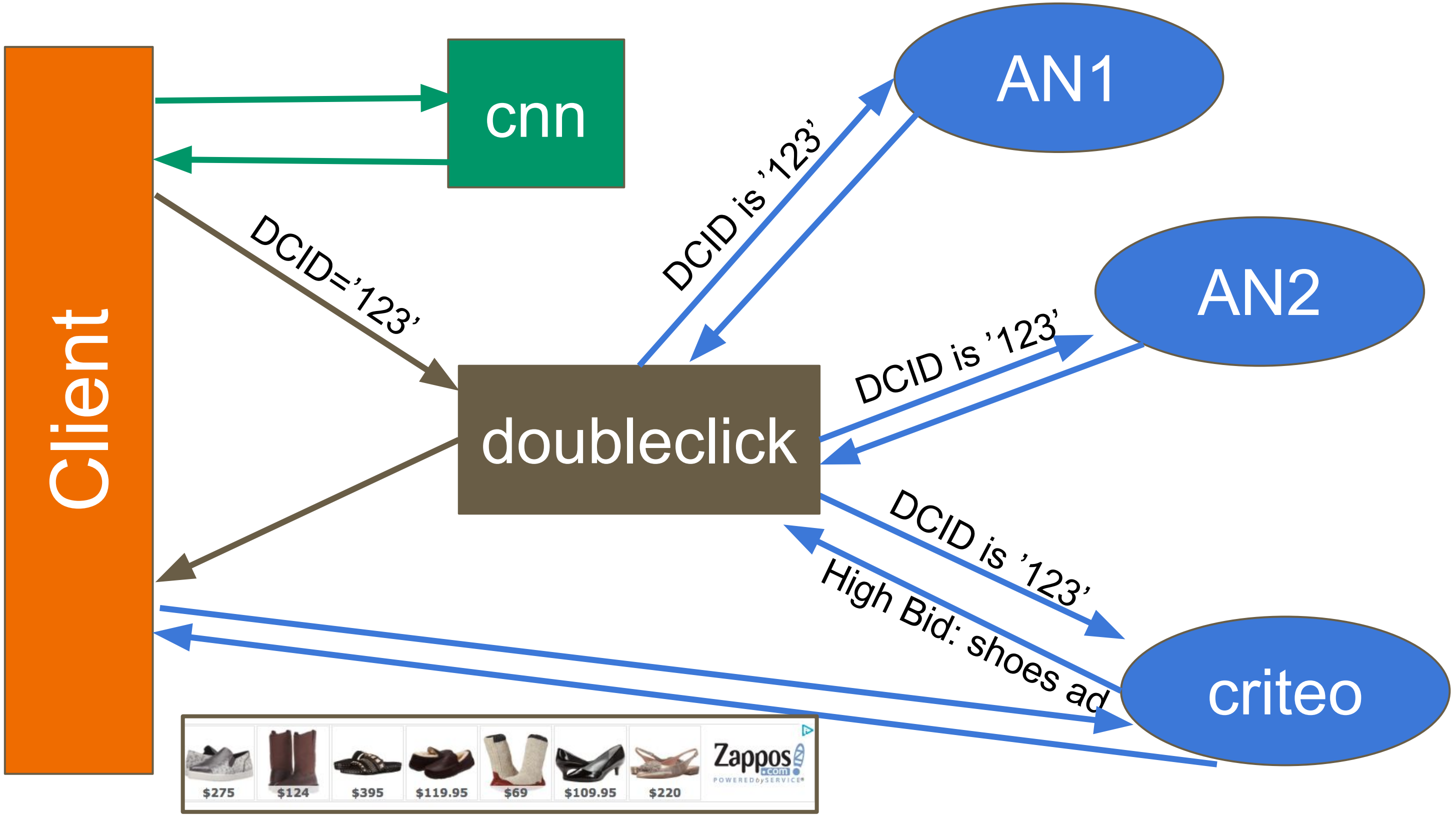
Visiting CNN pt. 1



Cookie Syncing (Piggyback Requests)



Visiting CNN pt. 2



Viewing the Ad

- Ads aren't just images
- Videos
- Flash Objects (good thing Flash is secure)
- JavaScript (maybe restricted by Ad-Network, maybe not)
- Web-beacons
  - Requests to paying entity's server (e.g. zappos.com) to combat fraudulent ad-networks
- Allowed formats vary by ad-network and ad-exchange (more later)

Mobile Ads

- Similar concepts, but simpler
- Single device-wide “Cookie”
  - Google Advertising ID (GAID)
  - ID for Advertisers (IDFA)
- No cookie synching necessary!
- Tracking/Advertising Libraries Integrated in Apps

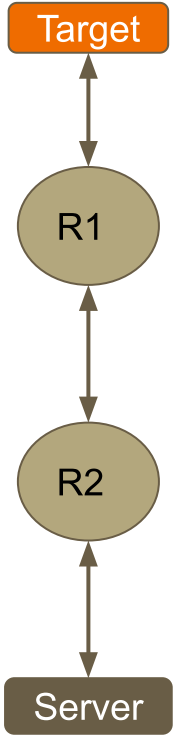
Malicious Ad Content

- Old Method: Trick or hijack users to visit your sketchy website
- Malvertising: Send it to them as an ad

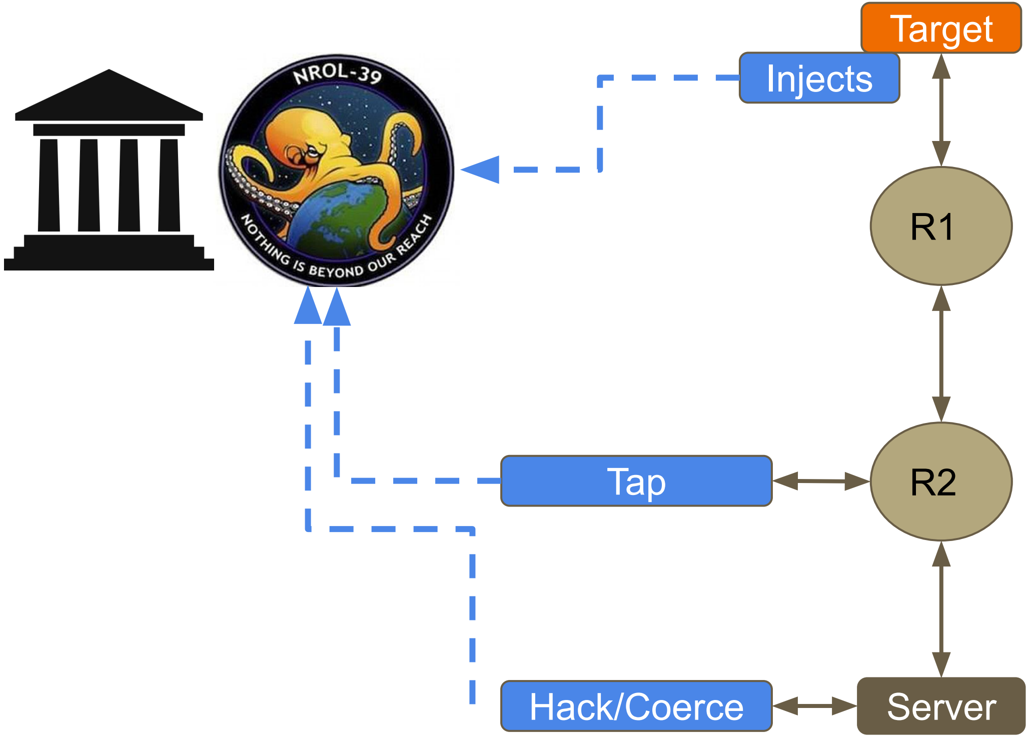
- Ad Networks claim to audit ad content
  - This varies in thoroughness
- Customers can self-host or use 3rd-party hosting services
  - Removes audit’s effect
- Continuing problem
  - Major sites hit ~2014
  - Targeting of DoD contractor IP addresses in a few cases

?

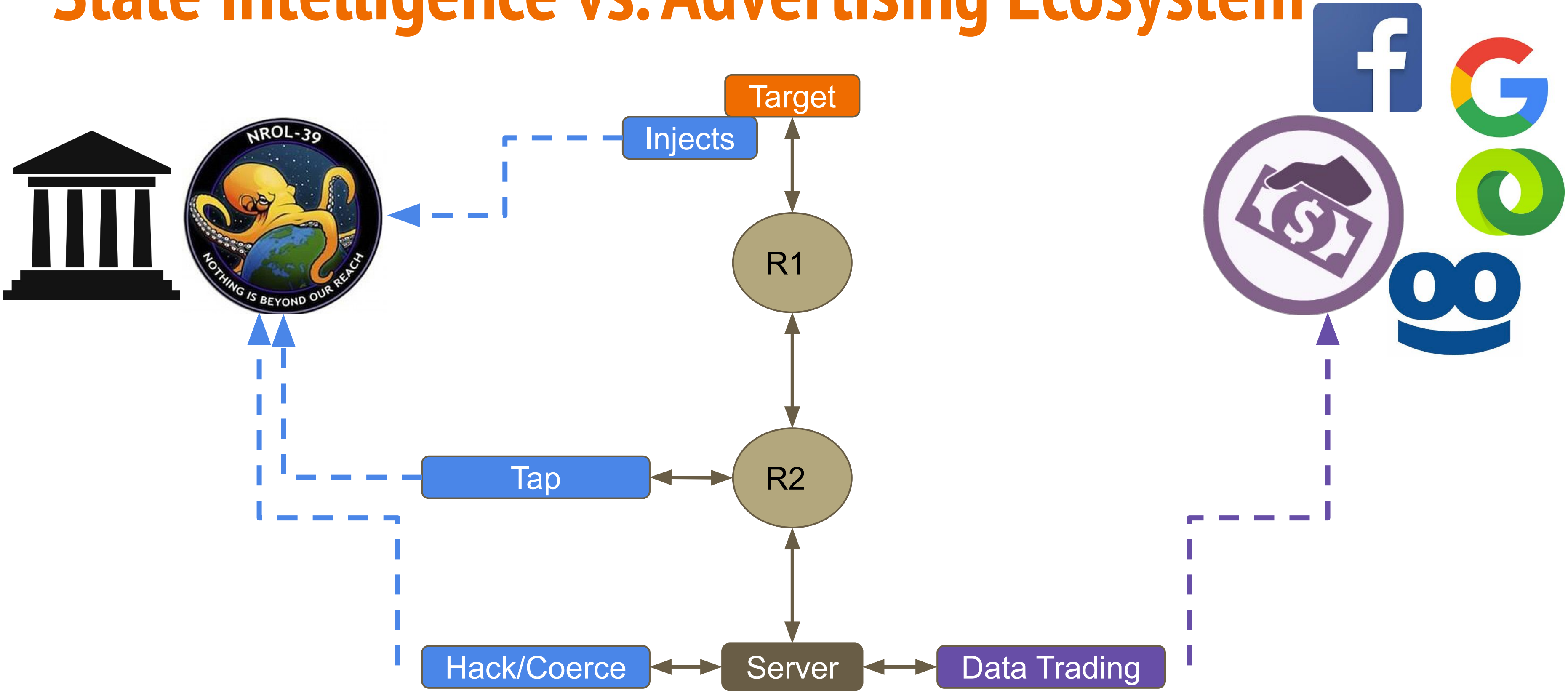
State Intelligence vs. Advertising Ecosystem



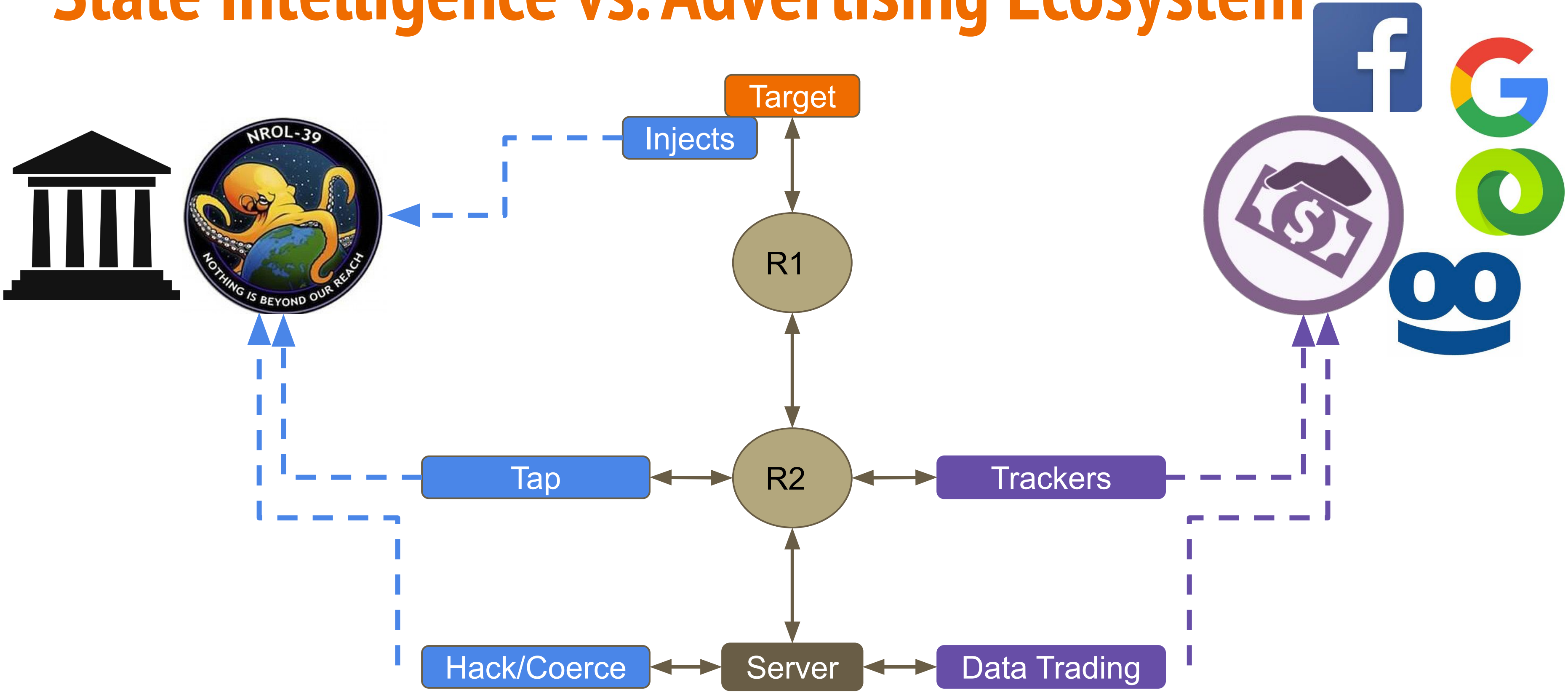
State Intelligence vs. Advertising Ecosystem



State Intelligence vs. Advertising Ecosystem



State Intelligence vs. Advertising Ecosystem



Target

Injects

R1

R2

Tap

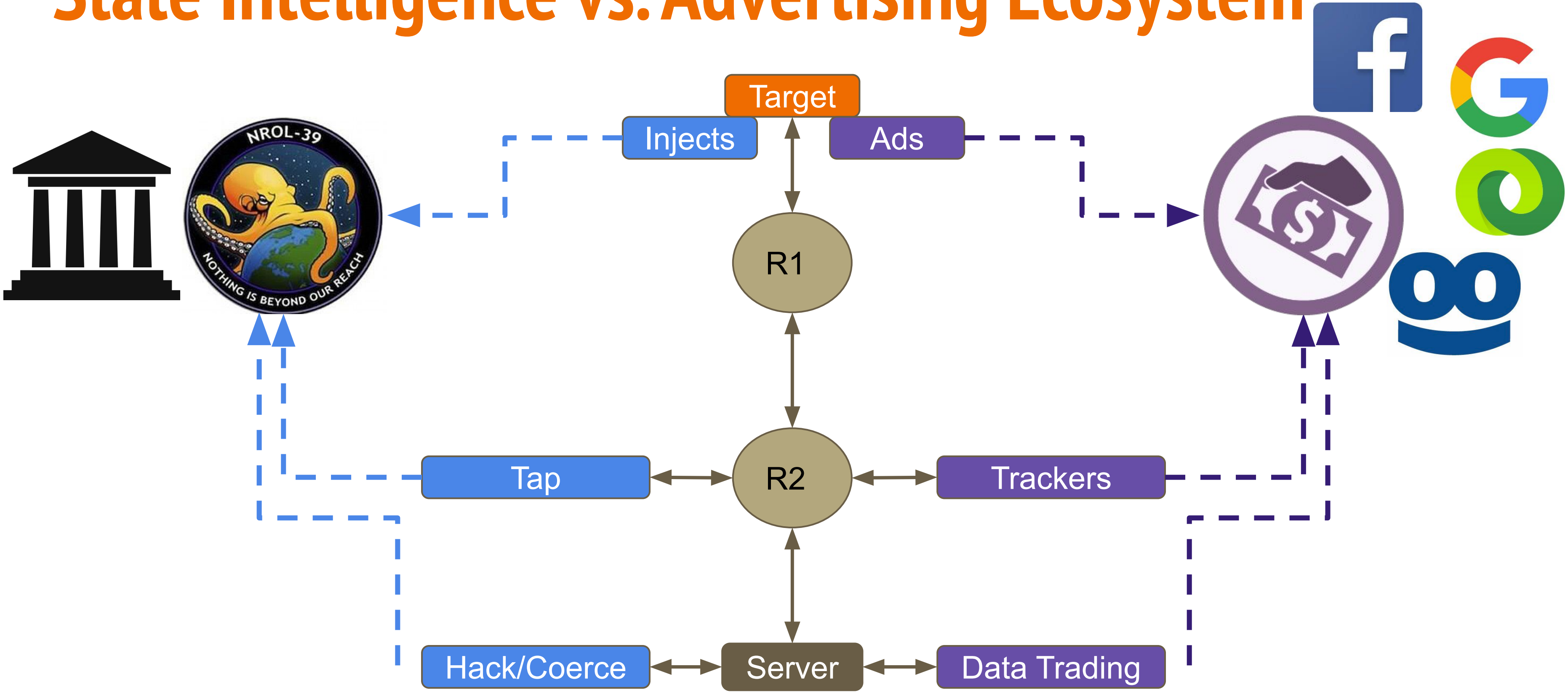
Trackers

Hack/Coerce

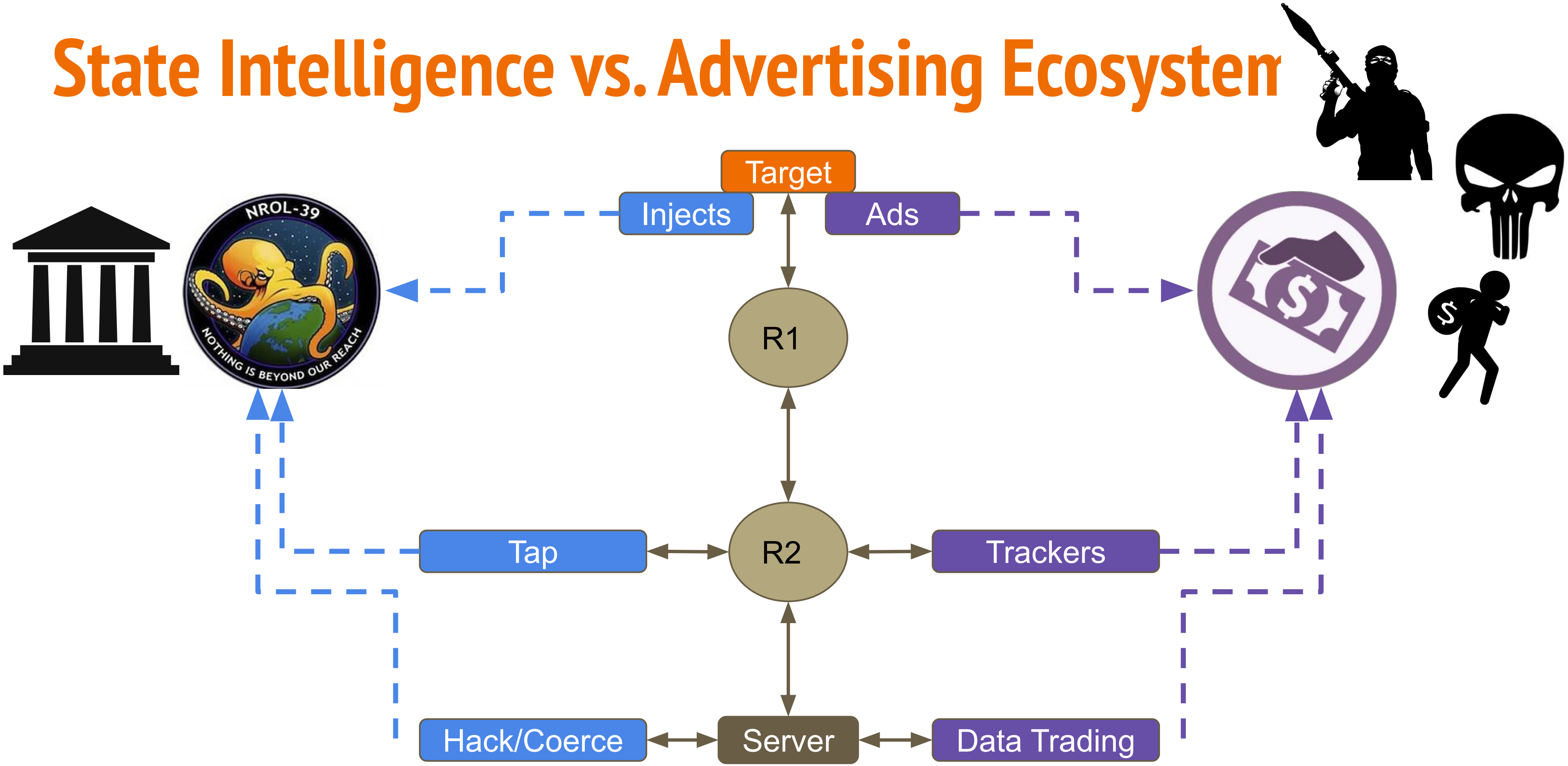
Server

Data Trading

State Intelligence vs. Advertising Ecosystem



State Intelligence vs. Advertising Ecosystem



ADINT: Surveillance via Advertising

What does the advertising ecosystem know? -- Not just websites you visit

Pages You Visit	Real Name
Email Address	Physical Addresses
Search Keywords	Gender and Age
Apps You Use	Sexual Orientation
WiFi Networks	Physical Location
Interests	Offline Purchases
Employer	Income
...	...

ADINT Core Concept: Ad Targeting as an Oracle

How old is alice@gmail.com → Make these ads:

Email=alice@gmail.com AND Age=18
Email=alice@gmail.com AND Age=19
Email=alice@gmail.com AND Age=20
Email=alice@gmail.com AND Age=21
Email=alice@gmail.com AND Age=22
Email=alice@gmail.com AND Age=23
...
Which one got served?

But Will ADINT Work? A Case Study

- [redacted]
  - Demand-Side Provider (DSP), facilitates advertisers buying ads
  - Specialized in Mobile
  - Offers “Hyperlocal” targeting
- Focused on Physical Location
  - Obvious use for surveillance
  - Concerning privacy implications
  - Dynamic targeting data

Benchmarking

Practical Operational Details Unknown

1. **How quickly do our ads get served?** -- 3m:30s
2. **How often will we see our ads?** -- 80% of opportunities
3. **How much will these ads actually cost?** -- Half-a-penny
4. **How precise and accurate is the location targeting?** -- 8-meters

Methodology

- Combination of Fake and Real User Devices
  - Found no differences in cost or frequency of ad serves
- Fake Users:
  - Android 4.4.4
  - 27 Year-Old Women
  - Created Gmail, Facebook, and Twitter accounts
  - Turned on Location, Logged into WiFi

Apps	Installs	Location Ads
The Chive	5-10M	✓
Grindr	10-50M	✓
iFunny	10-50M	
Imgur	5-10M	✓
MeetMe	1-5M	✓
My Mixtapez	10-50M	✓
Talkatone	10-50M	✓
TextFree	10-50M	✓
TextMe	10-50M	✓
TextPlus	10-50M	✓
Words with Friends	50-100M	

Case Study Threat Model: A Stalker

Stalker/Adversary:

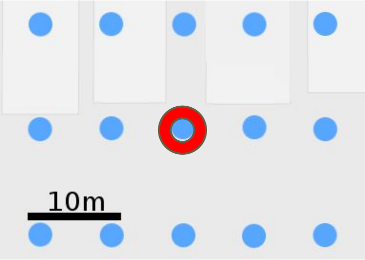
- Access to DSP → $1,000
- Knows Target's Mobile Advertising ID
  1. Sniffing Network Traffic
  2. Target clicked on ad in past
  3. Served ad to certain ad-libraries or exploited

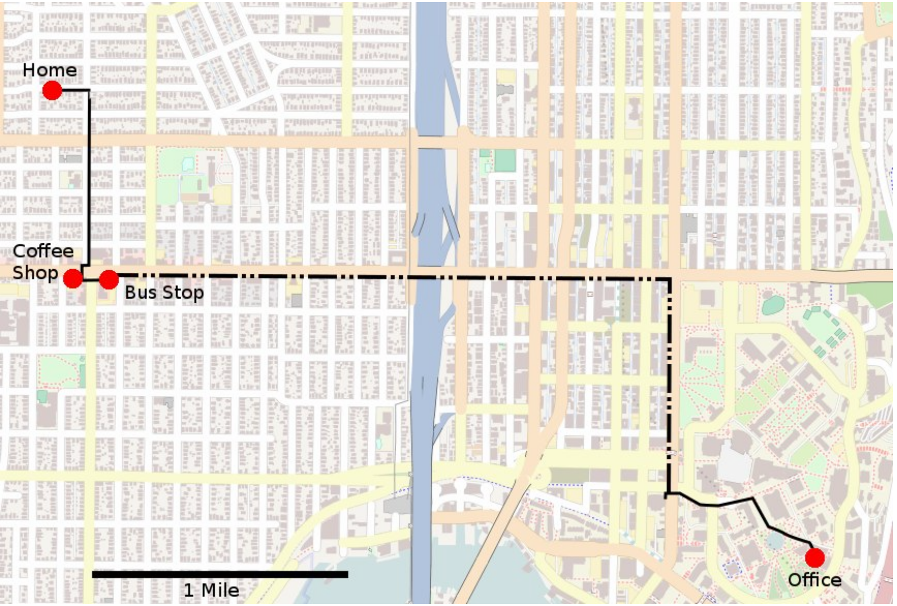
Target:

- Uses an app ads can be served to (Talkatone)

Attack 1: Such a lovely home, but that commute!

- **Goal: Determine the Location of:**
  - **Home**
  - **Office**
  - **Frequent Hangouts**
- **Method:**
  - **Create grid of location ads**
  - **Observe which are served and when**



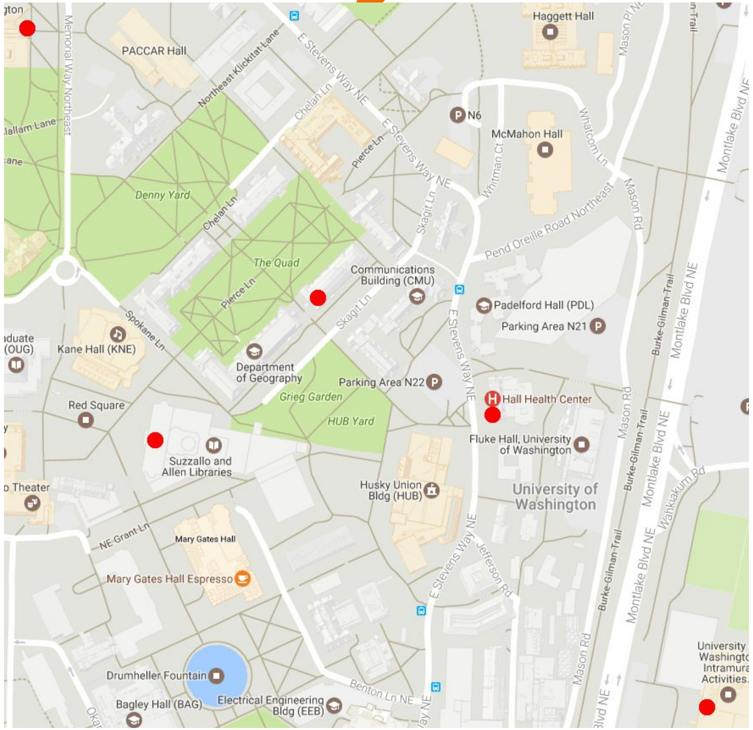


Attack 2: I really wish you wouldn't go there...

Hypothesize the target *might* go
But maybe only once

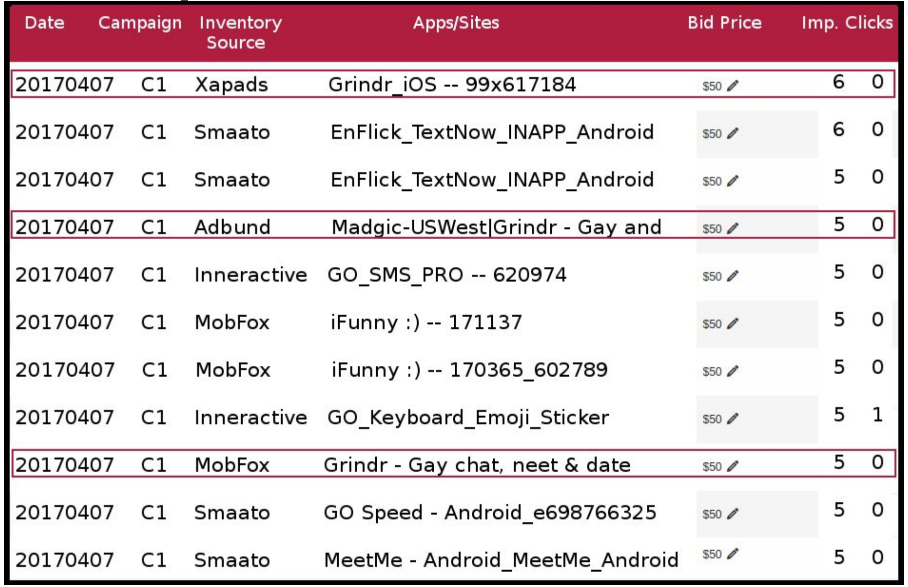
- **Specialized Medical Centers**
- **Police Station**
- **Religious Centers**
- **Rival Businesses**

All ads served within <u>5 minutes</u>
Some within <u>1 minute</u>



Attack 3: Whatcha up to on there?

DSP reports where our ads are shown:



Date	Campaign	Inventory Source	Apps/Sites	Bid Price	Imp.	Clicks
20170407	C1	Xapads	Grindr_iOS -- 99x617184	$50 ✎	6	0
20170407	C1	Smaato	EnFlick_TextNow_INAPP_Android	$50 ✎	6	0
20170407	C1	Smaato	EnFlick_TextNow_INAPP_Android	$50 ✎	5	0
20170407	C1	Adbund	Madgic-USWest|Grindr - Gay and	$50 ✎	5	0
20170407	C1	Inneractive	GO_SMS_PRO -- 620974	$50 ✎	5	0
20170407	C1	MobFox	iFunny :) -- 171137	$50 ✎	5	0
20170407	C1	MobFox	iFunny :) -- 170365_602789	$50 ✎	5	0
20170407	C1	Inneractive	GO_Keyboard_Emoji_Sticker	$50 ✎	5	1
20170407	C1	MobFox	Grindr - Gay chat, neet & date	$50 ✎	5	0
20170407	C1	Smaato	GO Speed - Android_e698766325	$50 ✎	5	0
20170407	C1	Smaato	MeetMe - Android_MeetMe_Android	$50 ✎	5	0

Dating Apps	**Torrenting Apps**
Grindr	BitTorrent
Hornet	FrostWire
Jack'D	uTorrent
Meet24	**Other**
MeetMe	Adult Diapering Diary
Moco	Hide My Texts
Romeo	Hide Pictures Vault
Tagged	Pregnant Mommy's Maternity
Wapa	Psiphon
Wapo	Quran Reciters

Case Study Summary

- Serve ads to real and fake users: 80% of auctions won, $0.005 / ad
- Find any location a target visits for more than 4 minutes while using apps
- Enumerate ad-containing apps a target uses
- Know *when* a target uses an app
- Enumerate members of a crowd

?

A Survey of DSP Capabilities

	Min. Cost	Targeting									Content			
DSP		*Demographics*	*Interests*	*PII*	*Cookie/MAID*	*Device*	*Network*	*Location*	*Domain/App*	*Search*	*HTML*	*Flash*	*3rd*	*Beacon*
Admedo	$5,000	✓	✓+	✓	✓	✓	✓+	✓+	✓	-	✓	✓	✓	✓
AdRoll	$0	-	-	✓	✓	-	-	-	-	-	-	-	-	-
AdWords	$0	✓+	✓+	-	✓	✓	✓+	✓+	✓+	✓	✓	✓	-	✓
CaseStudyDSP	**$1,000**	**✓**	**✓**	**-**	**✓**	**✓**	**✓+**	**✓+**	**✓**	**-**	**✓**	**-**	**✓**	**✓**
Bing	$0	✓	✓	-	✓	✓	✓	✓	✓	✓	-	-	-	-
Bonadza	$300	✓+	-	-	✓	✓	-	✓+	✓	-	✓	-	✓	✓
BluAgile	$1,000	✓+	✓	-	✓	✓	✓+	✓+	✓	-	✓	✓	✓	✓
Centro	$5000 / month	✓+	✓+	✓	✓	✓	✓+	✓+	✓	✓	✓	✓	✓	✓
Choozle	$99 / month	✓+	✓+	✓	✓	✓	✓+	✓	✓	-	✓	✓	✓	✓
Criteo	$0	-	-	-	✓	-	✓	✓	-	-	✓	✓	✓	✓
EactDrive	$50	✓	✓	-	✓	-	✓	✓+	✓	✓	✓	✓	✓	✓
Facebook	$0	✓+	✓+	✓	✓	✓	✓	✓+	-	-	-	-	-	-
GetIntent	$0	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Go2mobi	$0	-	-	-	✓	✓+	✓+	✓	✓+	-	✓	✓	✓	✓
LiquidM	$1,000	✓	✓	-	✓	✓	✓	✓+	✓	-	✓	-	✓	✓
MediaMath	$50,000 / month	✓+	✓	✓	✓	✓	✓	✓+	✓	-	✓	✓	✓	✓
MightyHive	$2,000	✓	✓	✓	✓	✓	✓	✓+	✓	✓	✓	✓	✓	✓
Simpli.fi	$10,000	✓	✓+	-	✓	✓	✓+	✓+	✓	✓	✓	✓	✓	✓
SiteScout	$500	-	✓+	-	✓	✓	-	✓	-	-	-	✓	-	-
Splicky	$0	✓	✓	-	✓	✓	✓	✓+	-	-	-	-	-	-
Tapad	$2,000	✓	✓+	✓	✓	✓	✓	✓+	✓	✓	✓	✓	✓	✓

Costs of ADINT

81% (17/21) DSPs cost ≤ $2,000 initially

Per-Ad Costs are Negligible

Technical Sophistication is Low

- **Only *active ad content* requires code**

Lots of Potential Users

Technical Targeting

Location:

- Zip-Code/City Targeting - 18 / 21
- Hyperlocal Targeting - 14 / 21

IP Address - 8 / 21

- Enumerate household devices
- Semi-unique identifier
- Dense Residential Location Targeting

Personal Targeting

Interests & Demographics

- 394 IAB Interest Categories - 17/21
  - Include: Religion, Health Issues, Addictions, Immigration
- Age/Gender/Language - 17/21
- More In-Depth:
  - Ethnicity, Sexual Orientation, Employer, Job Title, Income, Finances, Personality Type...

Personally Identifying Information (PII)

- Target by Email - 8/21
- Target by Real Name or Physical Address - 1/21

Active Ad Content

- Web Beacon can report to us when an ad is displayed
    - IP Address
    - User-Agent String
- Using JavaScript, we could:
    - Fingerprint the device
    - Exfiltrate Location if permitted
    - Exfiltrate Mobile Advertising ID
- Web Beacon to other entities → Set Tracking Cookies

Putting It All Together

- Key challenge: strong identifiers (MAID is best, but harder to get)
- An ADINT campaign operation loop:
  1. Target aspect of interest
     - interest, location, app, searches…
  2. Obtain identifiers from ad serves
     - IP address, device fingerprint, location, MAID…
  3. Target useful info *using* identifiers
     - home/office location, employer, associations…

Who Would Use ADINT? Burglars

- **Previously: Find social media posts about vacations**

- **Find targets by financial status, employer, or luxury purchases and interests**
- **Use location targeting to find residence**
- **IP target to enumerate all residents**
- **Location targeting to determine when they are all away**

Who Would Use ADINT? Ideological Vigilantes

- Previously: Social media posts or social app usage

- Find targets by whatever is relevant:
    - Sensitive app use
    - Searches and website visits
    - Interests and demographics
- Use more targeting to confirm improper behavior
- Location targeting to find and confront
- Ad content to harass

Who Would Use ADINT? Who Wouldn’t?

- Journalists
- Paparazzi
- Investors
- Employers
- Law Enforcement (esp. Local)

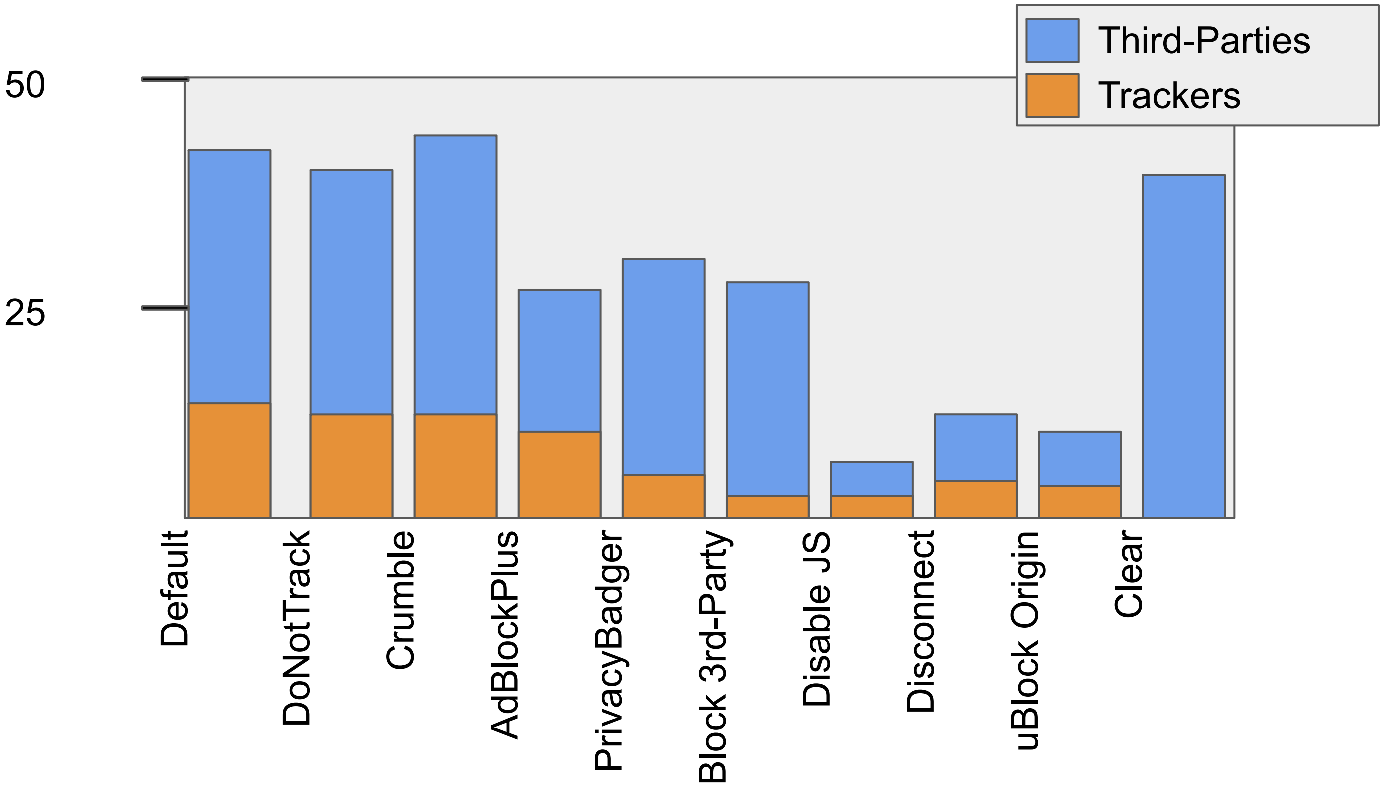
ADINT Summary

- Low Entrance Costs ($0-$2,000 for 80% DSPs)
- Online Surveillance
    - Websites visited
    - Apps used
    - Searches made
- Personal Surveillance
    - Demographics
    - Interests
    - Finances
- Location Surveillance
- Users don’t intend to share this with anyone

Defenses and The Future

1. User defenses
2. Ad Network Self-Policing
3. Regulations

User Defenses



User Defenses Cont.

- Tracking defenses
  - Prevent some information leakage (browsing history)
  - Also prevent ADINT by preventing ad targeting and ad serving
- Mobile is a problem
  - Typically no extensions in browsers
  - Lots of ads in apps

Ad Network Self-Policing

- Never received complaint or suspicion from case study DSP
- 2010 Korolova et al. prompted Facebook to institute 20-person minimum
- Preventing trivial use and detecting suspicious ads can help
  - Raise the bar until ADINT is as costly as conventional surveillance tech
- *Some* user-based ad networks have motivation
  - Google and Facebook have users to worry about
  - Most ad networks have no “users”

Regulations

- Current backlash against advertising ecosystem
    - Unclear this will lead to anything
- GDPR in the EU increases transparency
    - Likely does not significantly affect ADINT capabilities
- Ad networks want customers to engage in ADINT-like behavior
    - Current market direction is still towards more information and specificity.

?