Privacy Risks with Facebook's PII-based Targeting

Auditing a data broker's advertising interface

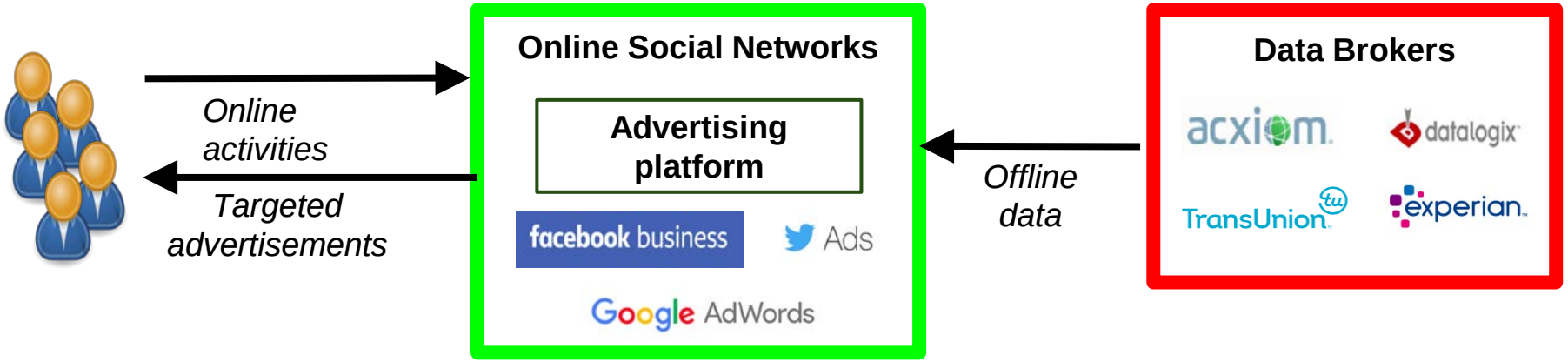
Giridhari Venkatadri[1], Athanasios Andreou[2], Yabing Liu[1], Alan Mislove[1], Krishna P. Gummadi[3], Patrick Loiseau[3,4], Oana Goga[4]

[1]Northeastern University, [2]EURECOM, [3]MPI-SWS
[4]Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG

PRIVACYCON

Traditional Data Brokers

21st Century Data Brokers



Attribute-based Targeting

1. **Attribute-based targeting**

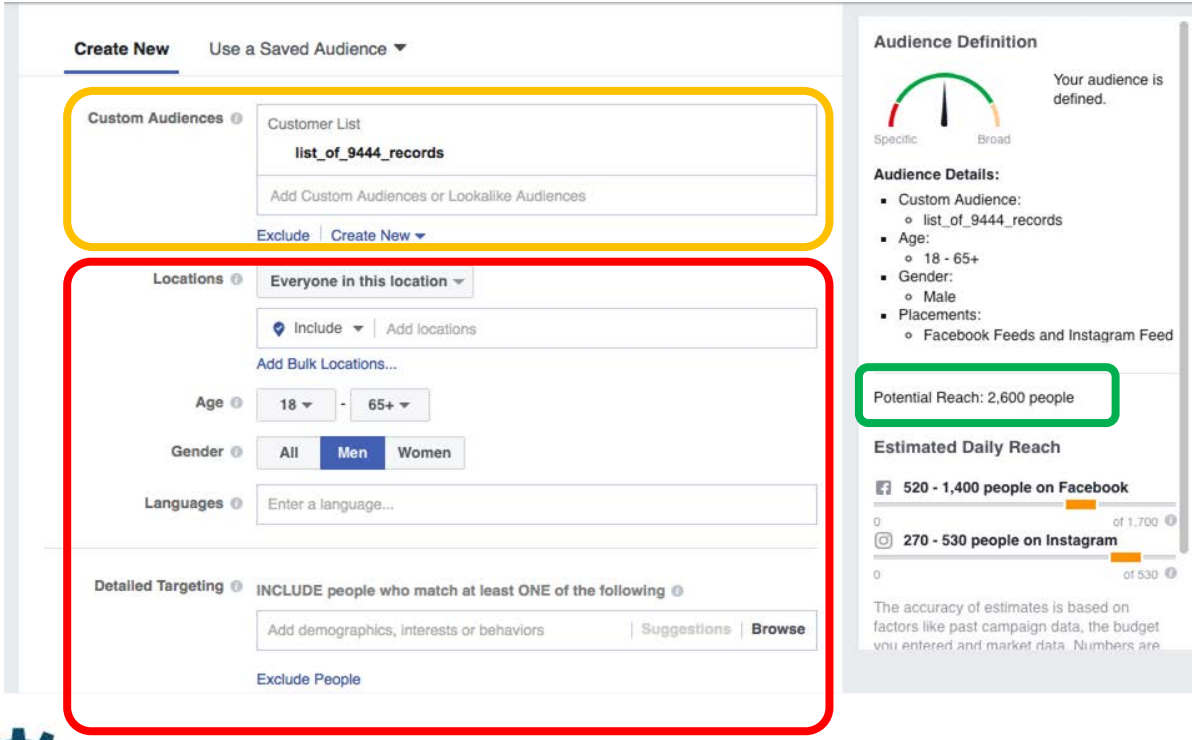
   Advertiser specifies attributes to create *audience* of users

2. **Audience size**

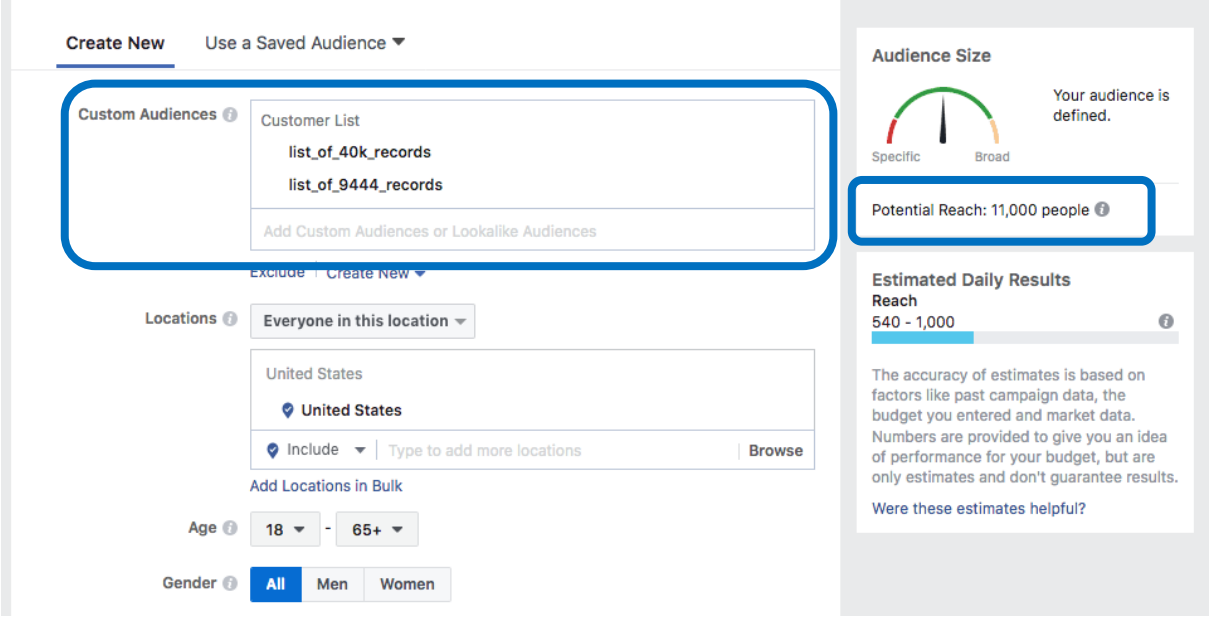
   Facebook provides *obfuscated* number of selected users

3. **Custom (PII) audiences**

   Advertiser can select via *PII of target users instead*



Custom Audiences



1. **Select fields to use**

   Advertiser specifies which fields they have on users

2. **Upload CSV file with user data**

   Advertiser uploads file to Facebook

3. **Facebook matches users**

   Advertiser is provided with an audience to advertise to

4. **Facebook provides statistics**

   Advertiser can obtain obfuscated size for audience combinations

What PII can be Used?

Site	Name	Email	Phone number	City or ZIP	State or Province	Birthday, Gender	Employer	Site user ID	Mobile advertiser ID
Facebook	✓	✓	✓	✓	✓	✓	✗	✓	✓
Instagram	✓	✓	✓	✓	✓	✓	✗	✓	✓
Twitter	✗	✓	✓	✗	✗	✗	✗	✓	✓
Google	✓	✓	✓	✓	✗	✗	✗	✓	✓
Pinterest	✗	✓	✗	✗	✗	✗	✗	✗	✓
LinkedIn	✗	✓	✗	✗	✗	✗	✓	✗	✓

Raises Concerns

Sites have **detailed data** from service, data brokers

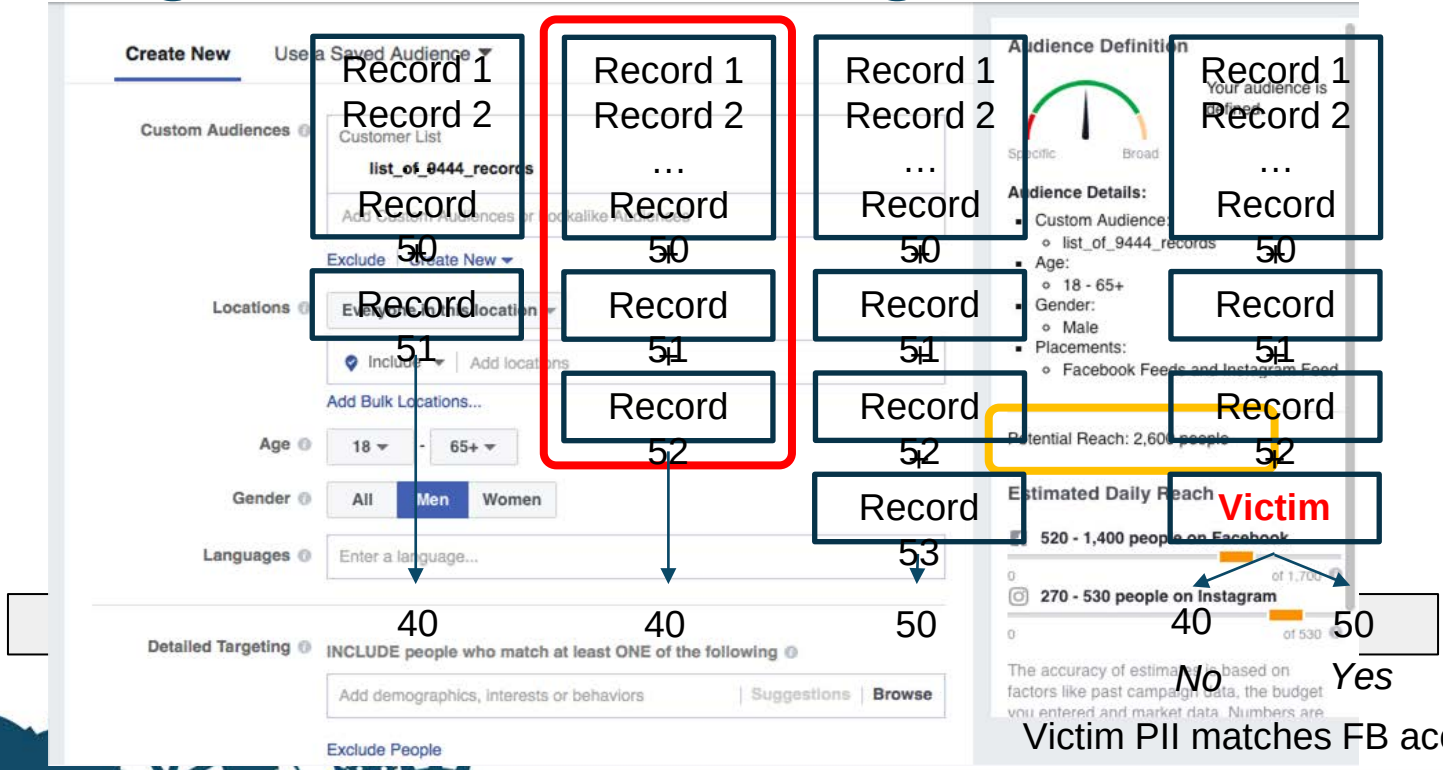
First time allow anyone to **link against that database**

Allow advertisers to upload data, obtain aggregate information

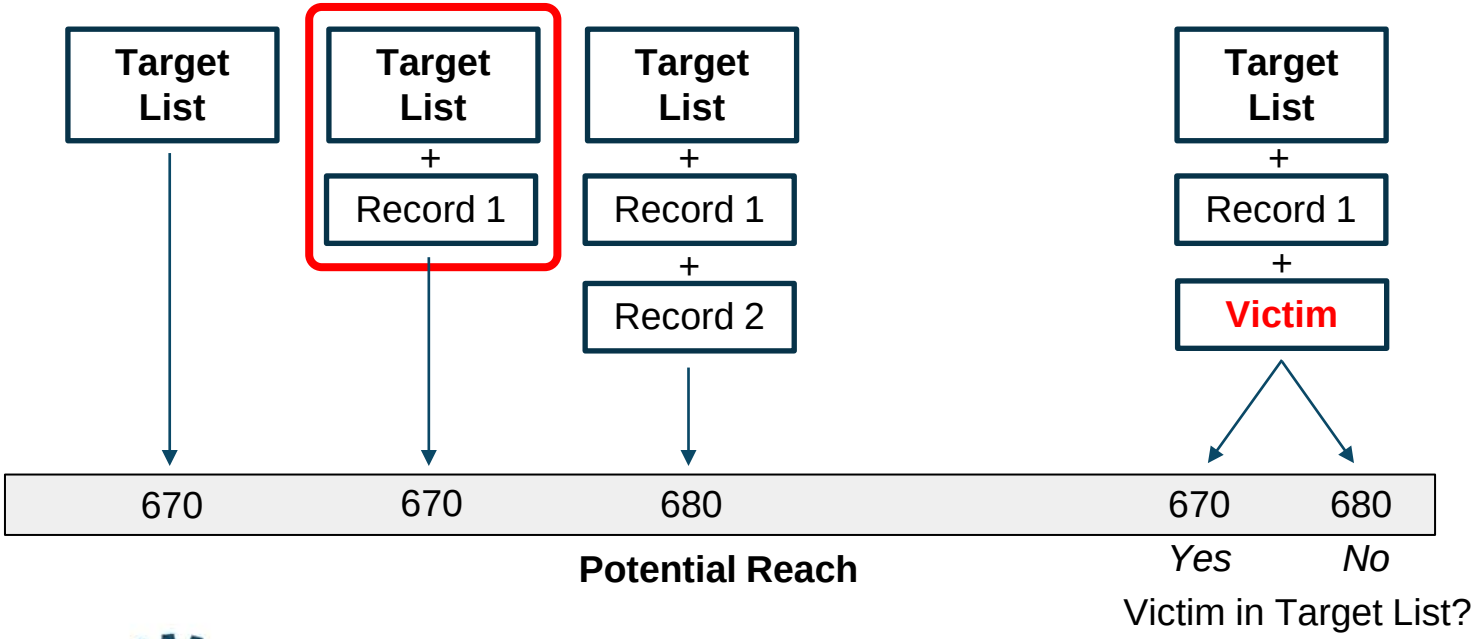
Question: Are attacks that leak user information possible?

Yes! With a few insights…

Insight 1: Overcoming Obfuscation



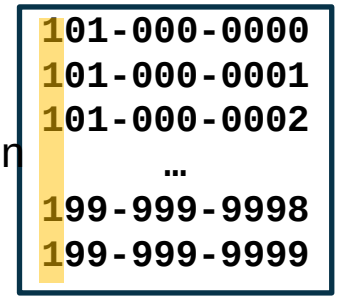
Insight 2: Determine if (matched) Victim is in List



Attack: Learning User's Phone Numbers

Can ask: Is **Victim** in **Target List** ?

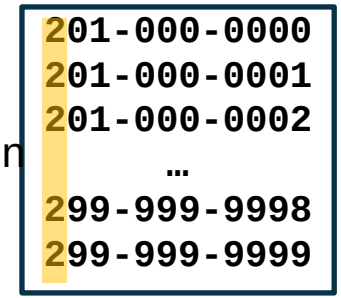
Is **Victim** in



*If **No***: First digit is not 1

*If **Yes***: First digit is 1

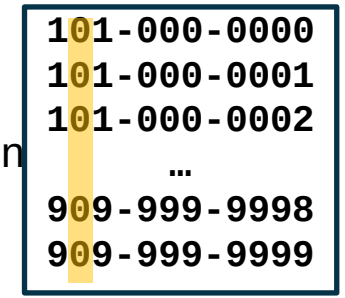
Is **Victim** in



*If **No***: First digit is not 2

*If **Yes***: First digit is 2

Is **Victim** in



*If **No***: Second digit is not 0

*If **Yes***: Second digit is 0

Resulting Attacks

The attacks we discover allow attacker to:

- Link multiple pieces of PII to a single user
- Infer any active Facebook user's phone number
- De-anonymize visitors to attacker's website

No ads placed, no victim interaction, no way for victim to detect attack

Reported to Facebook, **mitigation in place**

Implications

Many online services now de-facto data brokers; use data for advertising

- Anyone can be an advertiser!
- Interfaces allow advertisers to query this database, in obfuscated form

Our work shows that **interfaces can inadvertently leak user data**

- Highlights the need to audit advertising interfaces for privacy leaks

More info? Check out our IEEE S&P'18, NDSS'18, and FAT*'18 papers

https://mislove.org