Targeted Advertising on Social Media: Transparency, Control, Fairness, Privacy

Till Speicher, Nina Grigic-Hlaca, Athanasios Andreou, Giridhari Venkatadri, Oana Goga, Patrick Loiseau, Krishna P. Gummadi, Alan Mislove

Max Planck Institute for Software Systems
This talk: Focus on Facebook

- By far, the largest social media platform
  - In terms of number of users
  - In terms of data aggregated on users
  - In terms of advertisers & ad revenues
  - In terms of introducing novel & provocative targeting practices

- However, many issues discussed generalize to other social media platforms
  - Like LinkedIn, Twitter, YouTube, Instagram etc.,
Background: Facebook ad platform
Ads on Facebook

Spotify Premium für 0,99 €. Für alle, die gerne abtanzen.

€29.39
http://www.lightinthebox.com/Herrenhalbschuhe

Stainless Steel Cooler Stones - 4Pcs
gearbest.com
Free Shipping #CoolerStones #iceball #coolsummer

English (US) · Español (España) · Català · Português (Brasil) · Français (France)
Data used for targeting ads

- Facebook gathers lots of data (features) on users

<table>
<thead>
<tr>
<th>Category</th>
<th>Behaviors</th>
<th>Demographics</th>
<th>Interests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anniversary</td>
<td></td>
<td>Business and industry</td>
</tr>
<tr>
<td></td>
<td>Consumer Class</td>
<td></td>
<td>Entertainment</td>
</tr>
<tr>
<td></td>
<td>Digital activities</td>
<td></td>
<td>Family/relationships</td>
</tr>
<tr>
<td></td>
<td>Exports</td>
<td></td>
<td>Fitness and wellness</td>
</tr>
<tr>
<td></td>
<td>Mobile Device</td>
<td></td>
<td>Food and drink</td>
</tr>
<tr>
<td></td>
<td>User Multicultural Affinity</td>
<td></td>
<td>Hobbies and activities</td>
</tr>
<tr>
<td></td>
<td>Seasonal and Events</td>
<td></td>
<td>Shopping and outdoors</td>
</tr>
<tr>
<td></td>
<td>Travel</td>
<td></td>
<td>Technology</td>
</tr>
<tr>
<td>Facebook</td>
<td>1 2 3</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>Axiom</td>
<td>1 5 25</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Epsilon</td>
<td>4 1 29</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>Experian</td>
<td>3 1 11</td>
<td>11</td>
<td>350</td>
</tr>
<tr>
<td>Datalogix</td>
<td>151</td>
<td>21</td>
<td>350</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1 2 39 74 81 6 2 16 152 9 26</td>
<td>3</td>
<td>175 5 29 13 3 24 36 20 10 16 27</td>
</tr>
</tbody>
</table>
Closer look at features: Examples

- **Demographical** (gathered by Facebook)
  - **Relationship:**
    - Interested In: Men and Women, Men, Unspecified, Women
    - Status: Separated, Widowed, Open Relationship, Divorced, In a relationship, Married, Engaged, Unspecified, Single, Complicated Civil Union, Domestic Partnership
  - Each user feature is a **boolean variable**

- **Demographical** (aggregated from data brokers)
  - **Financial:**
    - Income: Geschätztes monatliches Nettoeinkommen 2.000 bis 2.600, 2.600 bis 3.600, 3.600 bis 5.000, über 5.000 Euro
## Data aggregation across countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Facebook</th>
<th>Epsilon</th>
<th>DLX</th>
<th>Experian</th>
<th>Acxiom</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>598</td>
<td>14</td>
<td>350</td>
<td>5</td>
<td>128</td>
<td>1105</td>
</tr>
<tr>
<td>UK</td>
<td>598</td>
<td>0</td>
<td>19</td>
<td>17</td>
<td>103</td>
<td>737</td>
</tr>
<tr>
<td>France</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>619</td>
</tr>
<tr>
<td>Germany</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>658</td>
</tr>
<tr>
<td>Australia</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>24</td>
<td>656</td>
</tr>
<tr>
<td>Brazil</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>618</td>
</tr>
<tr>
<td>Japan</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>615</td>
</tr>
<tr>
<td>South Korea</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>598</td>
</tr>
<tr>
<td>Canada</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>598</td>
</tr>
<tr>
<td>India</td>
<td>598</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>598</td>
</tr>
</tbody>
</table>
1) Traditional targeting: Advertisers specify a boolean formula over the features.

- Typically, in a restricted CNF form:
  - \((F_1 \lor F_2 \lor F_3 \ldots) \land (F'_1 \lor F'_2 \lor F'_3 \ldots) \land \ldots \land \neg F_K \land \neg F'_K\)

- Users are targeted, when their feature values inferred by Facebook satisfy the targeting formula.
How Facebook advertisers target users

2) Custom audience targeting: Advertisers upload PII of users they wish to target

- Advertisers love custom audience
This talk: 4 Aspects of Ad Targeting

1. Transparency
2. Control
3. Fairness
4. Privacy
This talk: 4 Aspects of Ad Targeting

1. **Transparency**
   - Can I know *what data about me is used to target ads?*

2. **Control**

3. **Fairness**

4. **Privacy**
Transparency of Facebook ad targeting

- Facebook provides transparency via explanations
- Largely voluntary
- At times to satisfy legal requirements
For ads targeting customer PII:

"One reason you're seeing this ad is that Booking.com wants to reach people who have visited their website or used one of their apps. This is based on customer information provided by Booking.com..“

Doesn’t state what PII Booking.com used!

E.g., Did booking.com use email? Or phone number? Or name/address?
How complete are the explanations?

- For ads targeting data broker info:
  - “One reason you’re seeing this ad is that Peugeot wants to reach people who are part of an audience created based on data provided by Acxiom. Facebook works with data providers to help businesses find the right audiences for their ads. Learn more about data providers.”

- Doesn’t state what Acxiom provided data was used!
  - E.g., is it based on financial data? Or purchasing habits?
For ads targeting FB user data:

- Beyond location, gender, age: picks exactly one of the several features used in targeting formula

  “One reason you're seeing this ad is that Peek & Cloppenburg wants to reach people interested in Shopping and fashion, based on activity such as liking Pages or clicking on ads.”

  “There may be other reasons why you're seeing this advert, including that Acer wants to reach people aged 18 to 45 who live or have recently been in Germany. This is information based on your Facebook profile and where you've connected to the Internet.”
Validation of incompleteness

- Ran a controlled ad targeting ourselves using a custom list and selecting millennials & expats
  
  - “One of the reasons why you're seeing this advert is because we think that you may be in the Millennials audience. This is based on what you do on Facebook..”
  
  - “There may be other reasons why you're seeing this advert, including that Acer wants to reach people aged 18 and above who live or have recently been in Germany. This is information based on your Facebook profile and where you've connected to the Internet.”
  
  - Unclear why millennials was chosen over expats!
Do explanations need to be complete?

- Should they specify **all features in ad targeting?**

  - **Arguments for:**
    - Avoids misleading / fake explanations:
      - Designed to gain consumer acceptance for a service
    - Builds trust and incentivizes cooperation

  - **Arguments against:**
    - Targeting formula may be a **business secret**
    - Overloads users with information
      - Need succinct explanations
Selecting features for explanations

- Example explanation:

  “One reason you're seeing this ad is that Peek & Cloppenburg wants to reach people interested in Shopping and fashion, based on activity such as liking Pages or clicking on ads.”

- Are the explained features the most important?
  - Is Shopping and fashion the most important of all the user’ features that Facebook and the advertised used to target the user?
  - Not clear what criterion Facebook is using!
Open challenges

- How to pick a few (K) features for explanations?
- How to determine the importance of a user feature?
  - Does it reveal privacy sensitive information?
- Is it a rare (or low prevalence) feature in population?
- Does it exert the most influence?
  - On sizing the target audience?
  - As measured by “Quantifying Input Influence” framework?
This talk: 4 Aspects of Ad Targeting

1. **Transparency**
   - Can I know what data about me is used to target ads?
   - **NO!** Only some (not all) data used is revealed

2. **Control**

3. **Fairness**

4. **Privacy**
This talk: 4 Aspects of Ad Targeting

1. **Transparency**
   - Can I know what data about me is used to target ads?
   - NO! Only some (not all) data used is revealed

2. **Control**
   - Do explanations offer consumers satisfactory control?

3. **Fairness**

4. **Privacy**
Consumer control over targeted ads

- Control **Inputs**: Determine *data used for targeting* ads
  - Help consumers protect their privacy

- Control **Outputs**: Determine *ads they receive*
  - Help consumers achieve specific decision outcomes
Consumer controls in Facebook

- Facebook claims that ad preferences page allows control over what ads users see

- In reality, it offers controls over inputs, not outputs!
Do explanations offer output control?

- Given the incompleteness of explained features:

  “One reason you're seeing this ad is that Peek & Cloppenburg wants to reach people interested in Shopping and fashion, based on activity such as liking Pages or clicking on ads.”

- Are the explained features necessary?
  - If the user asked Facebook to remove their data on shopping and fashion, would they stop receiving the ad?
  - No guarantees!
Do explanations offer output control?

- Given the **incompleteness** of explained features:

  “One reason you're seeing this ad is that *Peek & Cloppenburg* wants to reach people interested in *Shopping and fashion*, based on activity such as liking Pages or clicking on ads.”

- **Are the explained features sufficient?**
  - If the user asked Facebook to change their data on features other than shopping and fashion, would they continue to receive the ad?
  - No guarantees!
Open challenges

- Given the incompleteness of explained features

- How to construct necessary/sufficient explanations?

  - Technically challenging for arbitrary targeting formulae!

- Finding min. set of necessary or sufficient features
  - Often, maps to solving NP-Complete SAT problem
  - But, can leverage recent advances on SAT solvers!
This talk: 4 Aspects of Ad Targeting

1. **Transparency**
   - Can I know what data about me is used to target ads?
   - **NO!** Only some (not all) data used is revealed

2. **Control**
   - Do explanations offer consumers satisfactory control?
   - **NO!** Constructing satisfactory explanations is non-trivial

3. **Fairness**

4. **Privacy**
This talk: 4 Aspects of Ad Targeting

1. **Transparency**
   - Can I know what data about me is used to target ads?
   - **NO!** Only some (not all) data used is revealed

2. **Control**
   - Do explanations offer consumers satisfactory control?
   - **NO!** Constructing satisfactory explanations is non-trivial

3. **Fairness**
   - Can we detect & prevent discriminatory ad targeting?

4. **Privacy**
Can ad targeting be discriminatory?

Online Ads for High-Paying Jobs Are Targeting Men More Than ...
Adweek - 7 Jul 2015
"We found small instances where there was discrimination and gender-based discrimination in job ads," said ... The issue of bias and discrimination in ad targeting isn't new, ... The Carnegie Mellon researchers also were alarmed by ads that ... (Google's rules forbid serving ads based on health information.).

Carnegie Mellon Study Finds Gender Discrimination In Ads Shown ...
Marketing Land - 8 Jul 2015

When Algorithms Discriminate
Probing the Dark Side of Google's Ad-Targeting System
Highly Cited - MIT Technology Review - 6 Jul 2015
Google's Ad System Has Become Too Big to Control
In-Depth - Wired - 9 Jul 2015
Google's algorithm shows prestigious job ads to men, but not to ...
Highly Cited - Washington Post - 6 Jul 2015

View all
Can ad targeting be discriminatory?
Discrimination via correlated features

- FB’s early defense: *Ethnic affinity is not ethnicity*

- Used *voter records from NC* to check correlations
  - Voter records have race information
  - Created separate customer lists for different races
  - Checked correlations between their race & ethnic affinity

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Voter Records</th>
<th>Facebook Users</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>White</td>
<td>5,303,383</td>
<td>70.1%</td>
</tr>
<tr>
<td>Black</td>
<td>1,694,220</td>
<td>22.4%</td>
</tr>
<tr>
<td>Asian</td>
<td>79,250</td>
<td>1.0%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>163,236</td>
<td>2.2%</td>
</tr>
</tbody>
</table>
Does banning “ethnic affinity” help?

- What about pre-filtered custom lists:
  - using offline info like voter records?

- What of other correlated features?

<table>
<thead>
<tr>
<th>no</th>
<th>feature name</th>
<th>selectivity</th>
<th>Blacks percentage</th>
<th>rest percentage</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Demographic &gt; Ethnic Affinity &gt; African American (US)</td>
<td>17.0%</td>
<td>77.0%</td>
<td>10.9%</td>
<td>7.06</td>
</tr>
<tr>
<td>2</td>
<td>Demographic &gt; Politics (US) &gt; US Politics (Very Liberal)</td>
<td>11.8%</td>
<td>49.8%</td>
<td>7.7%</td>
<td>6.44</td>
</tr>
<tr>
<td>3</td>
<td>Interests &gt; Entertainment &gt; Music &gt; Gospel music</td>
<td>14.4%</td>
<td>48.3%</td>
<td>14.6%</td>
<td>3.32</td>
</tr>
<tr>
<td>4</td>
<td>Interests &gt; Shopping and fashion &gt; Beauty &gt; Hair products</td>
<td>12.2%</td>
<td>40.8%</td>
<td>12.9%</td>
<td>3.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>no</th>
<th>feature name</th>
<th>selectivity</th>
<th>Blacks percentage</th>
<th>rest percentage</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Demographic &gt; Politics (US) &gt; US Politics (Very Conservative)</td>
<td>14.4%</td>
<td>4.8%</td>
<td>26.5%</td>
<td>0.18</td>
</tr>
<tr>
<td>2</td>
<td>Demographic &gt; Politics (US) &gt; US Politics (Conservative)</td>
<td>16.6%</td>
<td>6.5%</td>
<td>29.7%</td>
<td>0.22</td>
</tr>
<tr>
<td>3</td>
<td>Interests &gt; Sports and outdoors &gt; Outdoor recreation &gt; Hiking</td>
<td>11.0%</td>
<td>8.0%</td>
<td>21.6%</td>
<td>0.37</td>
</tr>
<tr>
<td>4</td>
<td>Interests &gt; Sports and outdoors &gt; Outdoor recreation &gt; Camping</td>
<td>11.4%</td>
<td>11.5%</td>
<td>22.8%</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Open challenges

- How to detect discriminatory targeting in ads?
  - Particularly, with customer lists?

- How to avoid discriminatory targeting in ads?
  - Detecting & avoiding algo. discrimination is a hot topic
  - But, even here ads pose unique challenges

- Fair targeting might result in unfair ad impressions!
  - Targeting 100 men & 100 women might result in unequal impressions, when costs of their impressions are different!
This talk: 4 Aspects of Ad Targeting

1. **Transparency**
   - Can I know what data about me is used to target ads?
   - **NO!** Only some (not all) data used is revealed

2. **Control**
   - Do explanations offer consumers satisfactory control?
   - **NO!** Constructing satisfactory explanations is non-trivial

3. **Fairness**
   - Can we detect & prevent discriminatory targeting?
   - **Unclear** – need new measures & methods for fairness

4. **Privacy**
This talk: 4 Aspects of Ad Targeting

1. Transparency
   - Can I know what data about me is used to target ads?
   - NO! Only some (not all) data used is revealed

2. Control
   - Do explanations offer consumers satisfactory control?
   - NO! Constructing satisfactory explanations is non-trivial

3. Fairness
   - Can we detect & prevent discriminatory targeting?
   - Unclear – need new measures & methods for fairness

4. Privacy
   - Are there any privacy risks with targeted ads?
Potential audience reach estimate

- A feature of Facebook’s advertiser interface
Privacy risks from audience estimates

- Assume exact estimates of audience size

- Then, given a user’s PII
  - Like phone-num. or email-id. or name-address

- Any advertiser can check if the user is on Facebook

- And retrieve all the user’s info Facebook inferred
  - Including the financial info provided by data brokers!
Precision of audience reach estimates

- Reverse-engineered how the estimates work

- No estimates given when the audience reach < 20

- Estimates are rounded
  - Audience reach < 1000, rounded down to closest 10
  - Audience reach < 10000, rounded down to closest 100
  - Audience reach < 100000, rounded down to closest 1000
  - ..... 

- Estimates are transient, but stable over short-term
Privacy risks from audience estimates

- Given any customer list S with and a user U’s PII
  - Like phone-num. or email-id. or name-address

- Create a new customer list with S + U

- Is audience reach for S + U is more than S?
  - If it does, user U is on FB
    - One can similarly retrieve all the info FB has on the user
  - If not, either U is not on FB OR it’s a rounding error
The probability of rounding errors

- For a list $S$ with audience reach $> 20 \ & < 1000$
  - Rounding error probability is 0.9

- For $K$-lists with audience reach $> 20 \ & < 1000$
  - Chance of every try suffering a rounding error is $0.9^K$
  - Chance of at-least one try not being rounded is $1 - 0.9^K$
    - For $K = 100$, this is chance is 99.999974%

- So by creating **100 lists** with reach $> 20 \ & < 1000$
  - One can w.h.p. retrieve all data FB has on any user
How to create such customer lists

- Use **public voter records** in the US!
  - Randomly sample names/addresses from records
  - Till you get a customer list of size $> 20$ & $< 1000$!
  - Repeat the process 100 times!

- **Validation**: Used it to retrieve all data FB has on us!
  - User transparency tool?
  - But, could be used to **retrieve data on others** as well!
Open challenges

- Audience estimates are very useful for advertisers

- How to preserve the estimates without data leaks
  - Differential privacy? Other noisy estimates?
    - But, how can one enforce privacy budgets?

- Other privacy risks with custom lists?
  - Creepy ads?
    - E.g., targeting sex-offenders using public criminal records
  - More potent attacks?
Summary: Ad Targeting on Social Media

1. Transparency
   - Limited! Only some (not all) data used for is revealed

2. Control
   - Limited! Providing satisfactory explanations is non-trivial

3. Fairness
   - Unclear! Need new measures & methods for fairness

4. Privacy:
   - Non-existent! Avoiding data leaks via ads is non-trivial!

- Lots of open challenges!