Circles, Post and Privacy: From A Feedback and Awareness point of view
“The way that people think about privacy is changing. What people want isn’t complete privacy. It isn’t that they want secrecy... It’s that they want control over what they share and what they don’t.”

--- Mark Zuckerberg, Co-Founder of Facebook

“Privacy isn’t a technological binary that you turn off and on. Privacy is about having control of a situation. It’s about controlling what information flows... People still care about privacy because they care about control.”

--- Danah Boyd, Senior Researcher at Microsoft Research
Problem

- Post something to the wrong people.
- Users want to control their information, not hide.
- Privacy settings need granularity.
Community Detection

- Help users make sense of their online friends.
- Many different ways of community detection:
  - Graph-based (betweenness, bridge, cliques, modularity, etc.)
  - Feature-based (mainly profile information to compute feature distances)
  - Mixture
Community Detection, does it make sense?

- Users create circles differently when asked differently.

Examples:

I know Julian McAuley because of stanford, australians

I know Jure Leskovec because of advisor, stanford, professors

Overlapping groups are encouraged.
Community Detection, does it make sense?

- Users create circles differently when asked differently.

Friends are categorized into non-overlapping groups.
Community Detection, does it make sense?

- Users create circles differently when asked differently.

<table>
<thead>
<tr>
<th>#Alters</th>
<th>OverlappingFrequency</th>
<th>DuplicatePercentage</th>
<th>#Ground-Truth Circles</th>
</tr>
</thead>
<tbody>
<tr>
<td>277</td>
<td>48</td>
<td>17%</td>
<td>24</td>
</tr>
<tr>
<td>480</td>
<td>21</td>
<td>4%</td>
<td>9</td>
</tr>
<tr>
<td>218</td>
<td>349</td>
<td>160%</td>
<td>14</td>
</tr>
<tr>
<td>134</td>
<td>44</td>
<td>33%</td>
<td>7</td>
</tr>
<tr>
<td>168</td>
<td>317</td>
<td>189%</td>
<td>14</td>
</tr>
<tr>
<td>51</td>
<td>34</td>
<td>67%</td>
<td>13</td>
</tr>
<tr>
<td>763</td>
<td>14</td>
<td>2%</td>
<td>17</td>
</tr>
<tr>
<td>706</td>
<td>359</td>
<td>51%</td>
<td>46</td>
</tr>
<tr>
<td>97</td>
<td>95</td>
<td>98%</td>
<td>32</td>
</tr>
<tr>
<td>52</td>
<td>6</td>
<td>12%</td>
<td>17</td>
</tr>
</tbody>
</table>
A gap

Community Detection Algorithm

Circles in our mind..
A gap

Community Detection Algorithm

Circles in our mind..
A gap

Community Detection Algorithm

Circles in our mind..
A gap

Community Detection Algorithm

Circles in our mind..
It’s hard to tell.
To bridge the gap - An explorative F&A approach.

- Find a Community Detection Algorithm that is sensible.
- Then enable the user to explore and make decisions.
Newman’s Modularity-based Community Detection
An interactive visualization
An interactive visualization
Users make visibility decisions

<table>
<thead>
<tr>
<th>Types of Regretted Posts</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. To give utterance to my bad mood or negative opinion</td>
<td>6</td>
</tr>
<tr>
<td>b. To ask for advice or help</td>
<td>1</td>
</tr>
<tr>
<td>c. Uploaded photos depicting me in a way that I do not want to show to everyone</td>
<td>15</td>
</tr>
<tr>
<td>d. Language-specific posts (shown to the ones speaking the same language)</td>
<td>2</td>
</tr>
<tr>
<td>e. Religion or politics-related posts</td>
<td>5</td>
</tr>
<tr>
<td>f. To avoid contacting with a group of people for a particular reason</td>
<td>6</td>
</tr>
<tr>
<td>g. To only contact with a group of people for a particular reason</td>
<td>6</td>
</tr>
<tr>
<td>h. Inappropriate jokes</td>
<td>9</td>
</tr>
</tbody>
</table>
Users make visibility decisions
Entropy of visibilities to a post in a circle

\[ Entropy(p) = \sum_{c \in C} \frac{|c|}{N} \cdot \text{Entropy}(c, p) \]

\[ \text{Entropy}(c, p) = -\frac{N_{c,p}}{|c|} \cdot \log_2 \frac{N_{c,p}}{|c|} \]

\[ -\frac{|c| - N_{c,p}}{|c|} \cdot \log_2 \frac{|c| - N_{c,p}}{|c|} \]
How useful are these generated circles?

Group MOD (8 people, 24 posts)
Group FB (8 people, 24 posts)

In general, Group MOD outperforms FB, but the deviation is unfortunately large, larger sample size is needed for more conclusive results.
Thank You!