CS 6474/CS 4803
Social Computing: Sociological Foundations III

Munmun De Choudhury
munmund@gatech.edu
Week 3 | September 5, 2018
Data Characterization

- Comparison of public and critics sentiment around movies (Twitter vs. Rotten Tomatoes + IMDB)
- Identifying hateful comments on YouTube
- Characterizing and understanding the tone and sentiment of politician and celebrity Twitter posts
Community Studies

- Differences between reputation mechanisms of Quora and StackOverflow – what leads to better reputation?
- Understanding the structure and dynamics of pro-eating disorder communities
- Identifying linguistic markers of relationship abuse in Reddit communities
Prediction

- Predicting emotionality and loneliness of Flickr photos by analyzing visual features
- Prediction of workplace stress by analyzing corporate email content (Enron)
- Predicting factors that lead to persistent use or quitting of social media (Twitter)
Tools and Visualization

- Classification and exploration of political bias in blogs via a visualization
- Topic-aware personalized trends on Twitter via a web application
- Content summarization tool to provide positive social support to mental health help seekers on Reddit
Domain/Population-Specific Exploration

- Inferring a “health score” of the eating activities of college students by analyzing content of Yelp reviews
- “Social listening” – using social media sentiment to infer stock market trends
- Discovering the grievances, emotional expression, and topics shared in an online community of incarcerated individuals
The Strength of Weak Ties
Strength of weak ties: Mark Granovetter:

"It is the distant acquaintances who are actually to thank for crucial information leading to your new job, rather than your close friends!"

- Mark Granovetter (born October 20, 1943): an American sociologist and professor at Stanford University.
- 1969: submitted his paper to the American Sociological Review—rejected!
- According to Current Contents, by 1986, the Weak Ties paper had become a citation classic, being one of the most cited papers in sociology.
Summary

• The strength of a tie is defined in terms of four dimensions: amount of time, intimacy, intensity and reciprocal services
• Finding: even weak ties can be valuable
  • Weak ties can also act as bridges, and can help information diffuse faster
  • Generally weak ties have role in social cohesion and mobility, e.g., findings a job, geographic move, forming interest communities
• Strong ties are high maintenance; weak ties prevent small highly connected cliques from fragmenting
• Content of ties not considered; theory not validated with data!
Are the gradations “strong”, “weak” and “absent” sufficient for characterizing ties?
According to [Rethinking Friendships: Hidden Solidarities Today (Princeton, 2006) by Liz Spencer and Ray Pahl], there are eight different types of relationships:

- Associates: don't know each other well, and only share a common activity, such as a hobby or a sport.
- Useful contacts: share information and advice, typically related to our work or career.
- Fun friends: socialize together primarily for fun without a deep relationship to provide each other with emotional support.
- Favor friends: help each other out in a functional manner but not in an emotional manner.
- Helpmates: display characteristics of both favor friends and fun friends; socialize together for fun and also help each other out in a functional manner.
- Comforters: similar to helpmates but with a deeper level of emotional support.
- Confidants: disclose personal information to each other, enjoy each others company, but aren’t always in a position to offer practical help.
- Soulmates: display all of these elements and are the people we are closest to.

We have a much smaller number of strong ties than weak ties.
Tie strength: the 5-15-50-150-500 rule

According to [How Many Friends Does One Person Need?: Dunbar’s Number and Other Evolutionary Quirks, Robin Dunbar, Harvard University Press (November 1, 2010)]:

- Most peoples social networks have a common pattern, unchanged for thousands of years. There are clear boundaries based on the number of connections we have; it starts at five and goes up by a factor of three.
  - Inner circle: 5
  - sympathy group: 12-15
  - Semi-regular group: 50
  - stable social group: **150** (the Dunbar number)
  - friends of friends group (weak ties): 500

Robin Ian MacDonald Dunbar (born 28 June 1947): a British anthropologist and evolutionary psychologist and a specialist in primate behavior at University of Oxford.

Best known for his Dunbar’s number: a measurement of the “cognitive limit to the number of individuals with whom any one person can maintain stable relationships”.

Figure: Credit: (Adams, 2011)
Predicting Tie Strength With Social Media
Summary

• First study to quantitatively measure social tie strength.
• Based on Facebook data of 35 participants, they define 74 predictor variables of tie strength
• Seven different categories of the predictor variables:
• 85% predictive accuracy based on an OLS regression model; findings supplemented with interviews
• Findings:
  • Intimacy strongest feature; structural balance i.e., number of overlapping networks least
  • Strong ties provide emotional support, weak ties are great for accessing new kinds of information
How strong is your relationship with this person? barely know them ____________ we are very close
How would you feel asking this friend to loan you $100 or more? would never ask ____________ very comfortable
How helpful would this person be if you were looking for a job? no help at all ____________ very helpful
How upset would you be if this person unfriended you? not upset at all ____________ very upset
If you left Facebook for another social site, how important would it be to bring this friend along? would not matter ____________ must bring them

### Predictive Variables

#### Intensity Variables

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Distribution</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall words exchanged</td>
<td></td>
<td>9549</td>
</tr>
<tr>
<td>Participant-initiated wall posts</td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>Friend-initiated wall posts</td>
<td></td>
<td>47</td>
</tr>
<tr>
<td>Inbox messages exchanged</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>Inbox thread depth</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Participant’s status updates</td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>Friend’s status updates</td>
<td></td>
<td>200</td>
</tr>
<tr>
<td>Friend’s photo comments</td>
<td></td>
<td>1352</td>
</tr>
</tbody>
</table>

#### Emotional Support Variables

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Distribution</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall &amp; inbox positive emotion words</td>
<td></td>
<td>197</td>
</tr>
<tr>
<td>Wall &amp; inbox negative emotion words</td>
<td></td>
<td>51</td>
</tr>
</tbody>
</table>

### Duration Variable

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Distribution</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days since first communication</td>
<td></td>
<td>1328</td>
</tr>
</tbody>
</table>

### Reciprocal Services Variables

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Distribution</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Links exchanged by wall post</td>
<td></td>
<td>688</td>
</tr>
<tr>
<td>Applications in common</td>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

### Social Distance Variables

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Distribution</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of mutual friends</td>
<td></td>
<td>206</td>
</tr>
<tr>
<td>Groups in common</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Norm. TF-IDF of interests and about</td>
<td></td>
<td>73</td>
</tr>
</tbody>
</table>

### Structural Variables

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Distribution</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age difference (days)</td>
<td></td>
<td>5995</td>
</tr>
<tr>
<td>Number of occupations difference</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Educational difference (degrees)</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Overlapping words in religion</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Political difference (scale)</td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>
Why should we care about measuring tie strength on social media? How does it impact our studies of social media? How does it impact design?
Class Exercise I: What are the implications?

For individuals?
For communities?

Identify two scenarios each where strong ties are more useful and where weak ties are more useful
Are there situations where both strong and weak ties may be useful in an online context?

- Losing a job is associated with increases in stress, while talking with strong ties is generally associated with improvements in stress and social support.
- Weak ties do not provide these benefits.
- Bridging social capital comes from both strong and weak ties.
- Contrary to the “strength of weak ties” hypothesis, communication with strong ties is more predictive of finding employment within three months.
“Felt a little better with supporting comments about losing my job. My friends gave me a better outlook on the situation.”

“Been able to share my worries, get help on Cvs, job hunting and general chit chat about what’s happening to others.”

“Much better it is way easier to stay connected when looking for employment t see what others found. If someone finds a lead they cannot use they pass it on in case anyone else can use it. We are finding jobs”

“Yes, [I] am able to commiserate with other colleagues on losing our jobs (due to Bank failure) and getting prospects for new opportunities.”
Class Exercise II
How can you generalize Gilbert and Karahalios’ tie strength prediction model beyond Facebook?
Situate strong and weak ties in the context of unidirectional ties like on Twitter.
New types of “tie”-less systems

What does it mean to have “social interaction” in a tie-less system like Yik Yak?
Multiple platform use

Today we use more and more types of networks, and our contacts are spread across them. What does therefore a strong (or weak) tie on one platform mean for the other?
After all, in online context, interface design/algorithmic manipulation may hugely impact whether a social tie eventually grows to be a strong or a weak tie. What do you think is the impact?
Signed Ties and Structural Balance

- In many online and offline contexts, ties can be signed.
- Reasons could span from trust/mistrust to voting and positive/negative perceptions of feedback and interaction

(a) $A$, $B$, and $C$ are mutual friends: balanced.

(b) $A$ is friends with $B$ and $C$, but they don’t get along with each other: not balanced.

(c) $A$ and $B$ are friends with $C$ as a mutual enemy: balanced.

(d) $A$, $B$, and $C$ are mutual enemies: not balanced.