Self-Censorship on Facebook

Sauvik Das and Adam Kramer
sauvik@cmu.edu, akramer@fb.com
Self-censorship?

• Self-censorship is the act of preventing oneself from sharing a thought.

• We’ve all done it. (We’re doing it right now).
On Social Media?

• Social media adds an additional phase of filtering: *after* a thought has been expressed.

• Last-minute self-censorship
Why study it?

- Last-minute self-censorship can be both helpful and hurtful.

- Ripe opportunity to explore design implications and understand user behavior.
WHAT WE DO KNOW
What we know

Scarcely studied in its own right: it’s hard to measure what’s not there!
Boundary Regulation Strategy

• People present themselves differently to different social circles (Goffman ‘59)

• People have trouble maintaining consistency of presentation across social-contexts on social media (Fredric & Woodrow ’12, Wisniewski et al. ’12)
Tied to Audience

• People have an “imagined audience” when they share content on social media (Marwick & boyd ‘10)

• But are bad at estimating their audience (Bernstein et al. ‘13)

• When given the right tools, people selectively share and exclude content from different audiences (Kairam et al. ‘12)

• People said they would self-censor half as much content if given the right audience selection and exclusion tools (Sleeper et al. ‘13)
What we don’t know

• How often do people self-censor?
• What sorts of content gets self-censored?
• What factors are associated with being a more frequent self-censor?
Methodology
Measuring censorship

Registered that input occurred after 5 characters entered.

10 minute reset time, or after submission.

Compared “possible” posts with “shared” posts. Difference was considered **self-censored**.
Measuring censorship

- Censorship measured as a per-user count.

- 3.9 million randomly selected U.S./U.K. Facebook users.

- Logged for 17 days (July 6th-22nd, 2012).
Descriptive Stats

n=3,941,161

mean age: 30.9 years (s.d. = 14.1)

mean experience: 1386 days (s.d. = 401)

57% female
FINDINGS
How often do people self-censor?
Scale

71% of our sample self-censored at least once.

• 51% censored at least one post.
• 44% censored at least one comment.
Scale

• 33% of all potential posts censored.

• 13% of all potential comments censored.
What sort of content gets self-censored?
What gets censored?

<table>
<thead>
<tr>
<th>Content Type</th>
<th>Censorship Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>38.2%</td>
</tr>
<tr>
<td>Status Updates</td>
<td>34.5%</td>
</tr>
<tr>
<td>Events</td>
<td>25.3%</td>
</tr>
<tr>
<td>Friend’s Timeline</td>
<td>24.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Content Type</th>
<th>Censorship Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photos</td>
<td>14.7%</td>
</tr>
<tr>
<td>Group Msg</td>
<td>14.5%</td>
</tr>
<tr>
<td>Shares</td>
<td>12.7%</td>
</tr>
<tr>
<td>Status Update</td>
<td>12.2%</td>
</tr>
<tr>
<td>Wall Post</td>
<td>10.8%</td>
</tr>
</tbody>
</table>
Audience Uncertainty

Low censorship

Low

Events
Friend Timeline
Status Updates

High censorship

Groups

Introduction > Background > Methodology > Findings > Conclusion 21
Summary

• High audience uncertainty correlates with higher self-censorship.

• Broader topicality correlates with higher self-censorship.

• Groups can help us understand how these dimensions intermix.
What factors are associated with being a more frequent self-censorer?
Factors of interest…

- Social Graph Diversity
- Audience Selection Tools

Also controlled for…

- Activity on Site
- Age
- Experience with the Site
- Privacy Settings
- Gender
Hypotheses
People with more diverse social graphs will censor more.
### Social Graph Diversity Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Expected Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number friends of friends</td>
<td>+</td>
</tr>
<tr>
<td>Biconnected components</td>
<td>+</td>
</tr>
<tr>
<td>Friendship density</td>
<td>-</td>
</tr>
<tr>
<td>Friend age entropy</td>
<td>+</td>
</tr>
<tr>
<td>Friend political entropy</td>
<td>+</td>
</tr>
</tbody>
</table>
People who use audience selection tools will censor less.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Expected Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group member count</td>
<td>-</td>
</tr>
<tr>
<td>Friendlist created</td>
<td>-</td>
</tr>
<tr>
<td>Private messages sent</td>
<td>-</td>
</tr>
</tbody>
</table>
Modeling self-censorship

• Employed a Negative Binomial Regression.

• Negative Binomial was favored over Poisson because of the presence of overdispersion.

• Response was the count of censored count, offset with amount of created content.

• Coefficients estimated separately for posts and comments. Only posts reported in this presentation.
<table>
<thead>
<tr>
<th>Category</th>
<th>Feature</th>
<th>Coefficient</th>
<th>Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Graph Diversity</td>
<td>Average number friends of friends</td>
<td>1.32</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Biconnected components</td>
<td>1.12</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Friendship density</td>
<td>0.97</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Friend age entropy</td>
<td>0.96</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Friend political entropy</td>
<td>0.92</td>
<td>+</td>
</tr>
<tr>
<td>Audience Selection Tools</td>
<td>Group member count</td>
<td>1.29</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Buddylists created</td>
<td>1.13</td>
<td>-</td>
</tr>
</tbody>
</table>

All significant at $p = 0.01$
<table>
<thead>
<tr>
<th>Category</th>
<th>Feature</th>
<th>Coefficient</th>
<th>Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Graph Diversity</td>
<td>Average number friends of friends</td>
<td>1.32</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Biconnected components</td>
<td>1.12</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Friendship density</td>
<td>0.97</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Friend age entropy</td>
<td>0.96</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Friend political entropy</td>
<td>0.92</td>
<td>+</td>
</tr>
</tbody>
</table>

- **Diversity appears to have two components:**
  - People with larger 2\textsuperscript{nd} degree networks and more distinct social circles censor more
  - People whose friends are more diverse censor less.
• Audience selection tools had the opposite effect that we expected!

• People who were part of more groups and people who created more friend lists actually self-censor more posts.
Conclusion
Magnitude

• Self-censorship occurs frequently in social media, as expected.

• 71% censored at least once.

• Frequency varies by the nature of the content (post vs. comment) and its context (group post vs status update).
Not Just Audience

• People censor more as audience uncertainty increases.
• People censor more as the breadth of the topicality increases.
• Groups are weird, and possibly the key to understanding the interaction.
Diversity

• People with more diverse friends censor less.

• But, people with more disparate social contexts censor more.
Boundary Regulation

• Self-censorship does appear to be a boundary regulation strategy.
• Users who have more distinct social circles self-censor more.
• Even controlling for the use of audience selection tools!
Boundary Regulation

Present audience selection tools are insufficient or untrusted by users who need to balance many different social contexts.
Limitations

• Our metric is only a correlate.

• Groups are still wildcards.

• We don’t actually know what gets self-censored.
Questions?
Boundary Regulation

References


Audience References


Politics References


Gender References


• Walther, J.B. Selective self-presentation in computer-mediated communication: Hyperpersonal dimensions of technology, language, and cognition. Comp. in Hum. Behav. 23 (2007), 2538.
ZINB References


## Full List of Features

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Behavioral</th>
<th>Social Graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Messages sent</td>
<td>Number of friends</td>
</tr>
<tr>
<td>Age</td>
<td>Photos added</td>
<td>Connected components</td>
</tr>
<tr>
<td>Political affiliation</td>
<td>Friendships initiated</td>
<td>Biconnected components</td>
</tr>
<tr>
<td>Media privacy</td>
<td>Deleted posts</td>
<td>Average age of friends</td>
</tr>
<tr>
<td>Wall privacy</td>
<td>Deleted comments</td>
<td>Friend age entropy</td>
</tr>
<tr>
<td>Group member count</td>
<td>Buddylists created</td>
<td>Mostly (C/L/M)* Friends</td>
</tr>
<tr>
<td>Days since joining Facebook</td>
<td>Checkins</td>
<td>Percent male friends</td>
</tr>
<tr>
<td></td>
<td>Checkins deleted</td>
<td>Percent friends (C/L/M)</td>
</tr>
<tr>
<td></td>
<td>Created posts</td>
<td>Friend political entropy</td>
</tr>
<tr>
<td></td>
<td>Created comments</td>
<td>Density of social graph</td>
</tr>
</tbody>
</table>

* Conservative/Liberal/Moderate
Measuring Censorship

• **Gold standard:** honest users reporting instances of self-censorship

• Practical constraints:
  • **speed:** slower site speeds would present a confound.
  • **invisibility:** had to run behind the scenes--manipulating the UI would require extensive user testing.
  • **privacy:** ethical consideration that prevents us from logging content that users do not want to share.
Males will censor more than females.
People with more opposite sex friends will censor more.
Male censor substantially more posts than females.

Interestingly, this is even true as more males are part of their social graph.
Take-aways

• User-specific factors do seem to be associated with self-censorship.
• Males censor much more than females.
• Further research will be needed to discern why.
Modeling Self-censorship
Connected Component

Biconnected Components