Governance in Social Media:
A case study of the Wikipedia promotion process

Jure Leskovec (Stanford),
Daniel Huttenlocher and
Jon Kleinberg (Cornell)
Social media sites are driven by a collective activity of a large population

- In many cases this is a small group of core participants
- Such core participants:
  - Assign tasks
  - Enforce norms
  - Organize information
  - Answer questions

**Governance** [Bikhchandani et al., Burke-Kraut, Kriplean et al.]

- **Deliberation**: Group decision making
- **Enforcement**: Carrying out decisions
Group decision making and the Wikipedia promotion process

- Users are nominated to become admins:
  - Highly trusted users with special privileges

- The promotion process has well-defined structure:
  - Candidate submits a case for promotion
  - There is a period of discussion and deliberation
  - Community votes

Note:

- Any user can vote (not just admins)
- Voting is public and recorded as part of Wikipedia
Three important features:
- Deliberative process yielding a single decision
- Is publicly recorded
- Consequential for the community

Similarity to the off-line world:
- People evaluate other people

We study the perspective of the voters:
- Burke&Kraut examine candidate’s perspective
- How voters evaluate candidates?
- How single voter behaves across elections?
- How voting unfolds over time?
Voter’s evaluation of the candidate reflects different types of relative assessment

- Let voter V vote on candidate C
- We find that the vote of V heavily depends on relationship and relative merit of V and C:
  - Past interaction
  - Number of edits
  - Number of “barnstars”

Response function of a voter V:

- Prob. V votes positively given the current positive fraction of votes in the election
- Striking diversity of voter response functions
Votes are time stamped and signed by users

- **2,794 elections:**
  - Sept ‘04 – Jan ‘08
  - 44.6% success rate
    - Successful: 94.7% support
    - Failed: 31% support votes
- **114,040 votes (78% support):**
  - Each vote can get commented:
    - Support votes: 7% get discussed
    - Oppose votes: 82% get discussed
### Dataset: User characteristics

- **8,298 distinct users voted**
  - **7,499 voters**
  - **2,539 candidates** (some go for promotion multiple times)

### Votes by user type:

<table>
<thead>
<tr>
<th>User type</th>
<th>N</th>
<th>$f_v$</th>
<th>$p_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrators</td>
<td>1,235</td>
<td>44%</td>
<td>0.794</td>
</tr>
<tr>
<td>Unsuccessful candidates</td>
<td>1,304</td>
<td>12%</td>
<td>0.748</td>
</tr>
<tr>
<td>Other users (voters)</td>
<td>5,759</td>
<td>44%</td>
<td>0.783</td>
</tr>
</tbody>
</table>

N ... number of users  
$f_v$ ... fraction of votes  
$p_s$ ... prob. of support vote
How do properties of voter V and candidate C affect V’s vote?

Two natural (but competing) hypotheses:

- Prob. that C receives a positive vote depends primarily on the characteristics of C
  - i.e., there is some objective criteria for a user to become an admin

- Prob. that C receives a positive vote depends on relationship between characteristics of C and V
  - i.e., C compares merit of V to his own merit
Two ways to quantify merit $M$:

- **Total number of edits of a user:**
  - The more edits the user made the higher merit she has

- **Number of barnstars received by user:**
  - A way to reward Wikipedia contributors for hard work
  - Anyone can award a barnstar: “They are free to give and bring joy to the recipient”

Relative merit:

- How does the prob. of $V$ voting positively depend on the difference in merit of $C$ and $V$?
Voter V votes on candidate C

Prob. of support vote of V as a function of merit difference: $M_V - M_C$

Hypothesis: Monotonically decreases
Relative merit

- Prob. of support vote of $V$ as a function of merit difference: $M_V - M_C$

- Observations:
  - $V$ is especially negative when merit equals: $M_V = M_C$
  - Rebound for $M_V > M_C$
Direct V-C interaction

- Prob. of positive vote as a function of prior interaction of V and C

- Observation:
  - Prior interaction increases prob. of a positive vote
Thresholds and diversity of voters

- Aggregate response function:
  - How does prob. of voting positively depend on frac. of positive votes so far?

![Graphs showing different types of response functions: Linear, Diminishing returns, Threshold.](image)
Thresholds and diversity of voters

- **Aggregate response function:**
  - Baseline: If voter were to flip a coin then \( f(x) = x \)

- **Observation:**
  - Voters more inclined to express opinion when it goes against the prevailing opinion
  - Consistent with [Wu-Huberman]
Thresholds and diversity of voters

- **Personal response functions:**
  - How does prob. of voter V voting positively depend on frac. of positive votes so far?
  - Enough data that we can build models of individuals

11 users that took part in >400 elections

![Graph showing personal response functions](image-url)
Personal response functions:

- Average is close to baseline but individual variation in shape of response function is large

28 users that took part in >300 elections
Thresholds and diversity of voters

- Personal response functions:
  - Over time voters become more conservative
    - Response functions shift down and left

78 users that took part in >200 elections
Elections unfold over time:

- Sequence of pairs \((s(t), o(t))\)
  - \(s(t)\) ... number of support votes at \(t\)th vote
  - \(o(t)\) ... number of oppose votes at \(t\)th vote
Elections over time

- Very negative elections end early
- Failed elections are “top-heavy”
  - Start very positive and slowly get negative
  - Successful elections remain stable over time
Elections are dynamic:

- Votes arrive one after another
- Each voter sees the complete history so far

Does the order of first few votes determine the outcome?
Dynamics of elections

- Election outcome as a function of time of first negative vote:
  - 7 positive, 1 negative vote
  - 14 positive, 1 negative vote

- False hypotheses:
  - Candidate’s friends vote early
  - Herding behavior (excessive influence of first few decisions) [Banerjee, Bikhchandani et al.]
Social media sites are governed by (often implicit) deliberation

Wikipedia voting process has an explicit, public and recorded process of deliberation

Main characteristics:
- Importance of relative assessment
- Diversity of individuals’ response functions
- No evidence for herding

Connections:
- Networks with positive and negative edges:
  - Theories of structural balance [Haider ‘46] and status [CHI ’10, WWW ’10]
THANKS!
Data + Code:
http://snap.stanford.edu