Networks with Signed Edges
The idea of the reaction papers is:
- Familiarize yourselves more in depth with the class material
- Do reading beyond what was covered
- You should be thinking beyond what you read, and not just take other people's work for granted
- Think of the paper as a way to start thinking about the project

Read at 2 to 3 papers:
- Anything from course site, last year’s site, Easley-Kleinberg,…

Logistics:
- Due in 1 week: Oct 20 in class!
- Can be done in groups of 2-3 students
- How to submit:
  - Paper copy in a box **AND** upload to HW submission site
    - Use the homework cover sheet
- See [http://www.stanford.edu/class/cs224w/info.html](http://www.stanford.edu/class/cs224w/info.html) for more info and examples of old reaction papers
On 3-5 pages answer the following questions:

1 page: Summary
   - What is main technical content of the papers?
   - How do papers relate to the topics presented in the course?
   - What is the connection between the papers you are discussing?

1 page: Critique
   - What are strengths and weaknesses of the papers and how they be addressed?
   - What were the authors missing?
   - Was anything particularly unrealistic?

1 page: Brainstorming
   - What are promising further research questions in the direction of the papers?
   - How could they be pursued?
   - An idea of a better model for something? A better algorithm? A test of a model or algorithm on a dataset or simulated data?
Recap: Signed Networks

- Networks with **positive** and **negative** links
- Structure of **signed triangles**
  - **Structural balance:**
  - **Status theory:**
    - $A \rightarrow B$ :: B has **higher** status than A
    - $A \leftarrow B$ :: B has **lower** status than A
- **How to compare the two theories?**
  - Triads provide **context**
  - **Surprise:** Change in behavior of A/B when we know the context

\[
s_g(t) = \frac{k - \sum_{i=1}^{n} p_g(A_i)}{\sqrt{\sum_{i=1}^{n} p_g(A_i)(1-p_g(A_i))}}
\]

$\text{pg}(A_i)$… generative baseline of $A_i$
$p_r(B_i)$… receptive baseline of $B_i$
**Status: Two Examples**

- **Two basic examples:**

  ![Diagram](Image)

  - Gen. surprise of A: —  
  - Rec. surprise of B: —

  ![Diagram](Image)

  - Gen. surprise of A: —  
  - Rec. surprise of B: —
Joint Positive Endorsement

- X positively endorses A and B
- Now A links to B

A puzzle:
- In our data we observe:
  - Fraction of positive links deviates
    - Above generative baseline of A
    - Below receptive baseline of B

- Why?
A Story: Soccer Team

- Ask every node: How does skill of B compare to yours?
  - Build a signed directed network
- We haven’t asked A about B
- But we know that X thinks A and B are both better than him
- What can we infer about A’s answer?
A Story: Soccer Team

A’s viewpoint:

- Since B has positive evaluation, B is high status
- Thus, evaluation A gives is more likely to be positive than the baseline

How does A evaluate B?

A is evaluating someone who is better than avg.

→ A is more positive than average

Y A B

Y... average node
B’s viewpoint:

- Since A has positive evaluation, A is high status.
- Thus, evaluation B receives is less likely to be positive than the baseline.

How is B evaluated by A?

B is evaluated by someone better than average.
→ They will be more negative to B than average.
Determine node status:
- Assign X status 0
- Based on signs and directions of edges set status of A and B

Surprise is **status**-consistent, if:
- **Gen.** surprise is status-consistent if it has **same** sign as status of B
- **Rec.** surprise is status-consistent if it has the **opposite** sign from the status of A

Surprise is **balance**-consistent, if:
- If it completes a balanced triad
### Predictions:

<table>
<thead>
<tr>
<th>$t_i$</th>
<th>count</th>
<th>$P(\cdot)$</th>
<th>$S_g(t_i)$</th>
<th>$S_r(t_i)$</th>
<th>$B_g$</th>
<th>$B_r$</th>
<th>$S_g$</th>
<th>$S_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>178,051</td>
<td>0.97</td>
<td>95.9</td>
<td>197.8</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_2$</td>
<td>45,797</td>
<td>0.54</td>
<td>-151.3</td>
<td>-229.9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_3$</td>
<td>246,371</td>
<td>0.94</td>
<td>89.9</td>
<td>195.9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_4$</td>
<td>25,384</td>
<td>0.89</td>
<td>1.8</td>
<td>44.9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_5$</td>
<td>45,925</td>
<td>0.30</td>
<td>18.1</td>
<td>-333.7</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_6$</td>
<td>11,215</td>
<td>0.23</td>
<td>-15.5</td>
<td>-193.6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_7$</td>
<td>36,184</td>
<td>0.14</td>
<td>-53.1</td>
<td>-357.3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_8$</td>
<td>61,519</td>
<td>0.63</td>
<td>124.1</td>
<td>-225.6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_9$</td>
<td>338,238</td>
<td>0.82</td>
<td>207.0</td>
<td>-239.5</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{10}$</td>
<td>27,089</td>
<td>0.20</td>
<td>-110.7</td>
<td>-449.6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{11}$</td>
<td>35,093</td>
<td>0.53</td>
<td>-7.4</td>
<td>-260.1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{12}$</td>
<td>20,933</td>
<td>0.71</td>
<td>17.2</td>
<td>-113.4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{13}$</td>
<td>14,305</td>
<td>0.79</td>
<td>23.5</td>
<td>24.0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{14}$</td>
<td>30,235</td>
<td>0.69</td>
<td>-12.8</td>
<td>-53.6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{15}$</td>
<td>17,189</td>
<td>0.76</td>
<td>6.4</td>
<td>24.0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$t_{16}$</td>
<td>4,133</td>
<td>0.77</td>
<td>11.9</td>
<td>-2.6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Number of correct predictions: 8, 7, 14, 13
Both theories make predictions about the global structure of the network

- **Structural balance – Factions**
  - Find coalitions

- **Status theory – Global Status**
  - Flip direction and sign of minus edges
  - Assign each node a unique status so that edges point from low to high
From Local to Global Structure

- Fraction of edges of the network that satisfy Balance and Status?

- Observations:
  - No evidence for global balance beyond the random baselines
    - Real data is 80% consistent vs. 80% consistency under random baseline
  - Evidence for global status beyond the random baselines
    - Real data is 80% consistent, but 50% consistency under random baseline
Predicting Edge Signs

Edge sign prediction problem
- Given a network and signs on all but one edge, predict the missing sign

Machine Learning Formulation:
- Predict sign of edge \((u,v)\)
- Class label:
  - +1: positive edge
  - -1: negative edge
- Learning method:
  - Logistic regression

\[
P(+|x) = \frac{1}{1 + e^{-(b_0 + \sum_i^n b_i x_i)}}
\]

Dataset:
- Original: 80% +edges
- **Balanced**: 50% +edges

Evaluation:
- Accuracy

Features for learning:
- Next slide
For each edge \((u,v)\) create features:

- **Triad counts (16):**
  - Counts of signed triads edge \(u \rightarrow v\) takes part in
- **Node degree (7 features):**
  - Signed degree:
    - \(d^+_{\text{out}}(u), d^-_{\text{out}}(u),\) \(d^+_{\text{in}}(v), d^-_{\text{in}}(v)\)
  - Total degree:
    - \(d_{\text{out}}(u), d_{\text{in}}(v)\)
  - Embeddedness of edge \((u,v)\)
Edge Sign Prediction

- **Classification Accuracy:**
  - Epinions: 93.5%
  - Slashdot: 94.4%
  - Wikipedia: 81%

- Signs can be modeled from local network structure alone
  - Trust propagation model of [Guha et al. ‘04] has 14% error on Epinions

- Triad features perform less well for less embedded edges

- Wikipedia is harder to model:
  - Votes are publicly visible

![Charts showing classification accuracy for Epinions, Slashdot, and Wikipedia](chart.png)
## Balance and Status: Complete Model

<table>
<thead>
<tr>
<th>Feature</th>
<th>Bal</th>
<th>Stat</th>
<th>Epin</th>
<th>Slashd</th>
<th>Wikip</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>1</td>
<td>1</td>
<td>-0.2</td>
<td>0.02</td>
<td>-0.2</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>0</td>
<td>-0.5</td>
<td>-0.9</td>
<td>-0.4</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>0</td>
<td>-0.4</td>
<td>-1.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>-1</td>
<td>-0.7</td>
<td>-0.6</td>
<td>-0.8</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.05</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>1</td>
<td>-0.01</td>
<td>-0.1</td>
<td>-0.01</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>-1</td>
<td>-0.9</td>
<td>-1.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>0</td>
<td>0.04</td>
<td>-0.07</td>
<td>-0.03</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>0</td>
<td>0.08</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>-1</td>
<td>-1.3</td>
<td>-1.1</td>
<td>-0.4</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>1</td>
<td>-0.1</td>
<td>-0.2</td>
<td>-0.05</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>0</td>
<td>0.08</td>
<td>-0.02</td>
<td>-0.1</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>-1</td>
<td>-0.09</td>
<td>-0.09</td>
<td>-0.01</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>0</td>
<td>-0.05</td>
<td>-0.3</td>
<td>-0.02</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>-1</td>
<td>0</td>
<td>-0.04</td>
<td>-0.3</td>
<td>0.05</td>
</tr>
<tr>
<td>(\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet)</td>
<td>1</td>
<td>1</td>
<td>-0.02</td>
<td>0.2</td>
<td>-0.2</td>
</tr>
</tbody>
</table>
Do people use these very different linking systems by obeying the same principles?

- How generalizable are the results across the datasets?
  - Train on row “dataset”, predict on “column”

Nearly **perfect generalization** of the models even though networks come from very different applications
Important Points

- Signed networks provide insight into how social computing systems are used:
  - Status vs. Balance
  - Role of embeddedness and public display

- Sign of relationship can be reliably predicted from the local network context
  - ~90% accuracy sign of the edge
Important Points

- More evidence that networks are globally organized based on status

- People use signed edges consistently regardless of particular application
  - Near perfect generalization of models across datasets
How do People Evaluate Others?
People express positive and negative attitudes/opinions:

- Through actions:
  - Rating a product
  - Pressing “like” button

- Through text:
  Sentiment analysis
  [Pang-Lee ‘08]
  - Writing a comment, a review
People Express Opinions

- **About items:**
  - Movie and product reviews

- **About other users:**
  - Online communities

- **About items created by others:**
  - Q&A websites
Users Evaluating Others

Any user A can evaluate any user B:

- Positive (+) vs. negative (−) evaluation

Data:

- Users to users:
  - Epinions: Does A trust B’s product reviews?
  - Wikipedia: Does A support B to become Wiki admin?

- Users to items:
  - StackOverflow: Up/down vote (6M votes):
    Does A think B contributed a good answer?
How do properties of **evaluator A** and **target B** affect A’s vote?

Two natural (but competing) hypotheses:

1. Prob. that B receives a positive evaluation depends primarily on the characteristics of B
   - There is some objective criteria for a user to receive a positive evaluation
How do properties of evaluator A and target B affect A’s vote?

Two natural (but competing) hypotheses:

- (2) Prob. that B receives a positive evaluation depends on relationship between characteristics of A and B
  - **Similarity**: Prior interaction between A and B
  - **Status**: A compares status of B to her own status
Ways to quantify status (seniority, merit) of a user:

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

- Total number of answers of a user:
  - The more answers given by the user the higher status she has
How does the prob. of A evaluating positively depend on the status of A and status of B?

- Model it as a function of status $S_A$ of A and $S_B$ of B separately?
- Model as the status difference $S_A - S_B$?
- Model as the status ratio $S_A / S_B$?
### Status: Relative Assessment (1)

- **How does status of B affect A’s evaluation?**
  - Each curve is fixed status difference: \( \Delta = S_A - S_B \)

- **Observations:**
  - **Flat curves:** Prob. of positive evaluation doesn’t depend on B’s status
  - **Different levels:** Different values of \( \Delta \) result in different behavior

**Graph**

- **X-axis:** Target B status
- **Y-axis:** Fraction of positive evaluations (P(+)i)
- Different colored markers indicate varying status differences:
  - Red circle: Delta from -1000 to -500
  - Blue square: Delta from -300 to -100
  - Green diamond: Delta from 100 to 300
  - Purple triangle: Delta from 500 to 1000

**Status difference remains salient even as A and B acquire more status**
How does status of B affect A’s evaluation?

- Each curve is fixed status difference: $\Delta = S_A - S_B$

Observations:

- Below some threshold targets are judged based on their absolute status
  - And independently of evaluator’s status
**Effects of Similarity**

- How does prior interaction shape evaluations?
  - (1) Evaluators are more supportive of targets in their area
  - (2) More familiar evaluators know weaknesses and are more harsh

- **Observation:**
  - Prior interaction/similarity increases prob. of a positive evaluation

**Graph:**
- Fraction of positive evaluations (P+) vs. similarity
- Legend: English (red), French (blue), German (green), Spanish (purple)

**Observation:**
- Prior interaction/similarity boosts positive evaluations
Observation:
- Evaluation depends less on status when evaluator A is more informed

Consequence:
- Evaluators use status as proxy for quality in the absence of direct knowledge of B

Status is a proxy for quality when evaluator does not know the target
Observation:
- Evaluators with higher status than the target are more similar to the target

Selection bias:
- High-status evaluators are more similar to the target
Evaluator A evaluates target B

Prob. of positive evaluation of A as a function of status difference: $\Delta = S_A - S_B$

- Hypothesis: Monotonically decreases

\[
\begin{align*}
P(\text{positive eval}) & \quad (S_A < S_B) \\
0 & \quad (S_A = S_B) \\
-10 & \quad (S_A > S_B)
\end{align*}
\]
Puzzle: The Mercy Bounce

- Prob. of positive evaluation of B as a function of status difference: $\Delta = S_A - S_B$

- Observations:
  - A is especially negative when status equals: $S_A = S_B$
  - “Mercy bounce” for $S_A > S_B$

How to explain the mercy bounce?
How to explain low aggregate evaluations given by users to others of same status?

- Not due to users being tough on each other
  - Similarity increases the positivity of evaluations

Possible explanation:

- Most targets have low status (small \( \Delta > 0 \))
- Low-status targets are judged on abs. status
  - The rebound persists even for high-status targets
Social media sites are governed by (often implicit) user evaluations

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

Main characteristics:
- Importance of relative assessment: Status
- Importance of prior interaction: Similarity
- Diversity of individuals’ response functions

Application: Ballot-blind prediction
Predict Wikipedia election results without seeing the votes

- Observe identities of the first $k (= 5)$ people voting (but not how they voted)
- Want to predict the election outcome (promotion/no promotion)

Why is it hard?

- Don’t see the votes (just voters)
- Only see first 5 voters (10% of the election)
**Idea:** Split the status-similarity space \((s, \Delta)\) into 4 quadrants

**Model deviation in voter’s behavior when they evaluate a candidate from a particular quadrant:**

- \(d(s, \Delta)\) ... avg. deviation in fraction of positive votes
  - When voters evaluate a candidate \(C\) from a particular \((s, \Delta)\) quadrant, how does this change their behavior
Ballot-blind: the Model

- \( d(s, \Delta) \) ... signed deviation in the fraction of positive votes when \( E \) evaluates \( C \) of similarity \( s \) and status difference \( \Delta \)
  - \( P(E_i=1) \) ... prob. evaluator \( E \) votes + in election \( i \)

The models:

- Global **M1**: \( P(E_i = 1) = P_i + d(\Delta_i, s_i) \)
- Personal **M2**:

\[
P(E_i = 1) = \alpha \cdot P_i(\Delta_i, s_i) + (1 - \alpha) \cdot d(\Delta_i, s_i)
\]

where \( P_i \) is empirical frac. of + votes of \( E \)
Results: Wikipedia

- Predictive accuracy of baselines:
  - Guessing: 52%
  - If we know votes: 85%
  - Bag-of-features B1: 69%

- Model based on status and similarity:
  - Does not see votes
  - Sees only first 5 votes (10% of the lection)
  - Global model M1: 76%
  - Personal model M2: 75%
Important Points

- Online social systems are globally organized based on **status**

- **Similarity** plays important role

- Audience composition helps predict audience’s reaction

- What kinds of opinions do people find helpful?
What do people think about our recommendations and opinions?
- People find conforming opinions more helpful

![Graph showing the relationship between absolute deviation and helpfulness ratio.](image-url)
Positive reviews are more helpful