I rate you. You rate me. Should we do so publicly?

Chun-Yuen Teng, Debra Lauterbach, Lada A. Adamic
School of Information,
University of Michigan, Ann Arbor
Ratings are an integral part of Web 2.0, but are they honest?

• Many sites use recommender/reputation systems to help users identify reliable content and services

• How can one elicit honest ratings?
  – when users review items
  – when users rate other users’ reviews
  – when users rate other users

• We study a variety of sites that cover this space of rating types and various design choices
Amazon

• Amazon.com provides a platform which allows users to review products
• Users can decide to use a pen name or real name to review products
• 15 thousand article recent reviews from top 1500 reviewers (about one half using pen names)
Epinions

• Epinions.com allows users to share product reviews.

• Users can write reviews, rate other users’ reviews, and specify which users they “trust” or “distrust”
  – ~800K user-to-user ratings (trust or not)
  – ~100K users and 3 million articles
CouchSurfing

• CouchSurfing is a service for travelers looking to meet new people while finding a “couch” to sleep on.
• data: 600K users, 3 million edges

• Users can do the following for other users:
  – specify friendship level (e.g. acquaintance, friend, best friend)
  – specify how much they trust them (e.g. “somewhat”, ”highly”)
  – vouch for them
  – leave positive, neutral, or negative references
Structure of rating systems

Product review/rating

Rating of other users

Rating of product reviews
Research Questions

• How do design choices in online social networking & recommendation sites influence ratings?
• Are there other factors affecting ratings?
Design choices
Design choices: displaying ratings

• Show ratings publicly or keep private

• Amazon
  – Product reviews are public

• CouchSurfing
  – Public (friend, vouch, reference), private for trust

• Epinions
  – Public for rating of reviews
  – Public for trust links to other users
  – Private for distrust links to other users
Design choices: anonymity

- Do users have to identify themselves when giving ratings?
- Amazon
  - Choice of pen name or real name
- Epinions
  - Choice of username or staying anonymous when rating other users’ reviews
- CouchSurfing
  - All public ratings are identified: friendship, references, vouches
Design choices: reciprocity

• Is there any potential for reciprocity?
  • Amazon
    – Not really: products don’t rate you back
  • Epinions
    – Yes
  • CouchSurfing
    – Yes
Privacy and ratings
Privacy and ratings: scarcity of public, negative ratings when identified

• CouchSurfing
  – Users leave a positive reference for 87.7% of those they host and for 90.1% of those who host them
  – Neutral/missing references are confounded in data
  – The ratio of positive to negative references is 2500:1!
Privacy and ratings: negative ratings do not occur when they are private

• Epinions
  – Users express trust publicly and distrust privately
  – non-trivial fraction (14.7%) are “distrust” ratings.
Anonymity and ratings

amazon.com

Epinions.com
Anonymity and ratings in absence of reciprocity

<table>
<thead>
<tr>
<th>attribute</th>
<th>pen name</th>
<th>REAL NAME™</th>
<th>statistically significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>product rating # stars</td>
<td>4.19</td>
<td>4.21</td>
<td>no</td>
</tr>
<tr>
<td># reviews</td>
<td>498</td>
<td>551</td>
<td>yes</td>
</tr>
<tr>
<td>length of review (words)</td>
<td>364</td>
<td>377</td>
<td>yes</td>
</tr>
<tr>
<td># of fan voters</td>
<td>28.6</td>
<td>37.1</td>
<td>yes</td>
</tr>
</tbody>
</table>
Anonymity and ratings when there is potential for reciprocity

- Anonymous ratings are lower (3.84) on average than identified ratings (4.71)
- For the same user, anonymous ratings still average lower (4.01) than identified ones (4.76)
Is there evidence of reciprocity in ratings?
Reciprocity in Epinions

• We aggregate the user-to-article ratings into user-to-user ratings.
  – e.g.: if user A rates two of user B’s articles with average rating of 4, rating(A->B) = 4

• # of Ratings and rating scores show reciprocity
  – Rating from A to B is correlated with rating from B to A (rho = 0.475)
  – # of ratings from A to B and B to A also displays reciprocity (rho = 0.49)
Reciprocity in CouchSurfing

- *Public* friendship ratings are more highly correlated (\(\rho = 0.73\)) than *private* trust ratings (\(\rho = 0.39\))

(a) alignment of trust ratings between user pairs

(b) alignment of friendship ratings between user pairs

We omit trust rating of 2 (I don’t know the person)
Reciprocity in CouchSurfing

• Vouching also demonstrates reciprocity
  – If A vouched for B, 70% of the time B also vouched for A
  – Mean *private* trust score for reciprocated vouches was higher (4.47) than unreciprocated ones (4.19)
    ➔ lack of rating could signal lower trust
Are truthful ratings reliable?

• Even if one were able to elicit truthful ratings, would there still be biases?
• To answer this we used demographic information from CouchSurfing.com
Gender effects for trust & friendship

- Men rate both men and women about equally on trust and friendship.
- Women rate other women more highly on both trust and friendship.
Age

- Trust is very slightly higher the smaller the age difference between rater and ratee ($\rho = -0.06$)
- Trust depends on age of ratee – typical CouchSurfing demographic preferred?
Geography

• Closer friends tend to be geographically proximate
  – Friendship for one’s countrymen (4.19) is higher than foreigners (3.65)
  – Trust for one’s countrymen is higher than for foreigners (4.33 vs 4.16)
Geography

(a) average within-country ratings

(b) average ratings received by country residents

(c) average ratings given by country residents
Geography

• Countries with similar cultural background tend to be trusting of one another (e.g. Austria and Germany)

• Sharing a border does not always correspond to greater trust (e.g. Canadians did not rate US contacts more highly)
Conclusion

• Ratings should not be taken at face value
• Public, identified ratings tend to be positive when there is potential for reciprocity
• Demographics are tied to how users give ratings
Future work

• Survey users as to when and why they choose to rate anonymously
• Identify the criteria users use in rating others
• Develop trust prediction algorithms accounting for biases

more info

• http://netsi.org