

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

Chun-Yuen Teng, Debra Lauterbach, and Lada A. Adamic
*School of Information
University of Michigan
Ann Arbor, MI, USA*

Abstract

We find that ratings are not absolute, but rather depend on whether they are given anonymously or under one's own name and whether they are displayed publicly or held confidentially. The potential to reciprocate produces higher and more correlated ratings than when individuals are unable to see how others rated them. Ratings further depend on the gender and nationalities of the raters and ratees. All of these findings indicate that ratings should not be taken at face value without considering social nuances.

1 Introduction

Increasingly online websites promote the use of recommender systems to help customers to identify reliable content and services. These recommender systems rely heavily on ratings contributed by users in order to predict other users' preferences [3]. Possible reasons why users take the time to contribute ratings include improving the recommendations they receive themselves, enriching their profiles, self-expression, and helping and influencing others. Crucially, therefore, the quality of the recommendations depends on the quantity and accuracy of the ratings, which in turn depend on the design choices made when eliciting ratings. For example, a service may allow users to express themselves anonymously, or it may encourage them to build up their identified reputations [10].

With the proliferation of Web 2.0 services that enable users to interact with other users, often through an explicit social network, recommendation systems have acquired additional dimensions. Users can not only review products, but they can rate how much they trust other users' reviews, as is the case on Epinions.com. They can furthermore specify how much they trust the user in general, forming a web of trust directly from user-to-user ratings.

In this paper, we examine several possible factors that may contribute to how users give ratings. We conduct extensive analysis with three different datasets – collections of user-to-user ratings from the CouchSurfing social network, user-to-user and user-to-article ratings from Epinions, and user-to-product reviews from Amazon.

Our analysis shows that users give different ratings based on the context in which the ratings were given. That is, whether the ratings were given anonymously or using a real name, and whether they are shown publicly or kept private. We find other important factors in how users rate each other, including whether there is a potential for reciprocity, as well as the demographics of the rater and ratee.

The remainder of the paper is organized as follows. We review related work in Section 2 and describe our datasets in Section 3. Section 4 contrasts ratings that are anonymous vs. identified and private vs. public. Section 5 examines the role of reciprocity in public ratings, while Section 6 probes biases that might stem from rater attributes. We conclude in Section 7 by discussing our findings and potential opportunities for future work.

2 Related work

Our prior work [12] examining trust and reciprocity on CouchSurfing.com (for a description of CouchSurfing, see Section 3.3) focused on CouchSurfing's "vouch" mechanism by which users can publicly vouch for a friend's trustworthiness [1]. We found that whether a friendship connection is vouched or not is best predicted by the strength of the friendship, followed by the reference rating and how the friends first met. Bialski et al. [4] also studied trust in the CouchSurfing network, looking instead at the private trust ratings for each connection to determine which factors were related to higher trust between friends. They found that higher trust ratings are most correlated with the origin and context of the friendship, as well as how long the pair had been friends. While

previous work on CouchSurfing has looked at either the public ratings or private ratings separately, to our knowledge, no research has looked at both the public and private ratings in tandem to explore whether the public trust displayed by vouching matches the private trust ratings.

Several studies have aimed to predict trust ratings between users by applying prediction algorithms to user-to-user trust ratings. For example, Jøsang et al. [11] examined transitivity in trust networks, while Leskovec et al. [13] utilized balance theory to predict link valance in a number of empirical networks with signed edges. Guha et al. [9] predicted trust and distrust on the Epinions dataset with an algorithm that propagated trust along trusted edges, but halted at distrusted ones. Massa [14] also analyzed the Epinions dataset and showed that a small number of controversial users have non-negligible impact on trust networks.

Other papers have aimed to derive trust from information other than direct trust networks. Caverlee et al. [5] recommended the use of users' behavior and feedback to generate a trust score. Golbeck [8] showed that trust scores between users are correlated with differences between movie ratings and the similarity between users' profiles. Skopik et al. [17] proposed determining trust relationships by mining communication patterns between users. Robert et al. [16] formulated trust as being based on initial information and accumulated information between individuals.

Finally, anonymity in students' evaluation of teachers has been found to encourage negative feedback [15], while confidentiality has been shown to elicit a higher number and more technically substantive comments of scholarly articles than open peer review [2]. Our study extends this prior work on anonymity and confidentiality in ratings, by additionally considering reciprocity and demographics, while taking the examination of ratings into the online realm with several large online data sets. We describe the data sets in the following section.

3 Datasets

Our selection of three diverse datasets is meant to capture different contexts in which users are asked to provide ratings and express trust, as shown in Figure 1. On Amazon.com users rate products, and can do so using their real name or an anonymous pen name. Epinions.com likewise allows users to rate products anonymously or by identifying themselves, but rather than examining the reviews themselves, we focus on users' ratings of *others' reviews*. Unlike rating a product, rating someone else's review can be considered a dyadic interaction, and can potentially result in the rated user reciprocating positively or negatively. The third dataset we study, from CouchSurfing.com, has users directly rating other users

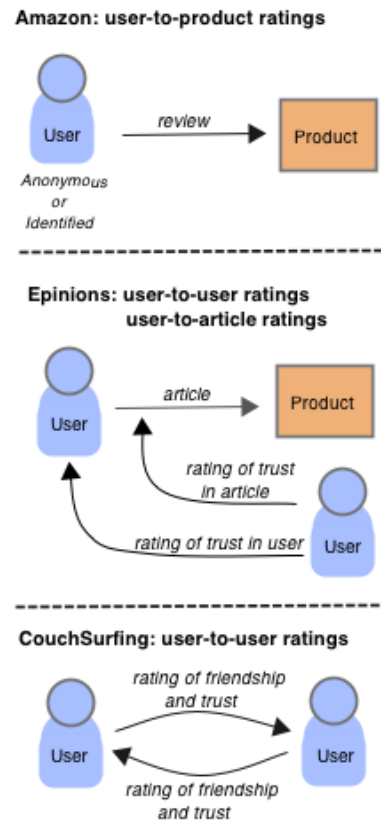


Figure 1: Types of rating activities for the three datasets used in this study.

on several dimensions related to friendship and trust. All public ratings are identified, but private trust ratings are not visible to other users. The datasets allow us to compare identified vs. anonymous and public vs. private ratings, and we describe them below.

3.1 Amazon.com

We downloaded the 10 most recent reviews posted by each of Amazon.com's top 1500 reviewers retrieved from <http://www.amazon.com/review/top-reviewers/> on Feb. 10, 2010. 740 of the top 1,500 reviewers did not use a real name, providing us with a roughly balanced sample of identified and potentially anonymous reviewers. Our data set includes for each review the username, a flag indicating whether the reviewer is using a pen name, a 1 to 5 star rating, a link to the product, product name, price, review title, text, and url as well as the date the review was made. We also recorded the mean rating and number of ratings for each product reviewed in order to evaluate whether a user's rating was low or high relative to others given for the

same product. We further record information about the users, including the number of helpful votes, percentage of helpful reviews, and number of other users who are fans of the reviewer.

3.2 Epinions

Epinions, a website that collects user-generated reviews for millions of products and services, has two types of ratings which we look at: user-to-user and user-to-article ratings. That is, users can express whether they trust a specific review that another user has written and they can also specify whether they trust another user’s reviews in general. The dataset we use, described in [14], contains 13,668,320 user-to-article ratings and 841,372 user-to-user ratings across 131,828 users and 3,119,947 articles.

3.3 CouchSurfing.com

CouchSurfing is a service that allows travelers across the globe to connect, not just for the purpose of finding a couch to sleep on, but to meet new people and learn about their culture. Because letting a stranger sleep on your couch is at best a rather risky business, CouchSurfing.com provides many features to identify trustworthy users. The most direct is the reference system, where users leave references for those they know, those who stayed with them and for those they stayed with. The vouch system [1] is another trust layer, which CouchSurfing wants its users to use with care: “The vouching system on CouchSurfing.com is a security measure. We take it VERY SERIOUSLY. Respecting the significance of vouching is essential to the integrity of the network. Once vouched for three times, you can vouch for any of your good friends. It is very important that you ONLY vouch for people that you have met in person and know well enough to believe that he or she is trustworthy”. CouchSurfing also collects private trust ratings from its users which are not displayed to other users in any form, but rather are used for internal purposes.

The dataset provided by CouchSurfing.com in anonymized form contains demographic information about the users, including, age, gender, city and country, the number of groups they belong to, how many times they have logged in to the site, and how many times their profile was viewed. In all, there were 648,099 users, with 3,115,548 directed edges, 3,011,487 of them having reciprocal ratings on both trust and friendship.

4 Anonymity and ratings

We first examine whether public and identified ratings tend to be more positive than ones that are expressed either anonymously or held private altogether.

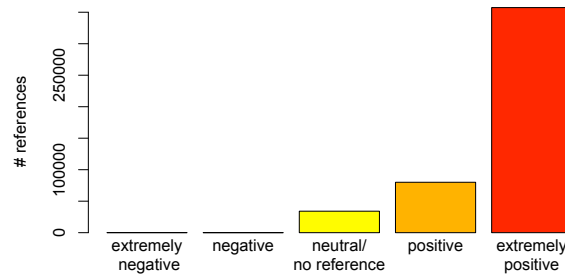


Figure 2: Balance of positive and negative references for hosting/surfing pairs on CouchSurfing.com before July 17, 2008, when “extremely positive” and “extremely negative” were eliminated as options.

4.1 Prevalence of positive public ratings

Overwhelmingly user-to-user public ratings are positive. On CouchSurfing.com the norm is to leave references for those one has had experience with. 73.1% of friendship ties have a positive reference, users leave a positive reference for 87.7% of those they host, and for 90.1% of those who host them. In contrast, negative references are extremely scarce, as shown in Figure 2. The ratio of positive to negative references that users leave for one another is 2,500:1. That is, while over 2,200,000 positive references were given, there were a mere 839 negative ones. Unfortunately, neutral references are conflated with the lack of a reference in our dataset. Since about 91.2% of neutral/non-reference stem from individuals who are friends, we believe that this category is composed mostly of missing data, and that neutral references are very rare as well. Therefore it appears that the primary utility of references is as a count of the number of positive experiences rather than a balance between positive and negative.

Epinions already recognizes the hesitancy users might feel in rating others negatively. When users rate whether they trust others, the “trust” ratings are public and form the “web of trust”. In comparison, the “distrust” ratings are private and are not shown on the site. Perhaps thanks to this insight on the part of the site designers, users do frequently give distrust ratings, and they comprise 14.7% of all user-to-user trust ratings. This makes one wonder whether there might be a higher number of negative references given on CouchSurfing, were the option of privately expressing dissatisfaction available.

4.2 Anonymous vs. identified ratings

In the previous section we noted the prevalence of positive ratings when they are identified and public. Next we compare ratings when users have the option of express-

ing their opinions publicly but anonymously. Both Epinions and Amazon allow their users to give ratings anonymously. Although CouchSurfing gives users the option of not verifying their identity, it is practically impossible to remain anonymous while identifying one’s friends. On Amazon.com reviews can either be written under a real name or a persistent pen-name which is unique to each account. Ratings of other users’ reviews are automatically anonymous, and are simply registered as a “helpful” vote. Therefore for the Amazon dataset we focus on the reviews themselves and whether they are penned under a real or a pen name.

The Epinions dataset does not include the product reviews themselves, but rather users’ ratings of others’ reviews. Users can choose whether to give these ratings anonymously. One might expect non-anonymous ratings in contexts such as these to be positive, where identifying oneself allows the individual being rated to reciprocate. For example, on Epinions, if user A rates one of user B’s reviews negatively, user B might find user A’s reviews and rate them negatively as well.

Of the two situations where one can choose to be anonymous, we expect anonymity to have a greater effect when a person rates another person’s review than when they are reviewing an item. This is because product ratings don’t carry the same likelihood of reciprocity. There exists the possibility that the manufacturer of a product could go in and rate a negative review as being unhelpful and furthermore find that user’s other reviews and rate them as unhelpful as well. However, they could do this whether or not the user is using their real name or their pen name, and it is not clear what utility the manufacturer would extract from this.

Indeed, we find no statistically significant difference in the average rating given in reviews penned under a pen name (4.19) compared ones to written with real name attribution (4.21). However, we do find several differences, albeit slight, between identified and anonymous reviews in terms of effort by reviewer. Top users using pen names tend to write on average fewer reviews (498) than those using their real names (551). Although there is no correlation between the valance of a rating and the length of the review ($\rho = 0.01, p = 0.459$), reviews written under a real name tend to be slightly longer (376.8 words as opposed to 364.1 words), suggesting that more effort goes into the reviewing task when one’s name is attached to it. However, this additional effort tends to go unrecognized, with anonymous and identified reviewers receiving helpful votes at an equal rate (85.7% vs. 86.0%).

One area where using a pen name does seem to be to one’s detriment is in acquiring fan voters. Fan voters are “people who consistently appreciate the author’s reviews”¹. One possible mechanism for this to occur is when a user is added as an “interesting individual”

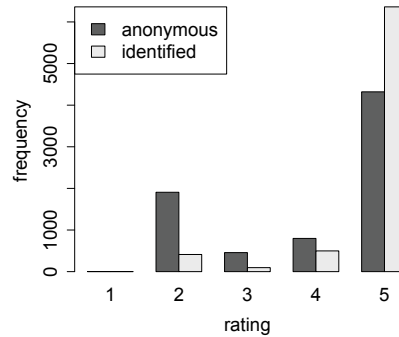


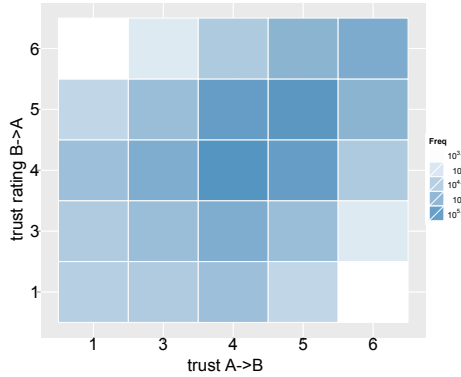
Figure 3: Comparison of ratings when the same Epinions user rates anonymously vs. when she identifies herself .

by others, meaning that those others will receive notifications when the user posts new reviews. In terms of number of fans, anonymous users on average have fewer ($p < 0.001$) fan voters (28.64) than real-name users do (37.12 fan voters). Perhaps other users are less likely to want to follow someone who they can’t identify as a person, regardless of the quality of reviews. Overall, these small differences suggest that in absence of reciprocity between rater and ratee, identified and de-identified ratings tend to be similar.

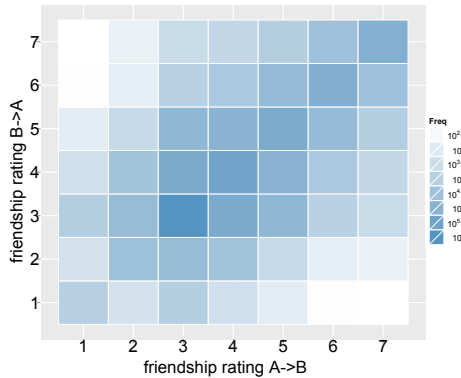
Interestingly, we find that few Epinions users choose to post their ratings of reviews anonymously. Only 115,058 (2.7%) of the user-to-article ratings are anonymous. However, while both identified and anonymous ratings tend to be positive, the mean of anonymous ratings, 3.84, is significantly lower, by nearly one star out of 5 (t-test $p < 0.001$), than that of the non-anonymous ones (4.71). Clearly users predominantly comment positively on others’ reviews, but when they have negative opinions, they tend to express them anonymously.

It might be possible that grumpier users, those who tend to give lower ratings, tend to also prefer to be anonymous, while more positive users tend to care less about anonymity. To check whether the *same* user prefers to be anonymous when rating negatively, we sampled one anonymous and one identified rating from all 7,484 users who had rated in both modes. We still found a significant ($p < 0.001$) difference between the mean of the anonymous (4.01) and identified (4.76) ratings for the same user, with the distribution shown in Figure 3. Any rating but the highest is more likely to be given anonymously, suggesting that individuals feel some discomfort in giving anything but highest praise to others publicly.

In short, we find that anonymous ratings do tend to be lower, but only in contexts where there is a reasonable chance of reciprocity between raters. We further examine reciprocity in the next section.



(a) alignment of trust ratings between user pairs



(b) alignment of friendship ratings between user pairs

Figure 4: (a) Users give each other trust ratings from 1 (I do not trust this person) to 6 (I would trust this person with my life). We omit “2” (I don’t know the person well enough to rate). (b) Friendship ratings vary from 1 (haven’t met) to 7 (best friend). 3 (CouchSurfing friend) refers to both the origin and context of the friendship.

5 Reciprocity in rating

Reciprocity is a natural tendency in both online and offline contexts. If someone does something nice or not so nice for us, we tend to pay them in kind.

5.1 CouchSurfing

Of the 4 types of ratings (vouching, references, friendship and trust) on CouchSurfing, only trust is not shown to the person being rated. If reciprocity plays a role, we expect higher correlation in the publicly displayed friendship level than in the privately held trust level. Figure 4 shows this to be the case, with friendship ratings between two users being more highly correlated ($\rho = 0.73$) than trust ratings ($\rho = 0.39$).

There are two possible explanations of the differences in correlation. One is that friendship is simply eas-

ier to evaluate than trust. However, our survey of 519 CouchSurfing users found no statistically significant difference in the difficulty level chosen by respondents in answer to the question “How difficult is it to select the [trust/friendship] level when adding a new friendship connection?”. The remaining explanation is that individuals adjust their friendship ratings because they can see how others rate them. For example, if A states that B is a best friend, while B states that A is a good friend, then A and B might feel uncomfortable about this discrepancy and adjust their ratings to align with what the other says. A may feel vulnerable for saying that B is a best friend if B’s feelings are not mutual. Or B might feel that she has insulted A by not saying that A is a best friend. The private nature of trust ratings may be preventing this type of reciprocity, as reflected in the lower correlation of trust.

In the CouchSurfing dataset, we can further compare the public web of trust in the form of vouches with the private system of trust ratings. In order to discern whether there is signal in the vouches about the individuals’ privately held trust beliefs, we took all pairs of users where both individuals are able to vouch for each other.

We find that vouches are indeed reflective of the underlying trust beliefs. Of the 94,546 instances where A vouched for B, and B was capable of vouching for A, the vouch is reciprocated 70% of the time. This demonstrates a high degree of public reciprocity. We do find that B is more likely to reciprocate if they hold a higher private trust rating of A. The mean trust score for reciprocated vouches was 4.47, compared to 4.19 for unreciprocated ones. Both fall between 4 (“I generally trust this person”) and 5 (“I highly trust this person”). This observation suggests that there is signal in missing ratings: at least some of the users that had an opportunity to reciprocate a vouch but did not may be communicating an implied distrust. One might be able to design reputation systems that could incorporate lack of reciprocity as implied distrust.

5.2 Epinions

To analyze reciprocity in the Epinions dataset, we aggregate the user-to-article ratings into user-to-user ratings. For example, if user A rates two of user B’s articles with an average rating of 4.5, we then just assign the user-to-user rating from A to B to be 4.5. We further filter out pairs of users who never rated each others’ articles. This leaves us with 74,747 user pairs.

Based on these mutual user-to-user ratings, we measure the degree of reciprocity. First we observe that the number of ratings from A to B and B to A displays moderate reciprocity ($\rho = 0.49, p < 0.001$). This could of course be due to users wanting to reciprocate, but also likely stems from shared interests by the users which

leads them to rate the same products and each others' articles on those products. We further find a moderate degree of reciprocity in ratings: if user A gives user B a higher rating score, user B also tends to give user A a higher rating score ($\rho = 0.475, p < 0.001$). Interestingly, and consistent with our results for CouchSurfing, anonymously given ratings have much lower correlation ($\rho = 0.135, p < 0.001$).

Finally, we find that the correlation in the semi-public user-to-user trust/distrust network is high. We recall that trust ratings are public, while distrust ratings are held private. Of the pairs of users who both rated one another, 97.1% publicly expressed mutual trust, 1.8% privately expressed mutual distrust, and 1.1% were mixed. The high degree of reciprocity might stem from the visibility of the public user-to-user trust network. Public trust ratings from A to B corresponded to B rating A publicly or privately 35.2% of the time. But when A distrusted B privately, this corresponded to B rating A only 6.0% of the time. While some reciprocity may stem from homophily, more likely users feel prompted to reciprocate when notified that someone holds their reviews in high regard.

6 CouchSurfing demographics and rating

So far we have examined how the options given to users within a rating system may influence their ratings. In this section we examine whether properties of the users themselves and those that they are rating correlate with the ratings given. The needed demographic data was available only in the context of CouchSurfing.

6.1 Gender

In the context of Couchsurfing, on average women are perceived to be more trustworthy (mean rating = 4.25) than men (mean rating = 4.12). What is even more intriguing is that men imparted trust ratings to both men and women that were on average equal (4.23), while women rated other women more highly (4.35) than men (4.05).

The same pattern repeats for friendship. Men rate the closeness of their friendship with other men on average about equally highly (3.88) as with women(3.82). However, women rate their female friends as closer (4.08) than they do their male friends (3.70). The above is summarized in Figure 5.

Furthermore, there is a greater degree of reciprocity in both friendship and trust ratings within genders, and specifically when both the rater and ratee are female, as detailed in Table 1. This could be due to women being more sensitive to social cues and context in determining appropriate behavior [7, 6]. Indeed, trusting someone

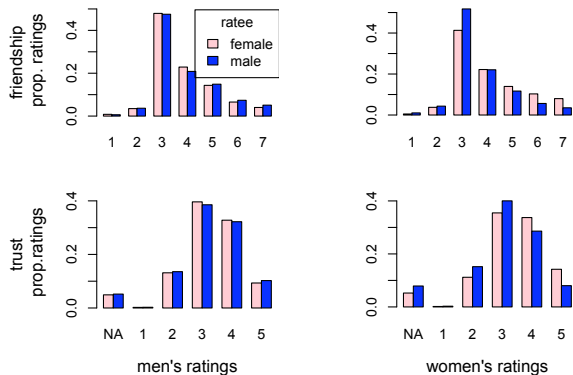


Figure 5: Friendship and trust ratings by gender of rater and ratee.

| | | friendship | | trust | | |
|---|--|------------|------|-------|------|------|
| | | M | F | M | F | |
| M | | 0.72 | 0.69 | M | 0.39 | 0.35 |
| F | | 0.69 | 0.82 | F | 0.35 | 0.45 |

Table 1: Correlations between how A rates B and B rates A, broken down by rating type and gender.

enough to let them sleep on your couch may very well produce different answers from men and women, while in other contexts the ratings may be indistinguishable or reversed. This gender asymmetry in trust would be an intriguing subject for further research.

6.2 Age

Age had very little bearing on how trustworthy a person was judged to be. Trust was slightly lower for older individuals ($\rho = -0.044, p < 0.001$) and those with greater difference in age from the rater ($\rho = -0.06, p < 0.001$). As shown in Figure 6, trust from others rises for users in their early 20's, then falls precipitously until plateauing in the 30s, after which point it starts rising. Unsurprisingly, the highest trust ratings are enjoyed by the typical CouchSurfing demographic, individuals in their 20s who are likely exploring the world and are both in need of a place to sleep cheaply and open to new experiences and social interactions.

6.3 Geography

Trust and friendship ratings also differ by geography. Closer friends naturally tend to be more geographically proximate; the average friendship degree for individuals in the same country is 4.19, as opposed to 3.65 for friends living in different countries. This also translates

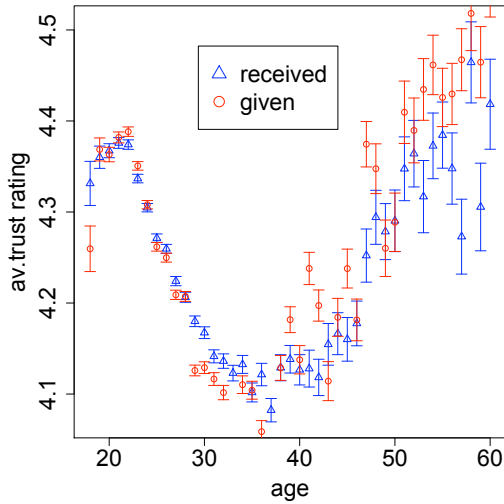


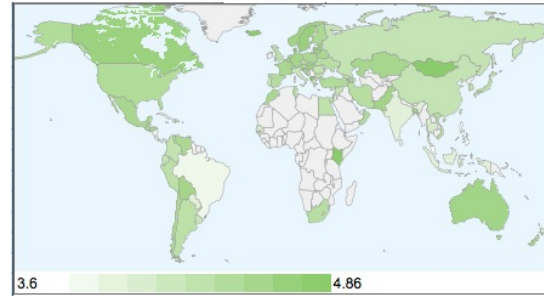
Figure 6: Average friendship ratings by age.

to a higher degree of trust for one’s countrymen than for foreigners (4.33 vs 4.16, $p < 0.001$).

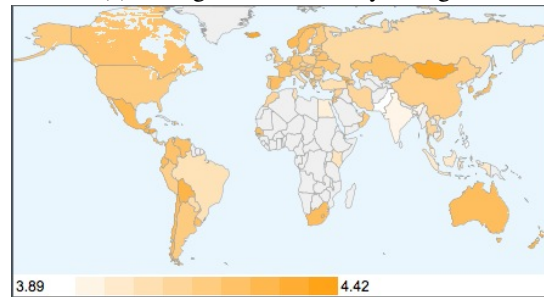
Next we aggregate trust at the country level. Figure 7 shows (a) how trustworthy residents of a country considered one another to be, (b) the trust residents of a country *receive* from foreigners, and (c) the trust residents of a country *give* to foreigners. The maps show that how residents perceive themselves is correlated with how they rate others and how others rate them. Possible explanations include different norms across countries in expressing trust, but also prevalence of crime and corruption. The one mild exception we noticed was that of Turkey, where residents were trusted slightly less by foreigners than they trusted those foreigners.

Pairwise trust ratings between countries also exhibit distinct patterns. For two countries, A and B, we take all the ratings from residents in A of residents in B, and compare it against ratings from residents in A to residents who are in neither A nor B. We then perform a t-test to identify pairs of countries where A gives B on average either unusually high or unusually low ratings. For robustness, we focus on countries that have both given and received 10,000 ratings.

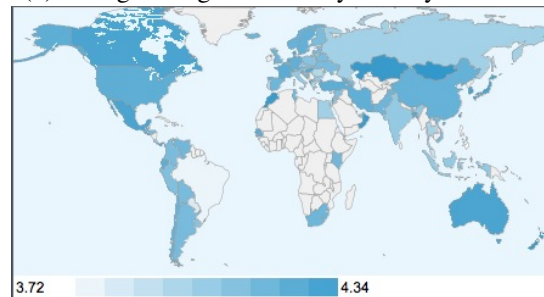
Some patterns that we saw on the aggregate scale are apparent in pairwise ratings as well. Brazil, India, Turkey, and Italy receive somewhat lower ratings from many other countries, and from one another. The United States is unpopular with Argentina, India is less trusting of China, while Americans give slightly lower trust ratings to Russians. Countries with similar cultural backgrounds tend to be more trusting of one another, e.g. Austria and Germany, Belgium and Spain, Brazil and



(a) average within-country ratings



(b) average ratings received by country residents



(c) average ratings given by country residents

Figure 7: Geography of trust: how users rate their countrymen and foreigners and how they are perceived by others.

Portugal, Canada and Australia, and Canada and France. The Scandinavian countries, Norway, Sweden, and Denmark, enjoyed higher ratings from one another. But sharing a border does not necessarily imply greater trust; Canadians did not rate the United States (4.24) more highly than they did other countries (4.26).

It is possible that the above demographic patterns in trust ratings are not a result of bias: that women should not trust men as much as women. That younger Couch-Surfers are more trustworthy than older individuals who choose to participate. That those who travel within their own country are more trustworthy than those who travel abroad. More likely, however, there are biases in how we humans evaluate each others’ trustworthiness, and these should be taken into account when trust ratings are gathered and utilized.

7 Conclusion

In this paper we studied factors in how users give ratings in different settings. Our three datasets, Amazon.com reviews, Epinions ratings, and CouchSurfing.com trust and friendship networks, represent a variety of design choices in how ratings are collected and shared. Less than truthful ratings can have different consequences, depending on the context. An inflated product rating might mislead a customer into buying a lower quality product, but an inflated trust rating on CouchSurfing.com, if not compensated for by other, more honest users, could lead one to host a rather unwelcome house guest.

Our findings indicate that ratings should not be taken at face value, but rather that one should examine the context in which they were given. Public, identified ratings tend to be disproportionately positive, but only when the ratee is another user who can reciprocate. Further evidence of reciprocity is in the alignment of public CouchSurfing friendship ratings, but far less alignment in the privately given trust ratings.

In future work, we are interested in surveying users as to when and why they choose to rate anonymously, and the criteria they use in rating others. We would like to design more robust ways of soliciting online ratings, and develop trust prediction algorithms that would account for the unavoidable biases we have identified. Both would inform design principles for trust and reputation systems in the context of online communities.

8 Acknowledgments

We would like to thank the NetSI group for helpful comments and suggestions. We are grateful to CouchSurfing.com for providing data used in this study. This work was funded by MURI award FA9550-08-1-0265 from the Air Force Office of Scientific Research.

9 Availability

The Epinions data set has been made freely available by Paolo Massa and may be downloaded from: www.trustlet.org/wiki/ExtendedEpinions_dataset

The Amazon reviews data set can be downloaded from www-personal.umich.edu/~ladamic/data/amazonratings/

References

- [1] Vouching on couchsurfing. <http://www.couchsurfing.org/vouch.html>.
- [2] Overview: Nature's peer review trial. doi:10.1038/nature05535, 2006.

- [3] ADOMAVICIUS, G., AND TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transaction on Knowledge and Data Engineering* 17, 6 (2005), 734–749.
- [4] BIALSKI, P., AND BATORSKI, D. Trust networks: Analyzing the structure and function of trust. In *International Network of Social Network Analysis SUNBELT conference* (2007).
- [5] CAVERLEE, J., LIU, L., AND WEBB, S. Socialtrust: tamper-resilient trust establishment in online communities. In *JCDL '08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries* (2008), pp. 104–114.
- [6] CROSON, R., AND GNEEZY, U. Gender differences in preferences. *Journal of Economic Literature* 47, 2 (2009), 448–474.
- [7] GILLIGAN, C. *In a different voice*. 1982.
- [8] GOLBECK, J. Trust and nuanced profile similarity in online social networks. *ACM Transactions on the Web* 3, 4 (2009), 1–33.
- [9] GUHA, R., KUMAR, R., RAGHAVAN, P., AND TOMKINS, A. Propagation of trust and distrust. In *Proceedings of the 13th international conference on World Wide Web* (2004), ACM, pp. 403–412.
- [10] HERLOCKER, J. L., KONSTAN, J. A., TERVEEN, L. G., AND RIEDL, J. T. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information and System Security* 22, 1 (2004), 5–53.
- [11] JØSANG, A., AND POPE, S. Semantic constraints for trust sensitivity. In *Proceedings of the 2nd Asia-Pacific conference on Conceptual modelling-Volume 43* (2005), Australian Computer Society, Inc., p. 68.
- [12] LAUTERBACH, D., TRUONG, H., SHAH, T., AND ADAMIC, L. Surfing a web of trust: Reputation and Reciprocity on CouchSurfing.com. In *Proceedings of the 2009 International Conference on Computational Science and Engineering-Volume 04* (2009), pp. 346–353.
- [13] LESKOVEC, J., HUTTENLOCHER, D., AND KLEINBERG, J. Predicting Positive and Negative Links in Online Social Networks. In *Proceedings of the International conference on World Wide Web* (2010), ACM.
- [14] MASSA, P., AND AVESANI, P. Trust Metrics on Controversial Users: Balancing Between Tyranny of the Majority. *International Journal on Semantic Web & Information Systems* 3, 1 (2007), 39–64.
- [15] MURPHY, K., CLEVELAND, J., SKATTEBO, A., AND KINNEY, T. Raters who pursue different goals give different ratings. *Journal of Applied Psychology* 89, 1 (2004), 158–164.
- [16] ROBERT, JR., L., DENIS, A., AND HUNG, Y.-T. Individual swift trust and knowledge-based trust in face-to-face and virtual team members. *Journal on Managing Information System* 26, 2 (2009), 241–279.
- [17] SKOPIK, F., TRUONG, H., AND DUSTDAR, S. Trust and reputation mining in professional virtual communities. In *Proceedings of the 9th International Conference on Web Engineering* (2009), p. 90.

Notes

¹<http://www.amazon.com/review/guidelines/fan-voters.html>