# TRUM 2013: The Role and Importance of Trust in User Modelling

Surya Nepal[1], Cecile Paris[1], Julita Vassileva[2], Jie Zhang[3]

[1]CSIRO ICT Centre, Australia
[2]University of Saskatchewan, Canada
[3]Nanyang Technological University, Singapore

surya.nepal@csiro.au, cecile.psris@csiro.au,
jiv@cs.usask.ca, zhangj@ntu.edu.

**Abstract.**. The 3rd international workshop on Trust, Reputation and User Modelling (TRUM 22013) was held with the International Conference on User Modeling Adaptation and Personalization (UMAP 2013). The purpose of the workshop is : (a) to bring researchers together from the communities of trust, reputation and user modeling, and online communities where trust plays an important role, (b) to provide a forum for cutting-age research possibly not yet well evaluated, and (c) to initiate and facilitate discussions on the new trends in trust, reputation and user modeling, and to move the trends forward. In this preface, we briefly introduce the workshop, present the summary of the papers presented in the workshop and ackwledged people who have helped for the success of the workshop.

**Keywords:** Trust, Reputation, User Modeliing

## 1 Introduction

The third Trust, Reputation and User Modeling (TRUM) workshop follows two successful previous workshops: TRUM'11 was held with UMAP 2011 at Girona, Spain and TRUM'12 - with UMAP 2012 at Montreal, Canada.

The workshops address an emerging area of overlap between user modeling and the area of trust and reputation modeling. This overlap has three aspects, illustrated in Fig. 1. First, decentralised and ubiquitous user modeling has sought inspiration from research in multi-agent systems over the last 10 years, resulting in a series of workshops on this topic at the User Modelling (UM) conference in 2005, 2007 and UMAP (User Modelling, Adaptation and Personalization) 2009. The current trend towards software applications using the cloud to store and process information that can be downloaded on social networks and mobile devices platforms brings new importance to the area of decentralised user modeling. Frameworks for dynamic and purpose based sharing of user model fragments among applicaions need to take into account the trust among these applications. The trust of one agent in another can be viewed as a simple user/agent model. Researchers in the area of trust and reputation mechanisms have studied for many years techniques allowing autonomous agents and peers to share, aggregate and make decisions based on these simple user models. User modeling researchers can gain useful insights from this area.

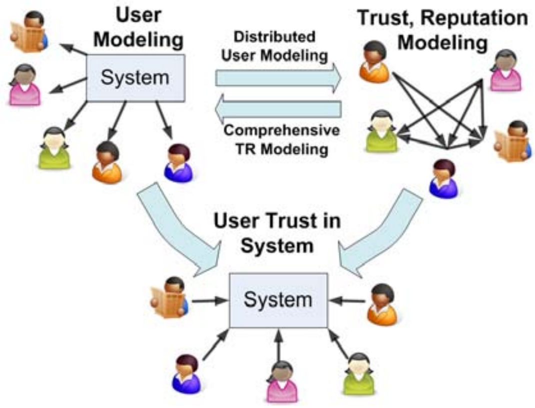



**Fig. 1.** Overalp of Trust, Reputation and User Modelling

Second, the area of trust and reputation modeling has experienced rapid growth in the past 7 years. Recently, two important trends have emerged in this area. One is the computational modeling of agents' cognition, such as subjectivity and disposition, to achieve more accurate trust and reputation modeling. The other is the modeling of agents' trust using a stereotype approach to deal with the problem of lack of experience. Both of these trends are closely related to studies in user modeling. The evidential success of these new trends inspires and encourages researchers in the trust community to make use of the rich literature in user modeling to develop more comprehensive trust and reputation modeling approaches.

A third important way in which research in user modeling overlaps with trust is the user’s trust in adaptive / personalised applications. In effect, it is a symmetrical area to that of user modeling: while user modeling suggests that the system models the user, here the user models the system. It relates to issues of user’s understanding of the application and of the privacy and integrity of the user model data, both of which are actively studied in the user modeling community. Facilitating the user’s understanding and trust in the system’s functioning and the way it manages the user’s data is very important, since it determines the user’s acceptance of the application’s recommendations or persuasion, the user’s satisfaction with the application’s functionality, and ultimately, its success.

While the papers presented in the first two TRUM workshops focused on formal models of users trust in systems / service providers, this workshop looks at trust in a more holistic way, that is manifested in online social networks. It involves three kinds of trust, as shown in Fig. 2 (trust triangle): (a) trust between members of the network, (b) trust between a member and the provided online service, and (c) the trust between a member and the service provider. This focus brings yet another intersection between trust research and user modeling, with respect to recommendation systems. Whereas recommendation systems typically rely on users' profiles or preferences, new types of recommendation algorithms are being designed based on trust behavior, thus further enhancing personalisation.

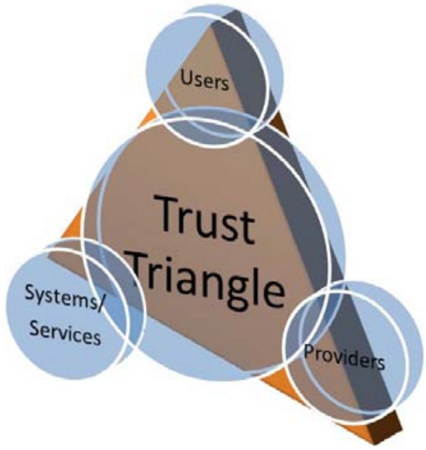



**Fig. 2.** Trust Triangle

To discuss the challenges related to this new holistic view and the potential solutions, the 3rd TRUM workshop was held with UMAP 2013 in Rome, Italy, with the following specific objectives:

- To bring researchers together from the communities of trust, reputation and user modeling, and online communities where trust plays an important role;
- To provide a forum for cutting-age research possibly not yet well evaluated;
- To initiate and facilitate discussions on the new trends in trust, reputation and user modeling, and to move the trends forward.

## 2 Organisation

The workshop was structured as a half a day event with a keynote speaker and four research paper presentations.

The keynote was given by Professor Alfred Kosba (University of California, Irvine, USA), on "Personalizing Privacy". It presented the results of recent studies on people's disclosure of personal data in smartphone and web shopping scenarios, showing a wide variety in individual privacy concerns across users. Further, providing adaptive, personalized privacy depending on the user individual privacy concerns. Ensuring a practical way to tailor the level of privacy according to the user's individual concerns and preferences is a novel and promising way of ensuring user trust in adaptive systems. This is particularly important for ensuring a better user experience and acceptance of recommender systems.

The research papers were as follows. The first paper, entitled "A User-Centric Study Of Reputation Metrics in Online Communities" by Hammer *et al*., discusses experimental work investigating whether users' trust in a reputation system is indeed positively affected by the system having more credible reputation values and more robustness against manipulation. The paper reports findings of an experiment carried out to investigate user perceptions of two reputation metrics, eBay and Neighbour-Trust Metric. The results could be of value to reputation metrics designers in making the system more user friendly. This is an important aspect of reputation systems as trusting reputation system is an essential to the successful and wide adaptation and deployment of reputation systems in ecommerce and online communities where users have to interact with unknown persons.

The second paper , entitled “Users’ motives shape trust in personalized applications: the importance of need satisfaction for perceived trustworthiness and risk” by Baer *et al*., looks at different user goals (in particular, "do-goals" and "be-goals"), and their respective effect on trust. The authors used two specific services (Facebook and Dropbox) to represent the different user goals and needs and conducted an experiment to examine whether the perceived trust and risks were also different.

The third paper explores the question of what constitutes trust in social networks and how people would characterise their conclusions of trust in these networks (e.g. according to which factors). It is entitled “Trust evaluation on Facebook using multiple contexts” and written by Švec and Samek. In the paper, the authors ask respondents (Facebook users) some questions in an effort to determine whether the authors' own proposal for trust modeling would coincide well with the views of the users. Graphs are presented which attempt to quantify the extent to which the authors' proposed model diverges from the users' opinions. Another interesting aspect of the paper is its exploration of literature that has likely not been discussed to a significant extent within artificial intelligence circles of trust modeling: theories from sociology. Its clarification of Marsh's original model is also insightful.

Finally, Bista *et al.* present a study of people's trusting behavior and expectation towards others within and out of a specific online community for welfare recipient in Australia. The paper is entitled “Know Your Members’ Trust”. The authors adapted a standard set of questions defined to capture trust attitude, trust experience and behavior, and trust expectation. Their results show that the members have overall positive expectation from the community, although they do not seem to have a trusting behavior towards strangers. There is a gap between members’ attitude and behavior about trust and their expectation from others. It is the authors’ hope that interactions within the community will help reduce this gap, leading to an increase in the social trust between members and towards governments.

## 3 Acknowledgements

We would like to thank workshop co-chairs, Shlomo Berkovsky and Pasquale Lops and the authors. Our gratitude also goes to the program committee members: Sanat Kumar Bista, Michael Fleming, Nathalie Colineau, Robin Cohen, Murat Sensoy, Thomas Tran, Wanita Sherchan, Julian Jang, Neil Yorke-Smith, Ebrahim Bagheri, Adam Wierzbicki.

# Personalizing Privacy

Alfred Kobsa

University of California, Irvine
U.S.A.

<last name> @ uci.edu

**Abstract.** Privacy concerns have been a nagging problem for the deployment of personalization over the past 15 years. Individuals' privacy preferences and behaviors vary widely though. Providers that collect personal information are therefore being asked or are even required to make their data collection and processing practices transparent, and to give users a say over how their personal data is treated. In practice though, people are overwhelmed by privacy choices and the rationality of their privacy decisions is limited.

We found regularities among people's disclosure of personal data in smartphone and web shopping scenarios, and identified subgroups who exhibit different disclosure behaviors with regard to different types of data. Some of these groups also have unique demographic or behavioral characteristics that are relatively easy to determine. We additionally found subgroups who react differently to different privacy "nudges". We discuss the implication of these findings on the dynamic adaptation of privacy to the presumed preferences of each individual user, an alternative to the above-described "transparency and user control" paradigm that imposes overly difficult and unwieldy privacy decisions on users.

# A User-Centric Study Of Reputation Metrics in Online Communities

Stephan Hammer, Rolf Kiefhaber, Matthias Redlin, Elisabeth Andre, and Theo Ungerer

Department of Computer Science, Augsburg University,
Universitaetsstr.6a, 86159 Augsburg, Germany
{hammer,andre}@hcm-lab.de
{kiefhaber,ungerer}@informatik.uni-augsburg.de
redlin@student.uni-augsburg.de

**Abstract.** With the growing importance of online markets and communities, users increasingly have to interact with unknown people. When choosing their interaction partners, they often lack direct experience and are forced to rely on ratings provided by others who are often unknown themselves. A number of reputation systems have been developed with the aim of improving the credibility of inferred reputation values. Most of these reputation systems proved their accuracy and robustness against manipulation in evaluations and therefore are believed to enhance the users trust in the system. However, what also matters is the users' experience with the reputation system. To investigate whether a reputation systems good functionality is sufficient to enhance the users' rating behavior and the users' trust in the provided reputation values and therefore also the entire system two substantially different reputation metrics were evaluated in an experimental game. The results obtained by this user-centric study are presented in this paper.

**Keywords:** Trust, Reputation Systems, User Study

## 1 Introduction

Today users interact in all kinds of online communities. They look for ratings for hotels, products or even experts, such as physicians. They trade in online marketplaces like eBay. They outsource tasks, such as the labeling of data, to online communities[1] and they arrange real-world interactions like carpooling[2] or small jobs like house cleaning or even babysitting[3]. In such communities users mostly have to interact with strangers. Therefore, it is crucial that they can trust in the benevolence and abilities of possible interaction partners. This reduces users' feeling of insecurity and risk [1] and increases their willingness to interact with unknown people [2].

[1] https://www.mturk.com/
[2] http://www.avego.com/
[3] https://www.taskrabbit.com/

The traditional approach of gathering information about someone's reputation entails asking only a small number of trusted people. This results in a small amount of information, but also in mostly credible information. In contrast, today's online approaches include a lot of information provided by a lot of mostly unknown people and thus the users are faced with uncertainty as to whether this information is reliable. Therefore, several reputation metrics, such as [3, 4] were presented to make inferred reputation more credible. All of these reputation metrics were evaluated on their accuracy, e.g., on the Epinions.com database [5], and proved their ability to overcome problems, such as manipulative ratings. Therefore, one could assume that the users trust more in these systems than in simpler ones. However, based on [6] it also matters how users think a reputation metric works and, more importantly, that users trust in the entire system's reliability, even if they do not know how it works. Therefore, two versions of an experimental game with substantially different reputation metrics, the Neighbor-Trust Metric (NTM) [7] and eBay's reputation metric[4], were designed. These versions were utilized in a user-centric study to investigate whether a reputation system's good functionality is sufficient to enhance users' rating behavior and users' trust in the reputation values provided and therefore the entire system. This paper presents the results gained from this study and possible steps to improve the users' experience with reputation systems.

The remainder of the paper is structured as follows: Section 2 gives a short evaluation of different reputation metrics from users' point of view. In Section 3 we introduce the most important aspects of the reputation metrics, eBay's reputation metric and the NTM, that were compared in the user study. The experimental game, the user study conducted with the game and a discussion of the results and experiences are presented in Section 4. Section 5 concludes the paper and presents future work.

## 2 Reputation Metrics

Because trust between interacting and cooperating subjects is a major issue in many fields of research several reputation metrics already exist. In general they are divided into global and local metrics. In this section, they are compared from a user's point of view.

Global reputation metrics, such as eBay's reputation metric, infer a unique global reputation for every user and do not take into account subjective perceptions of users. This is contrary to the diverse characters and opinions of all kinds of people that take part in online communities. If inferring the reputation of users with a lot of ratings this seems to be no problem, because the global reputation consists of many ratings provided by diverse users and therefore generally fits most of the users' opinion. Furthermore, users that received many ratings, in general, also received mostly positive ratings. However, for users that received only a small number of ratings it is difficult to infer, if the assessed reputation

[4] http://pages.ebay.com/help/feedback/scores-reputation.html

will match the actual experience. Since most of the users in online communities only received few ratings [6], this is a big issue.

In comparison to that, local metrics take into account that users' opinions on others' statements or trustworthiness can differ and are very subjective. To assess the trustworthiness of so far unknown users, TidalTrust [3] and Moletrust [5], for example, take into account that people feel more confident about information provided by known and trustworthy people than about information provided by unknown people. Therefore, they include only ratings provided by trustworthy users. That again is a problem, if we think about the reality in online communities in which users often have to interact with people that probably are unknown to the users' trusted people, too. In this case, a user assesses people's benevolence, competence or trustworthiness without any provided information.

Other metrics, such as the FIRE metric [11], consider the ratings provided by all former interaction partners of the target user. However, without a mechanism that verifies the accuracy of these trust statements, this approach is vulnerable to attacks and manipulations. Malicious participants or groups, for example, could offer false ratings to promote untrustworthy partners or blur the reputation of other users [10].

The Eigen-Trust metric [4] as well as the Neighbor-Trust Metric (NTM) [7] enhance this approach by the identification and isolation of manipulating participants. Thereby, both are able to infer the reputation for unknown participants based on the assessment of trusted as well as of unknown participants in a trustworthy way. However, the NTM extends the Eigen-Trust metric by separating the trust values for the direct interaction between users and for the reputation users provide about each other. The reason for this is, that a bad interaction partner nevertheless could be a good informant and vice versa. The details of this approach will be explained in Section 3.

## 3 The Evaluated Reputation Metrics

For the study, two substantially different reputation metrics were chosen to investigate whether a reputation system's good functionality is sufficient to enhance users' rating behavior and users' trust in the reputation values provided and therefore the entire system.

Since eBay's Marketplaces ended the first quarter 2013 with 116 million active users[5], eBay's reputation metric[6], despite the already mentioned drawbacks, is one of the best-known reputation metrics and seems to be accepted by the users. Furthermore, it is also one of the few metrics to be analyzed with regard to their influence on the users' behavior [8, 9]. The results of these studies, for example, showed that only half of all trades on eBay were rated and that the majority of provided ratings were positive. Although at first sight the last result could be interpreted as a success, a closer look at the data revealed two problems: because there was a high correlation between the ratings provided by buyers and sellers,

[5] http://investor.ebay.com/releasedetail.cfm?ReleaseID=757272
[6] http://pages.ebay.com/help/feedback/scores-reputation.html

Resnick and Zeckhauser supposed that the users (1) reciprocated and (2) feared retaliation [6]. To address these problems, sellers no longer are allowed to give negative or even neutral ratings, so as to alleviate buyers' fears of retaliation or unfair ratings. Instead, sellers can only leave comments on unfair ratings and can request a revision of the rating by the buyer[7]. This does not seem to be a trustworthy approach to handle possible manipulations of ratings. However, the users seem to accept the reputation system. Therefore, it was chosen to be one of the utilized reputation metrics in the study.

In comparison, our Neighbor-Trust Metric (NTM) [7] gathers the direct trust values $t_{ic}$ from all former interaction partners $i$ of a target user $c$, called "neighbors", and aggregates them by a weighted mean metric to assess an individual, local reputation value $r_{ac}$ for every user $a$:

$$r_{ac} = \frac{\sum_{i \in neighbors(c)} w_{ai} \cdot t_{ic}}{\sum_{i \in neighbors(c)} w_{ai}}$$

The weights $w_{ai}$ represent the trust of the user $a$ in the trust values the neighbors $i$ provide. The reason for the separation of the trust values for direct interactions between users and for trust ratings users provide to each other is, that a bad interaction partner could nevertheless be a good informant, and vice versa. The weights are adapted after every interaction. When a user $a$ had a direct experience with a user $c$ and provided a trust rating $t_{ac}$, this rating is compared to the trust rating $t_{bc}$ a user $b$ provided before the interaction. If $b$ gave information that corresponded with $a$'s own experience, then the future statements of $b$ will be weighted higher than before. Correspondingly, if the ratings differ, the weight will be lowered. Thus, the metric is not only able to learn about the trustworthiness of the interaction partner, but also to identify users that provide false or non-conformist ratings. Furthermore, by weighting down these users' ratings, inferred reputation values later will be more trustworthy and accurate. Therefore, by overcoming the vulnerability to manipulation the NTM should be more trustworthy for users than, for example, eBay's metric.

## 4 The User Study

### 4.1 Experimental Design

We investigated the influences of different reputation metrics on users' trust and rating behavior by comparing two versions of an experimental game that was inspired by other experimental games [12–14]. The two versions of the game differed only in the utilized reputation metrics. One version used eBay's metric and the other version used the NTM. We believed that two results could be possible: (1) The NTM's robustness against manipulation and the more credible reputation values (A) increase users' trust in the system and (B) cause more

[7] http://pages.ebay.com/help/feedback/feedback-disputes.html

honest ratings. (2) There is no difference in using eBay's metric or the NTM, for instance, because users do not recognize the different functionalities, since there are too short and too few interactions in the game and in online communities in general.

The experimental game was designed as a collaborative quiz (see Fig. 1). We assumed that collecting points and the chance to win prizes would be engaging and emotive.

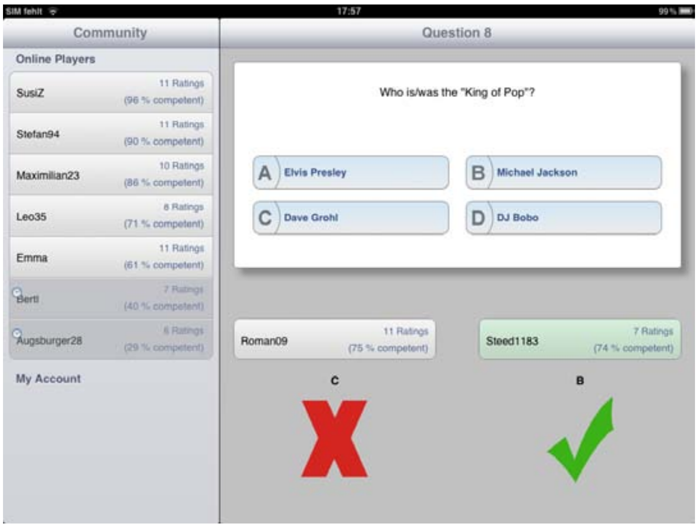



**Fig. 1.** Collaborative quiz

To enable a realistic comparison with cooperations in online communities, the following process sequence was designed:

1. A user has to choose an interaction partner. (In the study the interaction partners (teammates), were simulated by seven virtual players (VP) that were available from the beginning and had reputation values (RV) from 40% to 100%.)
2. The requested user has to confirm the cooperation. (In the study the decision of the VPs depended on their own reputation and the reputation of the requesting user. The user was rejected if her reputation was 20% lower than the VPs reputation (see Fig. 2 top)).
3. An interaction is successful if both users complete it successfully. (That is, both players have to answer a question correctly to get a point. The probability of a correct answer by a VP was *RV - 10%* for easy questions and *RV - 30%* for difficult questions. Therefore, players with a high reputation answered correctly more often than players with a low reputation.)

4. To increase all users' chances of gaining a higher benefit, interaction partners have to rate each other after each cooperation. This enables all users to distinguish between good and bad interaction partners.
5. When starting a new interaction, each user has to choose an interaction partner again. (Since the users were allowed to choose the same VP again, a VP that was chosen three times in a row entered an "idle" state, to prevent the participants from choosing the same VP throughout the entire study (see Fig. 3). This status lasted for three rounds.)

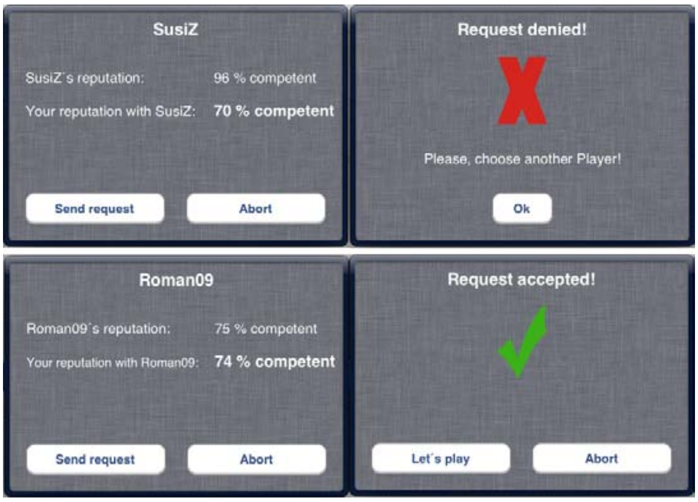



**Fig. 2.** Confirmation of user request depending on reputation. Top: rejection, bottom: acceptance

To investigate the users' reactions in different situations, based on [6], a variety of hardcoded behaviors for the VPs was implemented: (1) In general, the rating of the VPs corresponded to the user's answer. (2) If a user answered wrongly and rated the VP positively (independently of his answer), some VPs returned this favor and rated positively, too. (3) In a few cases some VPs rated a user negatively out of revenge if they received a negative rating.

### 4.2 Experimental Setting

Both versions of the quiz were played by half of the participants. In both versions the participants had to answer the same 10 easy and 10 difficult general knowledge questions. Based on the results of [6], such a small number of interactions corresponds to the actual conditions in online communities like eBay. Furthermore, it can be assumed that users that do not establish trust in a system during the first interactions will not use the system. The total number of received ratings in both versions of the quiz was shown for every player (see Fig. 3). Additionally, in the eBay-version a global unique reputation value, equal to that provided on eBay, was shown to support the user's selection of the next teammate (see Fig. 3 left). In comparison, in the NTM-version an individual local reputation value calculated by the NTM was shown (see Fig. 3 right).

To analyze the accuracy of the provided ratings and the users' selection of their teammates, the names of the chosen teammates, and the answers and ratings of the user and the current teammate for each question were logged. Moreover, interesting behavior was documented by hand. To analyze the participants' experiences with the respective quiz-version, they had to fill in questionnaires after they completed the quiz.

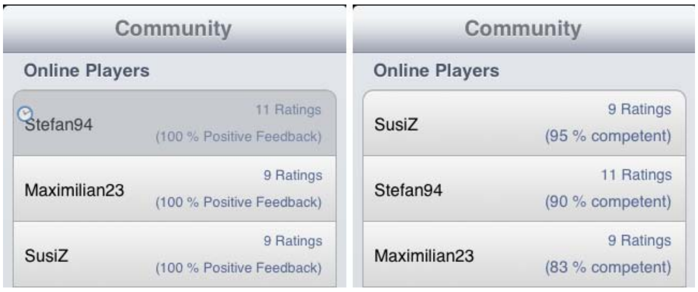



**Fig. 3.** Ranking of Virtual Players. left: eBay's reputation metric; right: Neighbor-Trust Metric

### 4.3 Conducting the Study

At first, the users had to fill in a questionnaire to provide general demographic information, and information about experiences with strangers and rating systems on the internet. Then, after a short introduction, the users had to play the quiz. The users were not informed about the functionality of the rating system. To increase the participants' ambition, they were told that they could win prizes depending on their results. After the quiz, the participants had to rate statements concerning their experiences with their teammates and the utilized rating system. All statements in the questionnaires had to be rated on a Likert scale from 1 ("not at all") to 5 ("definitely"). Ratings lower than 3 were interpreted as disagreement with a statement and ratings higher than 3 were interpreted as agreement.

### 4.4 Results and Experiences from the Experimental Game

Overall 16 women and 26 men aged between 22 and 56 (mean: 31.5) took part in the user study. The participants studied and worked in all kind of professions related (43%) and not related (57%) to computer science.

All participants already had interacted with unknown persons, e.g. on eBay, or had trusted in reviews on products or holiday destinations. Asked for their frequency of interactions with unknown people, the largest proportion of participants answered with "several times a year (29%)" or "several times a month (26%)". More than half of the participants reported on good (45%) or excellent (12%) experiences with unknown persons. All other participants rated their experiences as "neutral" and explained their ratings, for example, with mediocre information provided by others. Most of the participants agreed with the statement "Whenever you meet strangers, you have to be on guard until they have proven that they are trustworthy.". The average rating (M) was 3.62 (Standard Deviation (SD) = 0.90). This matches the fact that most of them also considered rating systems important (M=3.67; SD=0.89), because they allow an objective assessment of the trustworthiness of unknown people and decrease the chances of negative experiences. However, half of the participants were in doubt about the honesty of the provided feedbacks and some criticized possible manipulation and the lack of transparency of reputation systems. Nevertheless, several of the participants declared that reputation systems at least provide an indication of a user's trustworthiness.

A two-sided dependent t-test showed no significant differences for users' trust in the utilized reputation systems. Neither users' trust towards the provided ratings (NTM: M = 3.19; SD = 0.73; eBay: M = 2.95; SD = 0.72; p = 0.31) nor the perceived usefulness of the reputations systems (NTM: M = 3.80; SD = 0.66; eBay: M = 3.85; SD = 0.55; p=0.81) suggest that the NTM's ability to identify false ratings was recognized by the users. This was confirmed by an average rating of 1.71 (SD=0.76) when asked if they believed that the system is able to identify false ratings (eBay: M=2.05; SD=1.05). However, in both versions of the quiz almost all users stated that they based their selection of the teammates on the provided reputation values. But 67% of all participants also showed confidence and repeatedly selected players with whom they already had positive experiences, such as right answers or generous ratings and half of the users even based their choices mainly on positive experiences. This matches the results in [12] that direct and repeated interactions between users are the primary reason for increased trust.

The comparison of users' rating behavior in the two versions of the quiz showed small differences (see Fig. 4). But since the users did not recognize the NTM's ability to identify false ratings, these small differences seem not to be caused by the utilized reputation metrics. In both versions the participants rated honestly and rated positively if their teammates gave a correct answer (NTM: in 98% of the cases; eBays: 97%) and rated negatively if their teammates gave a wrong answer to easy questions (NTM: 73%; eBays: 65%) (see Fig. 4 (top)). However, there were many generous ratings (NTM: 50%; eBays: 55%) if the VPs gave wrong answers to difficult questions (see Fig. 4 (bottom)). Some of the users explained overly good ratings in general by the saying "To err is human". Furthermore, 33% of the participants admitted that they reciprocated, because of former positive experiences with the regarding teammate, such as right answers or prior generous ratings towards themselves. In this regard, 79% of all participants agreed that users in online communities can be convinced to rate positively if they received a positive rating in return (M=3.83; SD=0.65). 20% of all participants also explained that they provided overly good ratings because they feared retaliation. However, half of the participants negated that they would fear retaliation in general and the average score was 2.95 (SD=1.0). In summary, almost half of all wrong answers by the VPs were rated neutral or even positively.

Ratings (NTM-Users)

Easy Questions:

| | positive | neutral | negative |
|---|---|---|---|
| right | 146 | 2 | 0 |
| false | 4 | 13 | 45 |

Difficult Questions:

| | positive | neutral | negative |
|---|---|---|---|
| right | 100 | 2 | 0 |
| false | 9 | 45 | 54 |

Ratings (eBay-Users)

Easy Questions:

| | positive | neutral | negative |
|---|---|---|---|
| right | 151 | 2 | 0 |
| false | 4 | 16 | 37 |

Difficult Questions:

| | positive | neutral | negative |
|---|---|---|---|
| right | 106 | 4 | 2 |
| false | 10 | 43 | 45 |

**Fig. 4.** Ratings provided for right and false answers. NTM-Users (left) and eBays-Users (right); Easy Questions (top) and Difficult Questions (bottom)

## 5 Conclusion

This paper presented an experimental game by which two substantially different reputation metrics, Neighbor-Trust Metric (NTM) and eBay's reputation metric, were investigated from a user-centric perspective. The comparison of the metrics showed only small differences for users' rating behavior and no significant differences concerning users' trust in the reputation systems. This indicates that accuracy and robustness against manipulation are not the only criterions for good reputation systems. In addition to the general vulnerability of rating systems to manipulation, most of the participants in the study criticized the lack of transparency of rating systems. An improved transparency could therefore enhance users' experience of reputation systems. For reputation systems, such as the NTM, which assesses the credibility of a user's rating behavior, it could be a good idea to display this additional information. It could help to explain the inferred reputation and thus users' trust in the assessed reputation values could be increased. Furthermore, an additional criterion would be introduced that could support the choice of future interaction partners based on users' preference to interact with people that provide honest ratings. Finally, the amount of overly good ratings could be reduced, too.

**Acknowledgments** This research is sponsored by *OC-Trust* (FOR 1085) of the German research foundation (DFG).

## References

1. McKnight, D.H., Choudhury, V., Kacmar, C.: Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. Info. Sys. Research. 13(3), 334–359 (2002)
2. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An Integrative Model of Organizational Trust. The Academy of Management Review. 20(3), 709-734 (1995)
3. Golbeck, J.: Computing and Applying Trust in Web-Based Social Networks. Ph.D. Dissertation, University of Maryland at College Park, College Park (2005)
4. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The Eigentrust Algorithm for Reputation Management in P2P Networks. In: 12th international conference on World Wide Web, pp. 640–651. ACM, New York (2003)
5. Massa, P., Avesani, P.: Trust-aware Recommender Systems. In: 2007 ACM conference on Recommender systems, pp. 17–24. ACM, New York (2007)
6. Resnick, P., Zeckhauser, R.: Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. In: Baye, M.R. (ed.) The Economics of the Internet and E-Commerce 2002. vol. 11, pp. 127-157. Elsevier Science, Amsterdam (2002)
7. Kiefhaber, R., Hammer, S., Savs, B., Schmitt, J., Roth, M., Kluge, F., Andre, E., Ungerer, T.: The Neighbor-Trust Metric to Measure Reputation in Organic Computing Systems. In: 5th IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops, pp. 41–46. IEEE Press, New York (2011)
8. Resnick, P., Zeckhauser, R., Swanson, J., Lockwood, K.: The Value of Reputation on eBay: A Controlled Experiment. Experimental Economics. 9(2), 79-101 (2006)
9. Canals-Cerd, J.: The Value of a Good Reputation Online: An Application To Art Auctions. Journal of Cultural Economics. 36(1), 67–85 (2012)
10. Dellarocas, C.: Immunizing Online Reputation Reporting Systems Against Unfair Ratings And Discriminatory Behavior. In: 2nd ACM conference on Electronic commerce, pp. 150–157. ACM, New York (2000)
11. Huynh, T.D., Jennings, N.R., Shadbolt, N.R.: An Integrated Trust and Reputation Model for Open Multi-Agent Systems. In: Autonomous Agents and Multi-Agent Systems. 13(2), 119–154 (2006)
12. Bolton, G.E., Katok, E., Ockenfels, A.: How Effective Are Electronic Reputation Mechanisms? An Experimental Investigation. Management Science. 50(11), 1587–1602 (2004)
13. Bohnet, I., Harmgart, H., Huck, S., Tyran, J.-R.: Learning Trust. Journal of the European Economic Association. 3(2-3), 322–329 (2005)
14. Keser, C.: Experimental Games for the Design of Reputation Management Systems. IBM Syst. J. 42(3), 498–506 (2003)

# Users' motives shape trust in personalized applications: the importance of be-goals for perceived trustworthiness and risk

Nina Bär, Jessica Göthel and Josef Krems

Chemnitz University of Technology, Chemnitz, Germany
nina.baer@psychologie.tu-chemnitz.de

**Abstract.** The achievement of instrumental goals (do-goals) in automated systems is essential in forming users' trust. However, the use of personalized applications is additionally linked to non-instrumental goals (be-goals). Be-goals include the satisfaction of needs like stimulation, relatedness or competence that makes the use of personalized applications so popular. In an experimental study (N = 34) we investigated how different levels of be-goal achievement affect users' trust in two applications, a social network and a cloud service. Results reveal that greater be-goal achievement is related to lower users' trust, a lower perception of trustworthiness and higher risk. This finding suggests that users associate a higher vulnerability with online situations which are closely connected to the self. However, the be-goals competence and security appear to be positively influencing users' trust. From these first findings we argue that for enhancing trust in personalized applications both do-goals and be-goals should be considered.

**Keywords:** users' trust, be-goals, trustworthiness, risk, personalized applications

## 1 Be-goals in personalized applications

The user acceptance of personalized applications depends on a multitude of factors with users' trust in the system being one of them. The understanding of the application's purpose and trust in its proper functioning are important elements for high user satisfaction and frequent use. Concepts like the dynamic model of trust and reliance on automation [1] or models of antecedents of online trust in e-commerce [for an overview, see 2] highlight this relationship. Accordingly, trust is enhanced when the application supports the users in achieving instrumental goals (do-goals). But the claim to facilitate the users' understanding of the system's purpose in order to enhance trust neglects the role of non-instrumental user motives (be-goals). Be-goals are closely connected to the individual experience during an interaction [3] and essential for the decision to use an application. They explain certain fundamental qualities of experience humans strive for, e.g. the feeling of being the cause of one's own actions (autonomy) or the impression of being capable and effective (competence) [4]. For experiences with technology the most salient be-goals can be categorized as stimulation, relatedness, competence and popularity [5]. Do-goals, in contrast, are concrete outcomes of actions. They are task-driven and closely related to the technology used to achieve them [3]. In general, all personalized applications mediate goal-directed actions, i.e. they fulfil do-goals. However, only a few applications specifically address be-goals beyond the level of do-goals. Applications like social networks and cloud services are much more than means to an end. Beyond do-goal achievement they allow additionally non-instrumental use. People use social networks and cloud services in the first place for be-goals like connecting with their friends, sharing personal information or presenting oneself. Therefore they fulfil be-goals to a higher extent than other applications.

## 2 Trust and the role of be-goals

In trust research the fulfilment of be-goals does not obtain sufficient attention when compared to do-goals. HCI focuses on providing help for the users to achieve their do-goals as the concrete outcomes of actions. Consequently, uncertainty in task-driven usage situations is reduced and trust emerges [1]. Users' trust in applications is defined as an attitude of confident expectation that one's vulnerabilities will not be exploited in a situation of risk [6]. In this definition the aspects of vulnerability and risk are crucial. Being vulnerable includes an exposure of the user [6]. Risk as the appraised likelihood of a negative outcome [7] comprises both the perceived probability of negative consequences resulting from the usage of a product or a service and the significance of these consequences [8]. Vulnerability and risk perception on do-goal level could evolve due to a lack of knowledge of the application's functioning, e.g. the user fails in achieving a concrete task like adding a person in a social network or a cloud service. On be-goal level the user becomes vulnerable when the meaning of an action within the application is of high personal relevance, e.g. a failure in adding a person who is important for the user. So the mere failure in achieving the do-goal "add a person" does not provide any significance to the goal in case it is a person the user hardly knows and does not care about. Only the be-goal "feel connected to an important person" provides the significance for the user and makes the consequences of not achieving this goal personally meaningful [3]. In other personalized applications which focus on do-goals the negative consequences of not achieving the goal (e.g. access personalized timetables) the personal meaning will not evolve to the same extent like in applications which focus on be-goals. Therefore, we assume that the users' exposure (vulnerability) and potential negative consequences (risk) of the use of personalized applications are more significant for the users when they feel an intense connection to the application, i.e. when the application helps to achieve a high amount of be-goals.

By focusing on be-goals in personalized applications we see a chance to narrow the understanding of trust. Components of users' trust in applications are the perceived trustworthiness, the perceived risk, general system trust and the users' propensity to trust. The first component, perceived trustworthiness, includes characteristics of ability which are indicated e.g. by the usability of an application. Ability is closely related to do-goals. The more the application enables the user to achieve do-goals, the more trustworthiness will be perceived. Additionally, the adherence to principles the user finds acceptable (integrity) and the benevolence of the application's provider or developer belong to perceived trustworthiness [6]. For instance, applications which offer accounts free of charge signal that they are not intended to serve profit motives in the first place and are therefore considered to be benevolent. Both integrity and benevolence relate to intentions and motives of the application's provider concerning the users' needs. Hence, they are considered to match be-goals. The second component of trust is perceived risk. It is defined as the appraised likelihood of a negative outcome [7] connected to the use of the application. As shown above, negative outcomes are of more relevance when the use of personalized applications contains personally meaningful be-goals. As a third component of users' trust in applications system trust refers to the trust in the underlying Internet technology [6] and contributes to the basic level of trust. Finally, propensity to trust has a direct impact on the use of applications and the formation of users' trust. The combination of the four components forms users' trust.

We hypothesize that applications with higher be-goal achievement are related to lower perceived trustworthiness and higher perceived risk. Users' trust resulting from perceived trustworthiness, perceived risk and the dispositional factors system trust and propensity to trust is assumed to be lower with a higher degree of be-goal achievement.

## 3 Method

To investigate the effects of be-goals on users' trust two personalized applications were tested in a within-design. The experimental study required interaction with both of the applications, a social network and a cloud service, as well as a rating of individual be-goal achievement and trust. The applications were chosen on the basis of a pre-test which defined the differences in be-goal achievement between them. One of the applications was facebook.com representing the social networks. The other application was the cloud service dropbox.com which is widely popular within the student population. All participants had an account on each of the systems and used them regularly. Both applications have a high relevance for the daily life of the participants and were considered to be equally accepted.

**Material**. In a pre-study the achievement of different types of be-goals was assessed. The scale of need satisfaction employed by Hassenzahl, Diefenbach and Göritz [4] was used. It consists of 30 items depicting each of the top ten psychological needs [5] by three items. We skipped the items for luxury and physicalness because they seemed inappropriate for the context of personalized applications. The remaining 24 items included the psychological needs autonomy, competence, relatedness, meaning, stimulation, security, self-esteem and popularity. Participants were asked to rate the level of need fulfilment they usually feel during the use of the application on a 6-point

Likert scale, e.g. “When using [the application] I feel that I am a person whose advice others seek out and follow”. For measuring users’ trust in the main study we used the scale on online users’ trust (SCOUT) [9, in preparation]. It contains 15 items and measures situational and dispositional aspects of trust on a 5-point Likert scale. Participants were asked for their level of agreement on four dimensions: perceived trustworthiness, perceived risk, system trust and propensity to trust. The items for perceived trustworthiness include the assessment of both interaction characteristics and characteristics of the application’s provider, e.g. “The application makes me think the provider is competent.” The items for risk include statements about the usage situation, e.g. “I feel it is insecure to use this application.”. Beside situational components of trust the scale measures also dispositional components like system trust (e.g. “For me the internet is a trustworthy environment.”) and the user’s propensity to trust (e.g. “I tend to quickly trust persons or things.”). Furthermore, the Web Analysis Measurement Inventory [10] was applied to control for differences in usability. It contains five dimensions (attractiveness, controllability, efficiency, helpfulness, learnability) with four items on each dimension. Agreement is measured on a 5-point Likert scale.

**Procedure.** In the pre-study the scale of need satisfaction [4] was administered online with a short introduction explaining both parts of the experiment. Participants rated their level of be-goal achievement for each of the applications (social network, cloud service) in randomized order. In the main study all of the participants were asked to interact with both applications in randomized order. For each application the initial task was to log in to the application with the participant’s own username and password. They were then asked to carry out a brief information search task to create a user experience immediately before they rated perceived trustworthiness and perceived risk of each application. By using their own accounts within the experiment the participants were meant to feel as close to a real usage situation as possible by still keeping up standardized conditions. Several questionnaires for assessing control variables were administered after the participants had finished the tasks.

**Participants.** A total of N = 34 students of Chemnitz University of Technology (23 female, 11 male) took part in the study. The mean age was M = 22.2 (SD = 2.9) years. They were all well-grounded in Internet use and spent about 26 hours per week online for private purposes. All of them had an account on both of the applications used. There were no significant differences in the reported personal importance of the applications. The participants felt equally connected to both their accounts on facebook.com and dropbox.com.

## 4 Results

The pre-study was conducted to check for different levels of be-goal achievement. Results show significant differences on the be-goals competence, relatedness, stimulation, security, self-esteem and popularity (Table 1) between the social network and the cloud service. The social network scored higher on relatedness, stimulation self-esteem and popularity whereas the cloud service showed higher values for competence and security. Thus, the social network clearly offers a higher number of be-goals achieved.

**Table 1.** Means, standard deviations and results from paired t-tests for the be-goal achievement (scale of need satisfaction, scale ranging from 1-6)

| | Social Network | | Cloud Service | | | |
|---|---|---|---|---|---|---|
| | M | SD | M | SD | t | p |
| Autonomy | 3.61 | 1.38 | 3.57 | 1.28 | .19 | .852 |
| Competence** | 2.74 | 1.25 | 3.68 | 1,28 | -4.09 | .000** |
| Relatedness** | 4.62 | 1.22 | 3.41 | 1.37 | 4.43 | .000** |
| Meaning | 2.44 | 1.17 | 2.21 | .98 | 1.44 | .159 |
| Stimulation** | 3.74 | 1.28 | 3.17 | 1.28 | 3.23 | .003** |
| Security* | 2.77 | 1.07 | 3.34 | 1.21 | -2.77 | .009* |
| Self-esteem* | 3.22 | 1.30 | 2.82 | 1.24 | 2.34 | .026* |
| Popularity* | 3.34 | 1.35 | 2.91 | 1.29 | 2.24 | .033* |

According to the different levels of be-goal achievement we expected differences in trust scores between the applications in the main study. The social network as the application with a higher level of be-goal achievement should evoke less perceived trustworthiness, more perceived risk and less users' trust than the cloud service which serves a lower level of be-goal achievement. For the analysis of the trust scores a t-test for related samples revealed significant differences in the total score for users' trust ($t(33)=-6.79$; $p<.001$). The total users' trust scale consists of the mean of the four subscales perceived trustworthiness, perceived risk and the dispositional subscales system trust and propensity to trust. The comparison of the situational subscales perceived trustworthiness ($t(33)=--5.73$; $p<.001$) and perceived risk ($t(33)=5.83$; $p<.001$) confirms the significant differences between both applications (Fig. 1).

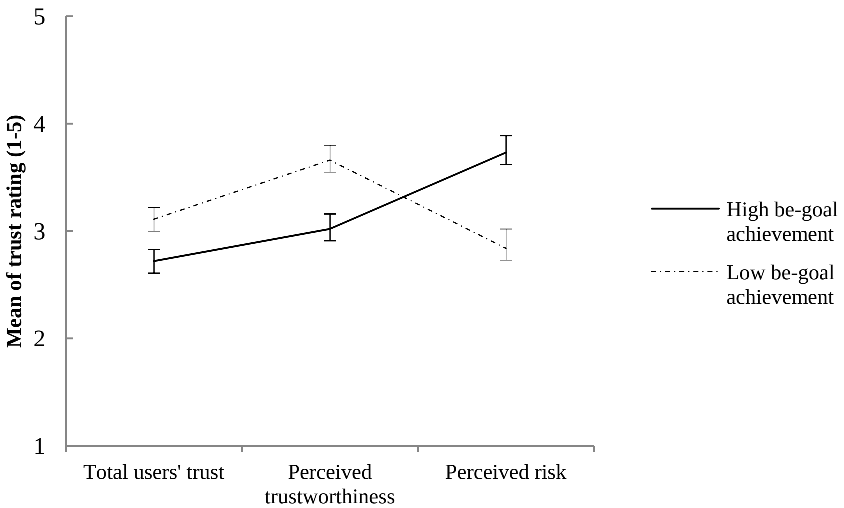



**Fig. 1.** Mean ratings on the scale of online users' trust, error bars indicate standard errors

Internal consistency for the scale on online users' trust (SCOUT) was α = .83 for the social network and α = .87 for the cloud service. The applications did not differ significantly on their usability scores ($t(33)=.769$, $p=.448$). Our results support the hypothesis that applications that fulfil more be-goals are related to less users' trust, less trustworthiness of the application and a higher perceived risk.

## 4 Discussion

Different degrees of be-goal achievement in personalized applications are related to different levels of users' trust. We found a higher be-goal achievement like relatedness, stimulation or self-esteem to be related to lower users' trust, lower perceived trustworthiness and higher perceived risk than a lower be-goal achievement. It seems that non-instrumental be-goals are associated with a higher personal meaning. People become more attached to applications the more intense the applications help to achieve their be-goals [3]. Therefore, potential negative consequences of interactions carry more weight and have to be considered when assessing trust in applications. If a user faces the risk of not achieving be-goals in an online situation of risk, the self will automatically be affected. That enhances vulnerability and hence, reduces perceived trustworthiness and aggravates risk. Furthermore, we discovered that different types of be-goals are differently related to users' trust in applications. Competence and security were fulfilled to a higher extend by the tested cloud service and were related to higher users' trust. Both competence and security refer to the feeling of being capable or in control. This connects directly to the concept of perceived trustworthiness, which is enhanced by the perception of ability. The distinction of be-goals between security and growth needs, which has been discussed by several authors [5], offers a possible explanation for the differences in users' trust related to those be-goals. Do-goals have not been an explicit part of this study. It is argued that for most products the fulfilment of do-goals can be seen as precondition for users' acceptance [11] and trust. The interplay between do-goals, be-goals and users' trust should be investigated in future research.

Our experimental study is to be seen as a first step in examining the relationship between be-goal achievement and users' trust in personalized applications. By using two existing applications with different levels of need satisfaction we created a setting of high external validity. Studies on users' trust gain validity by creating a real-life setting though we trade-off internal validity in return. For internal validity we controlled for influences on users' trust as far as possible. All of the participants had an own account on both of the applications. They used both accounts regularly on a voluntary basis. Both applications are well-known. Although facebook.com suffered from bad publicity lately we believe the high usage rates indicate that general acceptance was not affected by that. Usability as situational antecedent of trust was the same for both applications. By using a within test design we could eliminate the influence of personality factors. Still, there are external factors like social influences or media influences we did not cover in this study. For further investigation of the relationship between different levels of be-goal achievement and trust in personalized applications the effects of such external factors should be regarded. For enhancing internal validity a comparison of applications of only one type is advisable. Additionally, other applications apart from social networks and cloud services should be systematically tested. Depending on the service offered by different applications factors like the voluntariness of use or specific content of the application might influence the formation of users' trust.

Particularly for personalized applications users' motives and be-goal achievement should be considered before designing the system. Users' understanding of the underlying structure and the way their data is managed by the system do without question contribute to the formation of users' trust. But the degree of the achievement of be-goals – that determine the decision to use the application in the first place –is the key to a more detailed trust assessment and can give implications for designing trust-enhancing interfaces.

# 4 References


1. Lee, J. D., & See, K. A. Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80 (2004)
2. Beldad, A., De Jong, M., Steehouder, M.: How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. Computers in Human Behavior, 26(5), 857–869 (2010).
3. Hassenzahl, M.: Experience Design: Technology for all the right reasons. Synthesis Lectures on Human-Centered Informatics, 3(1), 1-95 (2010)
4. Hassenzahl, M. , Diefenbach, S., Göritz, A.: Needs, affect, and interactive products – Facets of user experience. Interacting with Computers, 22, 353-362 (2010)
5. Sheldon, K. M., Elliot, A. J., Kim, Y., Kasser, T.: What Is Satisfying About Satisfying Events ? Testing 10 Candidate Psychological Needs. Journal of Personality and Social Psychology, 80(2), 325–339 (2001)
6. Corritore, C. L., Kracher, B., Wiedenbeck, S.: On-line trust: concepts, evolving themes, a model. International Journal of Human-Computer Studies, 58(6), 737–758 (2003)
7. Zuckerman, M.: Behavioral expressions and biosocial bases of sensation seeking. Cambridge university press. (1994)
8. Vlek, C., & Stallen, P. J. Rational and personal aspects of risk. Acta Psychologica, 45(1), 273–300 (1980)
9. Bär, N., Jank, K. & Krems, J. SCOUTing trust in human-computer-interaction: The develop ment of the Scale for Online Users' Trust in websites. (in preparation)
10. Kirakowski, J., Claridge, N., & Whitehand, R.: Human centered measures of success in web site design. In Proceedings of the Fourth Conference on Human Factors & the Web. (1998)
11. Hassenzahl, M.: Hedonic, emotional and experiential Perspectives on Product Quality. In C. Ghaoui (Ed.), Encyclopedia of Human Computer Interaction (pp. 266–272). London: Idea Group Reference (2006)

# Trust evaluation on Facebook using multiple contexts

Tomáš Švec[1], Jan Samek[2]

Brno University of Technology, Faculty of Information Technology,
Božetěchova 2, 612 66 Brno, Czech Republic
[1]xsvect00@stud.fit.vutbr.cz, [2]samejan@fit.vutbr.cz
http://www.fit.vutbr.cz

**Abstract.** This paper applies the term trust from the point of view of artificial intelligence to social network analysis methods. It evaluates current available interactions for a model of trust considering various social networks. A mathematical model of trust for Facebook is designed. This model is implemented in Python programming language. Experiments are conducted on a sample amount of Facebook users and furthermore analysed from the perspective of both artificial intelligence and social psychology.

**Keywords:** social network, trust, multi-context trust, Facebook

*In a networked world, trust is the most important currency.*

Eric Schmidt

## 1 Introduction

The CEO of Google accurately commented on the current state of human emotions in a networked world in his speech for the University of Pennsylvania. The definition of social interaction has been radically transformed more than once in the past and present century. We reveal more and more of our inner selves on the Internet and there are a growing number of people in our vicinity called *friends* who we have never actually met. Although the artificial intelligence is still miles away from passing the Turing test [8], we still begin to answer the question whether it is possible to use patterns of human behaviour to simulate emotions.

This paper is aimed at creating a model of trust from the point of view of AI which would make use of social psychology in social networks. Basically, it is assumed that as the term *trust* originates in sociology and social psychology, it should be possible to apply this principle in its original field after 20 years and observe the differences. To achieve this, several terms have to be defined both in social psychology and artificial intelligence – similarities are observed and highlighted. Several examples of current social networks will also be briefly analysed and a representative network will be chosen for implementation.

The designed model of trust itself will be mathematically described, keeping in mind the necessity to minimize specific dependencies to be able to implement this model in a number of other networks. Reasons for correlation between various types of interactions and trust between entities will also be considered. Most importantly, the whole model and its implementation will be validated on real social network users and consequently summarized in the form of an exploratory investigation.

## 2 Social network analysis, trust and reputation

As the field of trust and reputation lies on the border of two scientific disciplines, sociology and computer science, it is sometimes impossible to adhere to strict technical description and mathematical definitions. A universal apparatus for describing human emotions has not been invented yet, after all. Despite these facts, this thesis leans toward computer science and therefore takes definitions from the field of artificial intelligence.

### 2.1 Trust

*Bruce Schneier*, a specialist on computer security and cryptography, considers the ability of building trust between individuals to be the cornerstone of modern society [9]. This mechanism may be tracked back to reciprocal altruism in some species. The following definition comes from one of the most renowned sociologists, *Anthony Giddens* [5]:

> *Trust is related to absence in time and in space. There would be no need to trust anyone whose activities were continually visible and whose thought processes were transparent, or to trust system whose workings were wholly known and understood. It has been said that trust is a device for coping with the freedom of others, but the prime condition of requirements for trust is not lack of power but lack of full information.*

In 1994 computer science was enriched by *Stephen Paul Marsh*, who influenced the field of artificial intelligence in a major way. He introduced trust into multi-agent systems in his doctoral thesis *Formalising Trust as a Computational Concept* [7]. His original understanding of the concept of trust came from the field of Humanities. Despite the precision and technical accuracy of his thesis, the term “trust” has never been fully defined in computer science, or to be more accurate, it has been defined in too many contexts and too many various situations. Marsh himself uses the definition from a famous psychologist, *Dr Deutsch* [2]:

1. The individual is confronted with an ambiguous path, a path that can lead to an event perceived to be beneficial ($Va+$) or to an event perceived to be harmful ($Va-$).
2. He perceives that the occurrence of $Va+$ or $Va-$ is contingent on the behaviour of another person.
3. He perceives the strength of $Va-$ to be greater than the strength of $Va+$.
4. If he chooses to take an ambiguous path with such properties, I shall say he makes a trusting choice; if he chooses not to take the path, he makes a distrustful choice.

### 2.2 Reputation

For the sake of readability, formal definitions of reputation are omitted in this text. It is, however, worth mentioning that the field of Humanities does not recognize reputation as a valid term. Social acceptance or trust perceived in groups of people or organizations is connected to social prestige instead. In the designed model, reputation could be derived from trust using an algorithm of arithmetic mean or similar techniques. In the context of trust and reputation, it is also important to describe what the *Dunbar number* is. We have seen a rapid rise of human society in the last few thousands of years. Biological evolution could, however, in no way compete against the pace of changes required for the human brain to adapt to modern society. As a result, we still have a fixed number of people we can keep track of in the matter of reputation. It happens to be the exact same number as the average population of a Neolithic settlement and also a rough average of the number of friends on *Facebook*. Today's scientists lean toward the value 150 [3].

### 2.3 Social network

This term is relatively new and dates back into the last century when *Barnes* described his stay in a *Norse village called Bremnes* [1]. Due to family traditions and isolation of this village from the rest of the world, Barnes was able to study some class phenomenon and categorize the inhabitants into groups. These relatively autonomous groups and their relationships were later described as a *social network*. The definition is as follows:

> *A social network is a social structure made up of individuals (or organizations) called "nodes", which are tied (connected) by one or more specific types of interdependency, such as friendship, kinship, common interest, financial exchange, dislike, sexual relationships, or relationships of beliefs, knowledge or prestige.*

## 3 Social networks: current situation

In spite of the prevailing endeavour to remain disclosed from any details of implementation that would concern a specific network, it was necessary to pick a deputy of social networks to demonstrate the formulated model of trust using real-life data. There were a few requirements concerning this deputy. Three most used social networks in Europe (*Facebook*, *VKontakte* and *MySpace*, according to InSites Consulting [10]) were amended with *Google+* and *Spoluzaci.cz*, two networks with bonds to the Czech environment. These were the desired treats of the deputy:

- more than one form of interaction on this network which can affect trust – these forms of interaction should also be actively used,
- it should be simple to use an API (Application Programming Interface) to access these services, the best option being an alternative from the service provider himself,
- location awareness would help us in the future to consider geographical factors in the analysis conducted using the model,
- the desired social network should be widespread, so that it is easier to collect representative data from real-life users,
- as the data would be collected in collaboration with people from the Czech Republic, it would also be desirable to have a number of Czech speaking users.

The results can be summarized in this table showing *Facebook* as the winner:

| | Interactions | API | GPS | Penetration | Czech |
|---|---|---|---|---|---|
| Myspace | x | | x | | |
| Google+ | x | x | x | | |
| Facebook | x | x | x | x | x |
| Twitter | | x | x | x | |
| Spolužáci | x | | | x | x |

**Table 1.** Social network properties.

## 4 Multi-context trust model for Facebook

The mathematical core of this model leans on a theory distributed by *Marsh* in his founding thesis [7]. This theory introduces so-called contexts of trust which represent the fields in which we are capable of trusting the entity. To explain this term in a simplified example, "I trust my brother to drive me safely to the airport, but I would feel very insecure if he were to fly my plane." Dividing trust into contexts is the only reasonable way to comprise a thing as complex as trust while maintaining the possibility of flexible changes and further development. Every context is normalised into the interval from 0 to 1 to facilitate future aggregation.

This model was designed for the possibility of implementation for multiple contemporary social networks. It was, however, necessary to implement and test this theory for a particular social network. Although the described contexts stand on functionality provided by Facebook, chosen interaction types are present in other networks as well. Please note that methods of computation were chosen according to the environment the tests took place in. Several optimizations aimed at robustness, accuracy or speed may be considered, including saturation of values (meaning extreme values shall be restricted not to distort obtained results), omitting larger groups that anyone has a high probability to be a member of, or analysing the content of text, not only its quantitative measures.

### 4.1 Trust contexts

A short description of the investigated *trust contexts* is described in the following sections.

**Interaction time span** This context seems to be the most intuitive one. The longer the time between the first and the last interaction, the higher trust we are likely to feel, even though there may be exceptions concerning people we contacted soon after joining the community.

**Number of interactions** The term *interaction* stands for one-way information channel, in this case – wall posts, comments and "likes". Their overall number should be counted and normalized using the following formula (equation 2):

$$A = \frac{1}{n} \cdot \sum_{x=1}^{n} I_x \tag{1}$$

$$T_N(x) = \frac{I_x}{A + \frac{1}{n} \cdot \sum_{x=1}^{n} |A - I_x|} \tag{2}$$

$I_x$ stands for number of interactions with person $x$, $A$ is the average number of interactions and the fraction divisor is a sum of the average number of interactions and absolute deviations of all acquired values. This formula provided the most reasonable results according to the first three respondents and was later empirically confirmed in the experiments.

Exclusion of extreme spikes represented by overly-active users is crucial here. Heuristics for this case include setting a maximum value.

**Number of characters** Several works in the area of trust study the relation between a number of characters in a message and the credibility of the writer [6]. As these works often belong to another application domain, this context is not given so much impact in the model. Setting a ceiling for the maximum number of characters is very important here, since copy & paste skills would be the easiest way to influence the model for educated users.

**Interaction regularity** Regularity differentiates people engaged in heated, yet scarce discussions which would normally boost a person's computed trust way above appropriate level. It is natural to trust people we communicate with on daily basis more than people that we had contact with in the past. One way to compute this context is using the formula introduced in the thesis [11] (equation 3):

$$x_v^{\Delta T}(A, B) = \prod_{i=1}^{n-1} |t_{i+1} - t_i| \tag{3}$$

There is an implementation issue, however, when we consider the amount of data and the necessity to express time in milliseconds. This formula would bring the most satisfying results at the cost of wide data type range. This context is therefore computed in a simplified manner. A set of perfectly regular intervals for the fixed number of interactions is computed and then compared to the real values.

As this statement may be a little unclear, a simple example shall be provided. Let us say we have four interactions to be considered, all of them occured shortly after the beginning of our friendship with the researched person. Our timespan for analysis is three months. Say we wanted to communicate regularly with this person. That would mean the first interaction occured at the beginning (which is correct and gives a small deviation). The second interaction should have occured after one month (which is still relatively close). The third and fourth interaction, however, should have occured at the end of the second and third month. If we compare these values to the ones close to the beginning, we get a very high deviation.

Based on the previous paragraph, we can see that the more interactions are irregular, the higher the deviation. This fact led to the necessity to invert the value to correspond with the rest of the contexts.

**Photo tagging** Photo tags have a very important meaning for trust. They usually indicate a link of people in the real world. There are special cases which should be considered (Christmas wishes would be a very good example, their informative value is next to nothing), but generally this context is very important for the resulting model.

**Group membership** A certain terminological ambiguity should be explained here. Groups and pages were not distinguished in Facebook initial times. Groups in this context represent a set of people who share a common trait, for example people who commute to the same city, people who work on the same project or people from one regional country unit. The more groups two people share, the more likely it is they trust each other. There is an inverse relationship between the size of groups and their importance. A shared smaller group usually means that these two subjects trust each other.

**Common interests** The only context which does not depend on any interaction and can be computed for any two people around the world. It builds upon the premise that people who share similar interests (like the same page here) are likely to trust each other more. This statement can be found in many papers on the subject, [9] serves as an example. A similar inverse relationship about size can also be applied here. This context, however, is the most time-consuming to compute and requires a lot of bandwidth. In case of time-critical operations, this is the part which should be omitted first, as it serves as more of an experimental feature.

**Number of friends** Due to the inconsistency in Facebook Graph API [4], this context was not implemented in the final model. It can be related to the previously mentioned "Dunbar's number". The deviation from the standard and widely accepted number of friends could also be considered an interesting factor for computing trust. People who have way more friends than the average number in their country may express similar traits. The same goes for the other extreme. This statement depends on many factors, though, and should be considered in connection with age groups.

### 4.2 Trust aggregation

These seven (eight) contexts should be aggregated in a way which allows us to establish an order relation. Marsh simply multiplied his contexts and used the resulting values. This approach fails here because of different importance of individual contexts. For this purpose, a priority vector (equation 4) is introduced in this model. It is a vector of numbers where $T_x$ represents the priority for given context.

$$P = (T_S, T_N, T_C, T_F, T_P, T_G, T_L) \tag{4}$$

The final value of trust can be obtained with this formula (equation 5):

$$T_x = \frac{S \cdot T_S + N \cdot T_N + C \cdot T_C + F \cdot T_F + P \cdot T_P + G \cdot T_G + L \cdot T_L}{S + N + C + F + P + G + L} \tag{5}$$

This method of aggregation enables us to attribute each context with its importance. If, for instance, we find a context less contributing to overall trust in our recent findings, we simply decrease the level of importance in the priority vector. Similarly, a completely new context may be added to the existing set and this expansion is also planned in the nearest future.

As for the particular model used in the experiments, the vector $(1, 3, 2, 2, 1, 2, 3)$ was used. Individual priorities were chosen based on empirical experience of the first 3 experimental users. The values were, however, retrospectively checked in the survey of participating users. Results showed apparent oscillation towards this choice of numbers as well.

## 5 Implementation and experimental results

As the model was intended to be deprived of any implementing details, the implantation itself shall only be described in a very brief manner. The particular example was implemented in the Python 3 programming language using the application interface supported by Facebook called Graph API [4]. Graph API produces data in JSON format, hence the need of Python's in-built libraries. Authentication is provided by the OAuth 2.0 technology.

The greatest issue encountered when collecting the data from users was how to get only limited access to their profiles and persuade them that no harm would come to their privacy. For this particular purpose, OpenGraph provides so-called access tokens, which can be generated on the developers' page and can be used to configure privileges for the holder of the token for a limited amount of time.

### 5.1 Exploratory investigation

As the research could be considered invasive by some users, quantitative research was not a valid option. Users with valid data for experiments consider their internet identities a part of their lives and therefore do not willingly provide access to their profiles. An exploratory investigation was a compromise and provided the possibility to work with a limited number of respondents and to ask relatively simple questions.

The exploratory investigation included 18 respondents randomly chosen in the age interval from 17 to 30 years. Men and women were both equally represented. The analysis was conducted for the time-span from 1.4.2011 till 1.5.2012. Results were verified by the respondents themselves using a questionnaire consisting of closed questions with the utilization of scaling.

Certain criteria had to be met in order for the user to participate in this investigation. The only condition was for the profile to be regularly used. Participating users were sent a short PDF file describing the procedure of generating their access token and also explaining which personal data they were making accessible. While the script was running (around 5 minutes for an average profile), they were given a simple command to record their answers for later use: "Name ten people you trust most on Facebook." Keeping this information to themselves was a key part of the investigation. They would perhaps try to obfuscate the initial guess if they were to show it to another person. This way, they were the only people who knew the answer.

After seeing the results of the scripts, users had to answer these questions with multiple choice answers:

1. How many people you listed actually occurred in the script's results?
   – Possible answers: 0–10.

2. How many people's trust was wildly mismatched?
   – Possible answers: 0–2, 3–5, 6–8, 9–11, 12 or more.
3. What actions among friends do you find most important for trust on Facebook?
   – Possible answers: Values 1–5 for these categories:
     - private messages,
     - comments,
     - "Like" tags,
     - common photographs,
     - common groups,
     - interaction regularity.

### 5.2 Experiment results

Question number 1 was the key element of this questionnaire. Resulting values form a fairly regular Gaussian curve. Most results converge to the number 5 and the arithmetic mean of all the values is 4.83. The figure 1 shows the number of respondents with each individual answer.

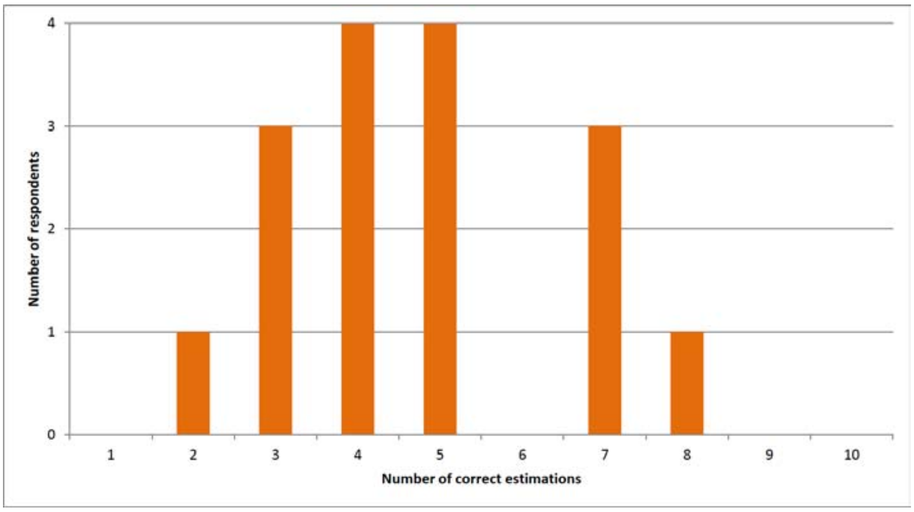



**Fig. 1.** Answers corresponding to expected results.

Question number 2 was designed to detect the most significant flaws of the model. Respondents were given the possibility to state whether someone's trust was way below/above expected values and this way they also verified the results themselves. Most respondents (11 answers 0–2) stated there were not as many deviations as one would expect. The figure 2 shows the deviation for respondents.

Question number 3 aimed at the credibility of the used priority vector. In this state, there must be a person setting the priority vector according to his/her preferences and acquired statistical data. One of possible expansions, however, relies on the possibility to change this model dynamically according to amount of collected data. So far, users seem to copy the initially set priority vector values.

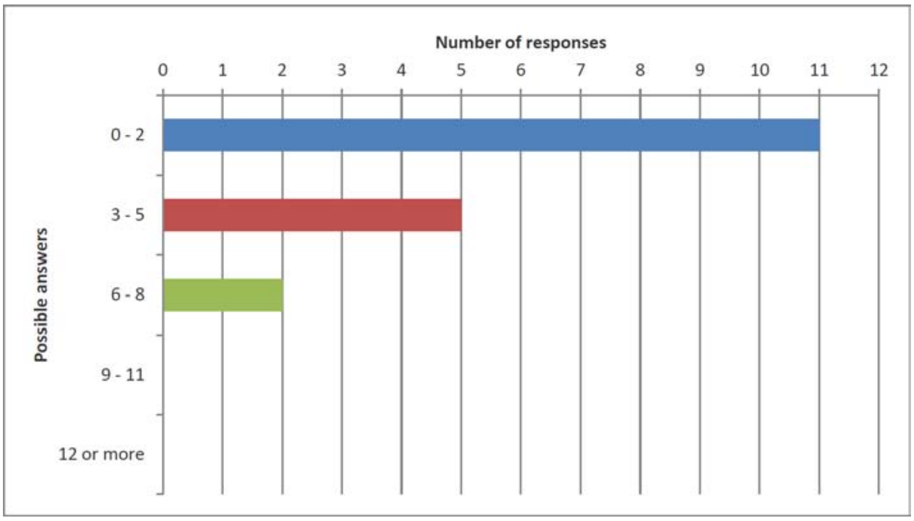



**Fig. 2.** Significant deviation of the model.

## 6 Conclusion

The goal of this paper was to analyse current situation in social networks from the point of interactions, design a model of trust for social networks, implement it and test its correspondence to the real world. The best evaluation of fulfilling these tasks is the experimental result:

Based on the respondents' answers, the model can evaluate correct trust with 48.3% probability. This number may seem like an unsatisfactory result. On the other hand, the model was given only information that (in most cases) is freely available on the web to anyone administering any Facebook account. Considering the best safety available, this information can still be seen by our friends, whose numbers, as we have learned, vary around the number 150. Would it be disturbing to the users that these 150 people can guess half the people they trust most on this network and use them for social engineering?

There are multiple paths this model could take in development. Since the very beginning, new contexts were intended to be added to this model, for example the similar number of friends or private messages analysis. Another possibility is to dynamically adjust the priority vector according to the amount of collected data. Users would also welcome an HTML interface for conducting the research themselves. Almost all respondents who participated in the exploratory investigation expressed this wish.

### Acknowledgment

This work was partially supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and by the project CEZ MSM0021630528.

## References

1. J. A. Barnes. Class and committees in a norwegian island parish. *Human Relations*, 7, 1954.
2. Morton Deutsch. *Cooperation and Trust: Some Theoretical Notes*, pages 275–320. Nebraska University Press, 1962.
3. Robin Dunbar. *Grooming, Gossip, and the Evolution of Language*. Harvard University Press, 1998. ISBN 0674363361.
4. Inc. Facebook. Graph API [online]. `https://developers.facebook.com/docs/reference/api/`, 2011-04-02 [cit. 2013-04-16].
5. Anthony Giddens. *The Consequences of Modernity*. Polity Press, Oxford, UK, 1990. ISBN 0745607934.
6. Audun Josang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2), March 2007.
7. Stephen Paul Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, University of Stirling, April 1994.
8. Graham Oppy and David Dowe. The turing test [online]. `http://plato.stanford.edu/archives/spr2011/entries/turing-test/`, 2011-01-26 [cit. 2013-04-16].
9. Bruce Schneier. *Liars and Outliers*. Wiley, 2012. ISBN 978-1118143308.
10. S. van Belleghem. Social media around the world 2011 [online]. `http://www.slideshare.net/stevenvanbelleghem/social-media-around-the-world-2011`, 2011-09-13 [cit. 2011-11-12].
11. Lizi Zhang, Cheun Pin Tan, Siyi Li, Hui Fang, Pramodh Rai Yao Chen, Rohit Luthra, Wee Keong Ng, and Jie Zhang. The influence of interaction attributes on trust in virtual communities. In *International Conference on User Modeling, Adaptation and Personalization (UMAP)*, 2011.

# Know Your Members' Trust

Sanat Kumar Bista, Surya Nepal, Cecile Paris

CSIRO ICT Centre
Marsfield, NSW, Australia

firstname.lastname@csiro.au

**Abstract.** This paper outlines the findings of a survey on trust, captured through attitude, experience, behaviour and expectation, of members in a Government run online support network. Overall, the results show that participants have high expectations for the behaviour of others in the community, expecting them to be trustworthy, helpful and supportive. There is a gap, however, between the respondents' own attitude and behaviour with respect to trust and what they expect of others in the community. The results of this survey will serve as a baseline against which to compare results obtained at the end of our community trial. To the best of our knowledge, this is the first time such a trust survey has been conducted in an online community to establish the initial baseline members' trust. We also present the initial results obtained at the end of the trial.

**Keywords:** Government Support Network, Social Network, Trust Survey

## 1 Introduction

Online social networking sites are often seen as a place for people to obtain social, emotional and moral support from others on the site. In the health domain, for example, they have been shown to have a positive impact (e.g., [1-4]). In partnership with the Australian Government's Department of Human Services (referred hereafter as *Human Services*), we have trialled an online community to investigate whether online communities could be beneficial to provide support to welfare recipients [5,6]. Our trial targeted parents transitioning from a parental payment to another income support benefit with the requirement to find a job, a transition that occurs when their youngest child is reaching school age. The transition is a difficult one for most parents.

Our online community was called *Next Step.* It was meant to be a place for people to find support from others in a similar situation, with the hypothesis that this would be helpful in their transition process. Individuals in the community are strangers to each other – but they all share the same situation. *Next Step* is also a place for the government to target its information and support services when dealing with this specific group of welfare recipients. In a community such as ours, it is important for its members to trust each other and the community provider. This is necessary for people to participate in the community, speak freely and share their experiences.

One of our hypotheses is that building online communities serves not only to provide informational and emotional support to target groups, but also to increase social trust through interactions in the community. To this end, we first sought to understand and establish initial trust values of individual members, i.e., their trust values before they join the community. This would provide a baseline against which to evaluate the increase of social trust values at the close of the community. We did this through a survey entitled ‘Knowing you better’, conducted within the first week of people joining the community. Trust was captured through a set of questions related to their attitude, experiences and behaviour. Individual members’ behaviour is mainly driven by their attitude. Similarly, their expectations about others are built through their experience. To the best of our knowledge, this is the first time such a trust survey has been conducted in an online community. It certainly is the first of its kind for an online community for welfare recipients. In this paper, we describe the design of the survey and present the corresponding results. We also conducted an exit poll to measure the change of social trust. We present the initial results of the exit poll and our observations.

The remaining of the paper is structured as follows: Section 2 provides some background on trust issues. Section 3 presents a brief review of the design of the survey and structure of its questions. Key observations are presented in Section 4, and discussed further in Section 5. Finally, Section 6 gives some concluding remarks.

## 2 Background

Trust is widely accepted as a major component of human social relationships and studied in different disciplines ranging from Sociology [7-9], Psychology [10,11], Economics [12,13] to Computer Science [14] and online service provisions [15]. In general, trust is a measure of confidence that an entity will behave in an expected manner, despite the lack of ability to monitor or control the environment on which they operate [16]. Trust plays an important role in the bootstrapping and sustainability of the online communities. Recently, there has been an increasing interest on trust and its role in social networking [17]. However, the majority of research in this area has focused on the computational aspects of trust, i.e., evaluating the reputation of a node or trust between the nodes using different features (e.g., rating, like/dislike, voting, social circle, etc.) of the social networks [18]. None of this research has focused on studying the impact of social networks on human aspects of trust (i.e., social trust). Social trust implies that members of a social group act according to the expectation that other members of the group are also trustworthy [19] and expect trust from other group members. Similarly, social capital is the quantity of trust a member has to other members in the society [20].

Our aim in *Next Step* is to understand social trust and see whether the use of online communities for delivering human services can eventually increase the social capital (i.e., the social trust between members and towards governments). To this end, we first need to measure the trust of an individual before coming to the community. We use questionaires developed and used in social science to measure the initial trust value.

How do you measure the trust value? Trust is measured using three human characteristics: attitude, behaviour and experience. We considered the following factors:

- People's trusting *attitude* towards *people in their own surrounding* (e.g., home, office, society, etc.).
- People's trusting *behaviour* towards *known people* (e.g., friends) in their own surrounding.
- People's trusting *behaviour* towards *strangers* in their own surrounding.
- People's trust *experience* from other people in their surrounding, including strangers.

It is important to understand these factors to establish the baseline trust values so that we can measure whether online communities could improve social capital. In addition to capturing people's attitude, behaviour and experience in their surrounding, we also need to consider the reciprocal attitude, behaviour and experience expected from other members in the community, so that we can also uncover the gap between individuals' own trusting attitude and behaviour and their expected trusting attitude and behaviour from others. Various tools have been used in social and behavioural sciences to measure these factors [21-24]. To the best of our knowledge, they have not been used for measuring the initial trust values of members in online communities.

## 3 Research Methodology

When possible, we adapted a standard set of questions defined and used in social and behaviour sciences. We added some questions dealing with interpersonal trust, a concept central to social sciences linked to collaboration and coordination between individuals within a network [25]. These new questions were adapted from [25].

### 3.1 Capturing Trust Attitude

With the intent to understand members' attitude towards trust in general, we adapted the *General Social Survey* (GSS) questions which act as a primary source of evidence on trust and social capital in the United States [26]. We used this instrument because of its wide use over time and space [27]. We took the three General Social Survey (GSS) questions on Trust, Helpfulness and Fairness shown in Table 1. In addition to the answer choices indicated in the Table, users could choose to answer: "don't know" or "don't want to answer".

**Table 1.** GSS Questions on Trust, Helpfulness and Fairness

| Questions | Answer Choices |
|---|---|
| Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people? | most people can be trusted;<br>can’t be too careful about<br>depends on the situation |
| Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves? | try to be helpful<br>just look out for themselves<br>depends on the situation |
| Would you say that most people would try to be fair or that they try to take advantage of you if they get the chance? | would try to be fair<br>would take advantage of you<br>depends on the situation |

### 3.2 Capturing Trust Experience and Behaviour

Six questions, adapted from [26], were employed to capture trust experience and behaviour. They are shown in Table 2. The first question captures the trust experience, and the others capture the trust behaviour. The answer choices to all six questions were: always, often, sometimes, rarely, never and don’t want to answer.

**Table 2.** Questions on Trust Experience and Behaviour [26]

| |
|---|
| Have you ever benefited from a person you did not know before? |
| You lend personal possessions (e.g., book, car, bicycle, etc.) to your friends. |
| You lend money to your friends. |
| You leave your door unlocked. |
| You lend personal possessions (CDs, book, car, bicycle, etc.) to a person you hardly know. |
| You lend money to a person you hardly know. |

### 3.3 Capturing Trust Expectation

We developed five questions (see Table 3) to capture members’ expectations about other members in the community. The first three questions capture one’s expectations about the attitude of others in the community, and the next two relate to one’s expectations about the behaviour of others in the community. The members were asked to answer the following questions by considering specifically the members of the *Next Step* online community. As with the first set of questions, the answers “don’t’ know” and “don’t want to answer” were also available.

Unlike previous questions which aimed to uncover the general trust attitude and experience of members, i.e., with respect to the world at large, these questions are specific to the other people one expects to meet (albeit virtually) in the *Next Step* online community.

**Table 3.** Questions on Trust Expectation

| Question | Answer Choices |
|---|---|
| You will expect them to be | very trustworthy, somewhat trustworthy, untrustworthy |
| You will expect them to be | very helpful, somewhat helpful, unhelpful |
| You will expect them to be | very supportive, somewhat supportive, unsupportive |
| They will generally share their knowledge with you. | Agree, somewhat agree, disagree |
| They will generally share their experience with you. | Agree, somewhat agree, disagree |

## 4 The results

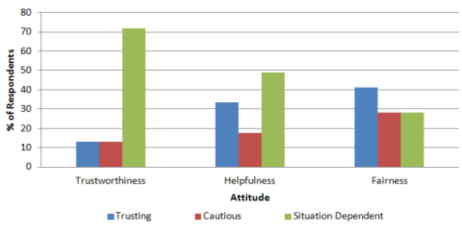
The survey 'Knowing you better' was done as a poll in the first week of people joining the community. The community was built over a period of twelve months through four phases of recruitment. Respectively, 55, 30, 26 and 152 members joined the community during these four recruitment phases, but only 99 of these visited the community at least once. Of those, 46 completed the survey from each recruitment (about half). This means more than 8% of the total registered members have completed the survey, which is nearly equal to the proportion of highly active and active members of the 90-9-1 Jacob Nielson's rule (a community often has 1% very active, 9% active and 90% passive members). We present some of the results here.

### 4.1 Trust Attitude

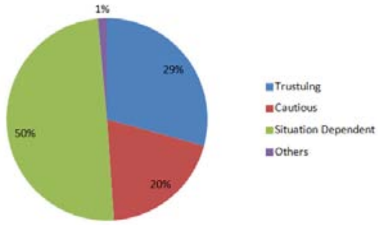
We grouped members in three categories based on their answers:

- "Trusting" for those who answered "Most people can be trusted", "Most people try to be helpful" and "Most people try to be fair";
- "Situation-dependent" for those who answered "Depends on the situation" for the three questions;
- "Cautious" for those who answered "Can't be too careful" for the three questions; and
- "Other" for those who answered "Don't Know" and "Don't Want to Answer" for the three questions.

We first look at the individual questions on attitude (Fig. 1 (a)). We note that the question about fairness received the largest number of trusting responses (41%) (i.e., "most people try to be fair") as compared to the questions about general trust (or trustworthiness) (13%) (i.e., "most people can be trusted") and helpfulness (33%) (i.e., "most people try to be helpful"). There was no response in the "other" group. We now combine the results from the three individual questions by computing their mean value in different categories. Fig. 1 (b) shows the proportion of people in each category. The largest category is "situation-dependent", i.e., people are not necessarily trustworthy, helpful and fair by default, and a situation or context plays a role.




(a)




(b)

**Fig. 1.** (a) Responses in the individual components of attitude, (b) General Community Attitude

Interestingly, 29% of respondents had a trusting attitude towards the world around them: they thought people could be trusted, were helpful and fair. So we have, from the start, a small core of people whose attitude is trusting. According to the Australian Bureau of Statistics results in the 2010 GSS [28], 54% of people (Australians) above 18 years say that most people can be trusted. This figure is much higher than the result for the members in our community. There could be many factors that could have influenced the response, including, for example, the specific demographics involved, or the fact that our respondents are in a transition phase and thus particularly stressed, and, as a result, cautious of the world around them. This, however, is *not* a conclusion and further research is necessary to understand such influence, if any, in detail.

### 4.2 Trust Experience and Behaviour

As the *Next Step* community is anonymous, we want to know people's *a priori* experience with strangers. Fig. 2 (a) presents the results. The majority (about 48%) reported having sometimes benefited from strangers, and 2.17% have often benefited from strangers. Overall, half of our community (if we combine "always" and "sometimes") has had reasonably good experiences from unknown people in the past. We also note, however, that 19.57% of respondents have never had any experience of altruistic behaviour from strangers.

We now look at our members' behaviour towards others, grouping the questions as follows:

- Behaviour with friends, i.e., lending personal possessions and money to friends;
- Behaviour with strangers, i.e., lending personal possessions and money to strangers; and
- General Behaviour: the question on leaving door unlocked.

Fig. 2 (b) shows the results. Unsurprisingly, the graph shows that members show more trust towards friends than towards strangers in terms of lending “things”. Interestingly, their general trust behavior in their own environment (“leaving door unlocked”) is higher than lending “things” to strangers.

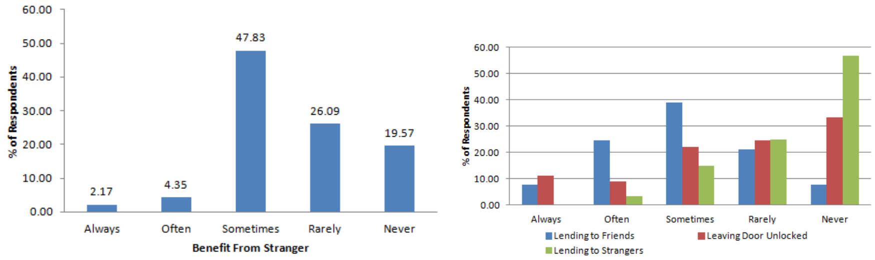



**Fig. 2.** (a) Trust Experience (Benefit from Strangers), (b) Trust Behaviour in Different Categories

### 4.3 Trust Expectations

We have so far discussed the aspects of trust that concerns someone’s attitude, experience and behaviour with respect to others. We now look at the questions of trust *of others in the community*.

We have two categories:

- Expectations about the attitude of other members’ in the community (e.g., “I expect others to be trustworthy/helpful/supportive”). We refer to this as the expectation about ‘community attitude’; and
- Expectations about behaviour: the behaviour people expect of other members in the community (e.g., “I expect others to share their knowledge and experience with me”). We refer to this as the expectation about ‘community behaviour’.

For the expectations about attitude, we first group members into the following four categories:

- “High Expectation” for those who answered “Very Trustworthy”, “Very Helpful” and “Very Supportive” for the three questions;
- “Cautious Expectation” for those who answered “Somewhat Trustworthy”, “Somewhat Helpful” and “Somewhat Supportive” for the three questions;
- “Bad Expectation” for those who answered “Untrustworthy”, “Unhelpful” and “Unsupportive” for the three questions; and
- “No Expectation” for those who answered “Don’t Know” and “Don’t Want to Answer” for the three questions.

Fig. 3 (a) shows the population distribution of the community responding to different categories. We see that a larger portion of the respondents expects at least some amount of trust, help and support from other members in the community. (If we put together the groups with “high” and “cautious” expectations, we get 76%.) The *Bad Expectation* group represents only 4% of respondents. 20% do not know what to expect.

Fig. 3 (b) presents the individual breakdown of the responses for the expectations of trust, helpfulness and support. We see that the majority of respondents had a *Cautious Expectation* in all categories.


**Fig. 3.** (a) Responses on the trustworthiness, helpfulness and supportiveness of the community, (b) Individual break-down of community expectation: attitude

## 5 Discussion

The survey analysis gave us a baseline for trust attitude and behaviour. It also provided further insights into members’ trust attitude, behaviour and expectations. Our community members, as a whole, seem to have relatively low trust attitude, comparing to the Australian average, and behaviour. Yet they expect a high trusting attitude and behaviour from other members in the community. A comparatively higher behaviour and attitude expectation is potentially a very positive indication that a community like *Next Step* has a potential to have positive effects on social capital and social trust. In the ideal case, the community members would, at the end of trial, have their own attitude and behaviour match the expectations they have from others.

We further analyse the trust behaviour of members. It shows that community members have benefited from strangers more than they are willing to lend to strangers. This reinforces the gap identified between the members’ own behaviour and the expected behaviour from others.

The high expectations from other members in the community comparing to expectations from strangers (shown by attitude towards strangers) might indicate that people do not see other community members as strangers – this is a fairly typical phenomenon in online communities, where people exhibit behaviour they would not normally exhibit with total strangers (such as sharing personal stories), even though people are strangers to each other, because of the connections people feel with each other by being in the same community.

In order to gain further insights, we have examined the data from the community in the light of the survey results. We gathered the login data of all members who have responded to the survey. We grouped them into two categories: “frequent visitors” and “overall respondents”. We define “frequent visitors” as those respondents who visited the community at least 15 times or more since registration.

Our first comparison is between the overall respondents’ trust attitude to that of the frequent visitors. We observed that there is no significant difference on “trusting” attitude between frequent visitors and overall community, see Fig. 4 (a). However, different results are observed in trust attitude expectations and behaviour (see Fig. 4 (b) and (c)). In both, the frequent visitors had high expectations from other members in the community. This means frequent visitors had similar trust attitude to that of overall community when the world around them is considered, but had higher trust behaviour of themselves, and more of them also had high expectations from others in the community.

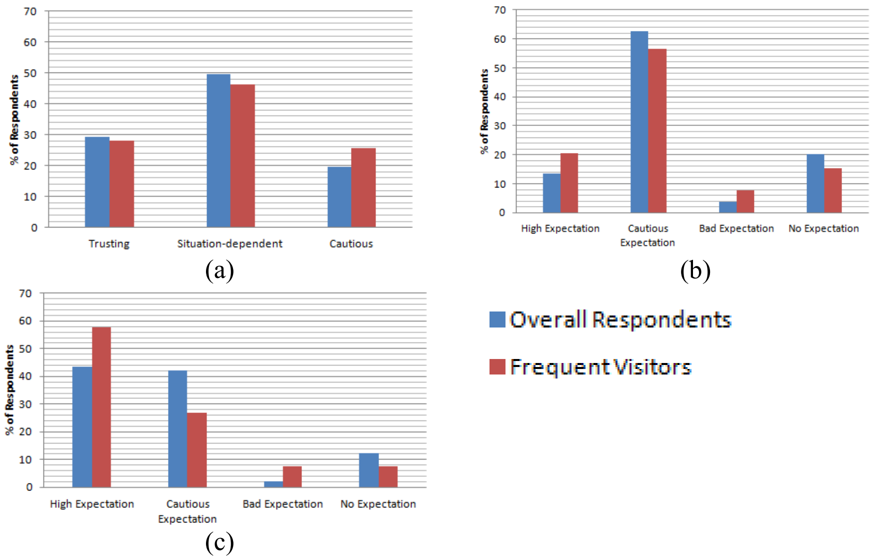



**Fig. 4.** (a) Trust Attitude, (b) Expectation Attitude, (c) Expectation Behaviour

In order to measure the increase in social trust, we ran an exit poll towards the end of the trial. We repeated the same set of questions that were asked in trust expectation as shown in Table 4. Purpose of the exit poll is to get an answer to the following:

(a) Does the online community like *Next Step* help to increase the overall trust of members towards each other and moderators? The answer to this question will help to understand the role of online communities in increasing social trust.

(b) Do members of online communities like *Next Step* value the role of the moderators? The answer to this question will help to understand and design the roles of moderator in online communities like *Next Step*.

Below we report an initial analysis of the exit poll to answer the first question.

**Table 4.** Exit Questions on Trust

| **Question** **Consider the members of this online community** | **Answer Choices** |
|---|---|
| Would you say that most people were: | very trustworthy, somewhat trustworthy, untrustworthy |
| Would you say that most people were: | very helpful, somewhat helpful, unhelpful |
| Would you say that most people were: | very supportive, somewhat supportive, unsupportive |

Exit poll was returned by 9 members, out of which 5 had also responded to the initial trust questionaries. Out of the 15 possible answers, 2 answers remained the same as initial ones whereas 13 answers moved to a more positive value, and none of the answers move to a more negative value. Though the result is based on a small number of responses and is thus not conclusive, it shows that the overall social trust in the community has increased.

# 6 Conclusion and Future Work

This report presented the findings from the trust survey that was carried out at the start of an online community project. Trust attitude, experience, behaviour and expectation as well as expectations about the behaviour of others with respect to fairness, helpfulness and support were presented based on the community members' responses. The results of the analysis show that the members had overall positive expectations from the community, although they did not themselves seem to have a trusting behaviour towards strangers. There is thus a gap between members' own attitude and behaviour about trust and their expectation from others. We hoped that the *Next Step* community would help reduce this gap, and that interactions in the community would lead to an increased social capital. We repeated the survey at the end of the trial for the community. The initial results show that the overall social trust in the community had increased.

**Acknowledgements**

This research has been funded under the Human Services Delivery Research Alliance (HSDRA) between the CSIRO and the Australian Government Department of Human Services. We would like to thank Nathalie Colineau, Payam Aghaei Pour, Brian Jin, Alex Sun and Bo Yan at CSIRO for their contribution in the design and implementation of the survey, and Gina Beschorner and her team at the Australian Government for their support and involvement in the *Next Step* community.

## 7 References


1. Skeels MM, Unruh KT, Powell C, Pratt W Catalyzing social support for breast cancer patients. In: 28th international conference on Human factors in computing systems, Atlanta, Georgia, USA, 2010. ACM, pp 173-182
2. Ma M, Agarwal R (2006) With a Little Help from Strangers: Social Support and Smoking Cessation in Online Communities. RHS School of Business, Univeristy of Maryland. http://www.rhsmith.umd.edu/news/releases/2006/120406.aspx. Accessed 18th July 2012
3. Newman MW, Lauterbach D, Munson SA, Resnick P, Morris ME It's not that i don't have problems, i'm just not putting them on facebook: challenges and opportunities in using online social networks for health. In: ACM 2011 conference on Computer supported cooperative work, Hangzhou, China, 2011. ACM, pp 341-350
4. Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, Thomas EJ (2010) Social support in an Internet weight loss community. International journal of medical informatics 79 (1):5-13
5. Bista SK, Colineau N, Nepal S, Paris C The design of an online community for welfare recipients. In: Proceedings of the 24th Australian Computer-Human Interaction Conference, elbourne, Australia, 2012. ACM, pp 38-41
6. Bista SK, Colineau N, Nepal S, Paris C Next step: an online community to support parents in their transition to work. In: Proceedings of the 2013 conference on Computer supported cooperative work companion, San Antonio, Texas, February 23-27 2013 2013. ACM, pp 5-10
7. Helbing D (1994) A Mathematical Model for the Behavior of Individuals in a Social Field. Journal of Mathematical Sociology 19 (3):189-219
8. Mollering G (2002) The Nature of Trust: From Geog Simmel to a Theory of Expectation, Interpretation and Suspension. Sociology 35:403-420
9. Molm LD, Takahashi N, Peterson G (2000) Risk and trust in social exchange: An experimental test of a classical proposition. American Journal of Sociology 5 (105):1396-1427
10. Rotter JB (1967) A new scale for the measurement of interpersonal trust. Journal of Personality 35 (4):651-665
11. Cook KS, Yamagishi T, Cheshire C, Cooper R, Matsuda M, Mashima R (2005) Trust building via risk taking: A cross-societal experiment. Social Psychology Quarterly 2 (68):121-142
12. Granovetter M (1985) Economic Action and Social Structure: the Problem of Embeddedness. American Journal of Sociology 91:481-510
13. Huang F (2007) Building Social Trust: A Human-Capital Approach. Journal of Institutional and Theoretical Economics (JITE) 163 (4):552-573
14. Maheswaran M, Tang HC, Ghunaim A Towards a gravity-based trust model for social networking systems. In: 27th International Conference on Distributed Computing Systems Workshops, 2007. ICDCSW'07. , Toronto, Ontario, Canada, June 25-29, 2007 2007. IEEE, pp 24-24
15. Josang A, Ismail R, Boyd C ( 2007 ) A Survey of Trust and Reputation Systems for Online Service Provision Decision Support Systems 43 (2):618 - 644
16. Singh S, Bawa S (2007) Privacy, Trust and Policy based Authorization Framework for Services in Distributed Environments. International Journal of Computer Science 2 (2):85-92
17. Dwyer C, Hiltz SR, Passerini K (2007) Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. Paper presented at the Thirteenth Americas Conference on Information Systems AMCIS 2007, Keystone, Colorado, USA
18. Golbeck J (2008) Computing with Social Trust. Springer Publishing Company, Incorporated,
19. Lewis JD, Weigert A (1985) Trust as a Social Reality. Social Forces 63 (4):967-985

20. Paldam M (2000) Social Capital: One Or Many? Definition And Measurement. Journal of Economic Surveys 5:629-653
21. Jones GR, George JM (1998) The experience and evolution of trust: Implications for cooperation and teamwork. Academy of Management Review:531-546
22. Cook J, Wall T (1980) New work attitude measures of trust, organizational commitment and personal need non-fulfilment. Journal of Occupational Psychology 53 (1):39-52
23. McKnight DH, Cummings LL, Chervany NL (1998) Initial trust formation in new organizational relationships. Academy of Management Review:473-490
24. Johnson D, Grayson K (2005) Cognitive and affective trust in service relationships. Journal of Business research 58 (4):500-507
25. Renno L, Castro H (2008) Assesing the Validity and Reliability of Interpersonal Trust Measures in Cross-National Surveys. University of Brasilia, Brazil.
26. Glaeser EL, Laibson DI, Scheinkman JA, Soutter CL (2000) Measuring trust. The Quarterly Journal of Economics 115 (3):811-846
27. Naef M, Schupp J (2009) Measuring Trust: Experiments and Surveys in Contrast and Combination. IZA Discussion Paper Series. Institute for the Study of Labor (IZA), Bonn, Germany
28. ABS (2011) General Social Survey Summary Results Australia 2010.