Social Relationships in Recommender Systems

Ofer Arazy, University of Alberta, ofer.arazy@ualberta.ca
Ibrahim Sana, Ben-Gurion University, elsanae@bgu.ac.il
Bracha Shapira, Ben-Gurion University, bshapira@bgu.ac.il
Nanda Kumar, City University of New York, Nanda_Kumar@baruch.cuny.edu

Abstract

The current industry standard for recommender system uses variants of collaborative filtering (CF), where recipient-source relationships are determined by the extent to which the recipient and source share interests. This research attempts to improve the performance of these CF recommender systems by identifying additional measures of relationship indicators based on theories from communication and marketing. We developed a social filtering model that incorporates these various social measures (e.g. Trust, Reputation, Interaction Frequency, and Relationship Duration), and conducted an empirical study to test the model. The results from the study show small, but significant, improvements for various social relationships. We plan to build on these preliminary results to further consolidate our research on using social relationship in recommender systems.

1. Introduction

Recommendation systems provide users with personalized information relevant to their specific interests, and are now an integral part of firms' information systems, serving both customers and internal employees. The collaborative filtering (CF) approach for recommendation systems emerged in the mid 1990s, and have since become an industry standard. Its greatest advantage over content-based systems is that explicit content descriptions are not required. A typical CF process consists of the three sequential steps: (1) identification of relevant sources for a user, (2) summarization of recommendations provided by these relevant sources (i.e. their evaluations of information items), and (3) generation of recommendations for the user, based on the evaluations provided by the relevant sources (Konstan 2004).

The initial identification of relevant sources is critical in assuring effective CF system performance. CF systems employ users' similarity of tastes to determine relevant sources (Shardanand & Maes 1995). However, additional indicators of social relationships are available online (e.g., 'trust' in ePinions.com), and these indicators could potentially be utilized to select relevant sources. We refer to recommender systems that incorporate additional indicators of social relationships – beyond the taste-similarity traditionally used in CF - as *social recommender system*. Research on social recommender systems is in its infancy, and to date studies have focused on one type of social relation – Trust. However, behavioral theory suggests that additional types of social relations impact recipient's advice taking, e.g. interaction frequency and reputation. Thus, the goal of this paper is to identify the types of social relations that are most valuable for social recommender systems.

2. Related Works

Using the past literature in marketing, applied psychology and organization, we identified four salient constructs relevant to the advice taking context: cognitive similarity (Gilly et al. 1998), tie strength (measured in terms of relationship duration and interaction frequency; Levin & Cross 2004), trust (Levin & Cross 2004; Smith et al. 2005), and social capital (Gilly et al. 1998). However, traditionally in social recommender systems, a recipient is associated with sources solely based on their **cognitive similarity** (and specifically, similarity in tastes; i.e. CF; Shardanand & Maes 1995). The main advantage of this approach is that it requires little effort from users: users may need to rate items they've consumed, but are not required to explicitly define their relationships to others. Its limitation is that in cases where little information is available about users and items (referred to as 'cold start'), prediction accuracy suffers.

Recently, research has begun exploring how additional types of source-recipient information could be elicited in online settings, and incorporated onto recommender systems. **Trust** and friendship relationships could be gathered from online social networks. Alternatively, instead of harvesting data from online communities, users might be asked to explicitly define the extent to which they trust other users (Goldberg et al. 1992). The main limitation of this approach (besides the effort required from users) is that users have on average very few links, and thus data is insufficient to improve recommendation quality. This limitation could be alleviated by 'propagating' trust across relationships, i.e. if user A trusts B, and B trusts C, then we could assume that A trusts C (at least to some extent). Numerous **trust-propagation** algorithms have been proposed (Guha et al. 2004), and most trust-based recommender systems employ some variation of trust propagation. Explicit trust relations and trust propagation were incorporated into a recommender system by Massa & Avesani (2004) and Golbeck & Hendler (2006). Massa & Avesani (2004) conducted an empirical evaluation on Epinions.com dataset, comparing the social approach (i.e. a combination of trust and CF) against standard CF. The result shows that without propagation, trust relations provide 0.7% improvements over CF, and when using propagation, enhancements are up to 4.5% (no statistical significance described). Golbeck & Hendler (2006) report that overall trust-based recommendation is *not* more accurate than traditional CF or Naïve prediction (i.e. prediction made based on the simple average of all users' ratings), although in some cases - specifically, when user's ratings differ from the item's common rating - social relations can provide accurate recommendation.

An alternative relationship indicator, source's **social capital** (or reputation), could be extracted in online settings using reputation systems, where a system records users' ratings on others' recommendations, and accumulates these ratings to calculate 'reputation' scores for recommenders. Early *reputation systems* were used for e-commerce (e.g., eBay), and recently this approach was adopted in recommender systems, presenting users' reputation score next to their recommendations (e.g., Amazon.com). Potentially, this approach could be extended to incorporate reputation score into the recommendation algorithm (Massa & Avesani 2004). However, to date there is no empirical evidence to demonstrate the benefit of reputation scores in recommender systems.

Tie-strength data could be extracted by tracking the frequency and duration of users' electronic communications (e.g., e-mail, text messaging), through a utility that is installed on users' computers. While this approach would require little effort from users, it may pose a risk to users' privacy, and to the best of our knowledge has not yet been implemented in recommender systems.

In summary, research on social recommendations systems is in its early phases, and to date there are no conclusive evidence for the benefit of social relations. CF systems often lack sufficient data to establish reliable recipient-source similarity relationship. Given that other types of social relationships (trust, friendship, interaction frequency and duration, and social capital) are correlated with taste similarity, we expect that including additional social relationship data would improve recommender systems accuracy. To date, no study provides a comparison of alternative social relations, and thus we are not sure which type of social relation is most beneficial for recommender systems.

3. Research Method

The objective of this study is to investigate the potential usefulness of social relationship information for enhancing the performance of recommender systems. Specifically, we study the effectiveness of alternative types of relational data, focusing on Reputation, Friendship, Trust, Relationship Duration, and Interaction Frequency. We choose to explore the research question in the context of movie recommendations, as this is an area where recommendation systems have been very popular (e.g. MovieLens, NetFlix). Specifically, the movie domain has been a popular test bed for studying social recommender systems (e.g. FilmTrust; Golbeck & Hendler 2006).

To study relationships between users, we recruited 99 participants from among undergraduate students of a large public university in Israel, and we applied strict measures to ensure the confidentiality and privacy of the participants. We asked students to rate movies (70 popular recent movies, as well as movies specified by individual users) resulting in a set of 240 unique movies and an average of 16.6

movie ratings per user. These ratings were used (a) to establish users' taste-similarity for the CF algorithm, and (b) to test system's predictions.

We implemented and compared different recommender algorithms - Naïve, CF, and Social - as follows. As a baseline, we tested a *Naïve algorithm*, where a recommendation is made based on all users' average rating for that item. The CF algorithm was implemented based on the standard user-user approach (Resnick et al. 1994; Herlocker et al. 2004), where taste-similarity was based on Pearson Correlation for user's rating vectors, and each recipient is associated with a set of close sources (using K-Nearest Neighbor algorithm).

In order to develop the social recommendation algorithm, we established various social relations between users. We ran a survey, where each student was asked to select three sources (within their cohort) from whom they would likely seek advice on choosing movies. The participants (i.e. recommendation recipients), then, rated their relationships with these sources (on a 7-Likert scale) on the various social dimensions described in Table 1.

Social Dimension	Measurement Item
Trust	I trust this person
Friendship	I would consider this person a friend
Interaction Frequency	How often did you communicate with this person
Relationship Duration	How long have you known this person
Social Capital	This person is reputable

Table1: Social dimensions and measurement items

To extend the number of social relations, we applied the MoleTrust propagation algorithm (Massa & Avesani 2005), with a propagation distance of 2, and decreasing weights assigned by distance from recipient. This procedure extended the average number of social links per user from 3 to 7.67[1].

We utilized social relations to enhance (rather than replace) traditional CF, and integrated taste-similarity (calculated based on movie ratings) and explicit relational information into one recipient-source similarity score (Massa & Avesani 2004). After standardization of scales, we used an integration procedure that gives preference to explicit social relations, as described in Figure 1[2]. Similarly to CF, each recipient was then associated with close sources using K-Nearest Neighbor algorithm.

A prediction for an item was based on the standard prediction formula (Resnick et al. 1994). Since, the formula can provide values outside the [0, 1] span, we've rounded these values to the [0, 1] range. For all recommender algorithms, if no sources were associated with a recipient, prediction was made based on the recipient's average rating (over all items he has rated). We used 80% of the data for training (i.e. establishing recipient-source taste similarity), and 20% for predictions, and made predictions for all users who've rated more than 5 items. In order to make sure results were robust, we used K-Fold cross validation (where K=10).

$$\text{Recipient-Source similarity} = \begin{cases} \text{Only explicit social relation exist} & \text{explicit social relation} \\ \text{Only taste-similarity exist} & \text{taste-similarity} \\ \text{Both explicit social relation and taste-similarity} & \text{explicit social relation} \end{cases}$$

Figure 1: procedure for integrating taste-similarity and explicit relations scores

[1] We also verified that propagation indeed improves performance.

[2] We tested this simple integration scheme and found it equal to other integration methods proposed in the literature.

Once the system made a prediction for an item (for each of the competing approaches: naïve, CF, and social), we compared it against the subject's actual rating. Performance was measured using the Mean Average Error (MAE; $MAE = \sum_{i=1}^{n} | P_i - R_i | \big/ N$, where P_i is the predicted rating for item i, R_i is the recipient's actual rating, and N is the number of predictions), which is the most appropriate measure for evaluating prediction accuracy in offline tests (Herlocker at al. 2004). In addition, we employed Precision, and Recall measures (where a 'hit' is considered when both the predicted and actual ratings are above 0.5)

4. Results

Table 2 below compares the performance of naïve prediction, standard CF, and various social schemes. When compared to the naïve approach, CF has higher MAE (2.6%) and Precision (0.3%), but lower Recall (5.1%).

Recommendation Approach	MAE	MAE improve over CF (sig.)	Precision	Recall
Naïve	0.2536		0.7398	0.8103
Traditional CF	0.2470		0.7422	0.7693
CF+ Trust	0.2413	2.3% ($P<0.000$)	0.7500	0.7796
CF+ Friendship	0.2416	2.2% ($P<0.000$)	0.7482	0.7777
CF+ Interaction Frequency	0.2414	2.3% ($P<0.000$)	0.7495	0.7827
CF+ Relationship Duration	0.2410	2.4% ($P<0.000$)	0.7493	0.7806
CF+ Reputation	0.2426	1.8% ($P<0.000$)	0.7481	0.7814

Table 2: Average MAE for the ten simulation runs

Social relations provide improvements over CF. MAE improvements are small, but statistically significant. Precision and Recall are slightly higher than those reported for CF (up to 1.1% and 1.7% respectively). In addition to accuracy enhancements, the social approaches improved coverage by 7.5%. I.e., using social relations allowed falling back on the default prediction (user's average rating) less times.

The differences in performance between the alternative social approaches are minimal. The highest MAE was attained with 'Relationship Duration', Precision is highest with 'Trust', and 'Interaction Frequency' provides the best Recall score. The MAE improvements for 'Relationship Duration' yielded consistent improvements over the ten simulation runs, as illustrated in Figure 2 below.

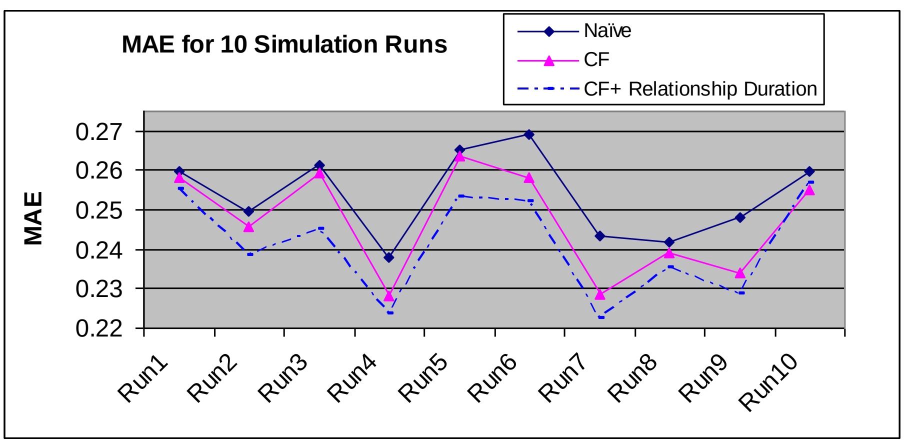

Figure 2: MAE over the ten simulation runs; naïve, CF, and 'Relationship Duration' approaches

In order to gain a better understanding of the situations where social relations are most useful, we performed two analyses. First, we classified users based on the number of movies they have rated. Contrary to the results reported in Massa & Avesani (2004), we found no relation between the number of movies rated and the benefit of social relations. Second, we analyzed the cases in which social relations provide accurate recommendation. Contrary to Golbeck & Hendler (2006) we found (statistically significant) positive relationship between social prediction accuracy and the accuracy based on the item's common rating. These findings suggest that we have not yet identified the salient factors that determine when social recommendations perform well (i.e. the factors identified to date do not generalize across settings), thus highlighting the need for additional research in the area.

5. Discussion and Conclusion

This study seeks to establish the benefit of social relations for enhancing recommender system performance. Consistent with Massa & Avesani (2004), we demonstrate that social relations could provide small improvements in accuracy over standard CF. To our knowledge, ours is the first study to show statistically significant improvements.

The comparison of the alternative types of social relations revealed small differences, where 'Relationship Duration' yielded the best MAE. This comparison provides the second contribution of this paper, and to our knowledge no prior study on recommender systems compared alternative types of social relations. When assessing the practicality of the proposed approach, one needs to also consider the effort involved in extracting this social information. Capturing communication frequency or duration could be calculated automatically (by installing a tracking utility on each users' computer), while establishing a social network (whether used to calculate trust, friendship, or reputation indicators) requires the users to explicate the relationships. Additional considerations include the technical difficulty of developing the software architecture and users' privacy concerns.

While our study shows small significant effects, we hope to leverage this preliminary study to further investigate the following avenues:

- Explore ways to enhance the social approach, specifically by considering alternative integration procedures (of CF and social similarities) and other trust propagation algorithms.
- Generalize the results by (a) conducting a larger-scale empirical study and (b) testing our approach on large movie databases (e.g. NetFlix).
- Repeat the empirical study and extract *actual* social relations data, rather than perceptions that were used in this study (e.g. interaction frequency).
- Extend the study to other user groups and recommendation domains.

In conclusion, recommender systems play a large role in the online world, by reducing information overload and enhancing shopping experience. The emergence of Web 2.0 tools (e.g. social networks, reputation mechanisms) – the theme of WITS'07 - presents an opportunity for recommender system, by providing relational data that could be utilized to associate a recipient with relevant sources. This study advances our understanding of the potential and limitation of this novel approach. Still, more research is warranted in order to realize the full potential of social relations in recommender systems.

Acknowledgement

This work was supported by Deutsche Telokom Co.

Bibliography

Gilly M., Graham J., Wolfinbarger M., and Yale L., 1998, A Dyadic Study of Interpersonal Information Search. *J. of the Academy of Marketing Science 26*, pp. 83-100.

Golbeck J. and Hendler J., 2006, FilmTrust, in *Proceedings of the IEEE Consumer Communications and Networking Conference*, pp. 282–286.

Goldberg D., Nichols D., Oki B., and Terry D., 1992, Using Collaborative filtering to Weave an Information Tapestry, Communications of the ACM, 35:12, pp. 61-70

Guha R, Kumar R., Prabhakar R., and Tomkins A., 2004, Propagation of Trust and Distrust, in *Proceedings of the 13th Conference on world wide Web (WWW2004)*.

Herlocker J., Konstan J., Terveen L., and Riedl J., 2004, Evaluating Collaborative Filtering Recommender Systems, ACM TOIS, 22:1, pp. 5–53.

Konstan J., 2004, Introduction to recommender systems: Algorithms and evaluation. *ACM Transactions on Information Systems*, 22:1, pp.1–4.

Levin D., and Cross R., 2004, The Strength of Weak Ties You Can Trust, *Management Science*, 40:11, pp. 1477-1490.

Massa P. and Avesani P., 2004, Trust-Aware Collaborative Filtering for Recommender Systems, in *Proceedings of the Int. Conf. on Cooperative Information Systems*.

Massa P. and Avesani P., 2005, Controversial users demand local trust metrics, in Proc. of AAAI-05, pp. 121–126.

Resnick P., Iacovou N., Suchak M., Bergstorm P., and Riedl J., 1994, GroupLens: An open architecture for collaborative filtering of netnews, in *Proceedings of the ACM Conference on Computer-Supported Cooperative Work*, pp, 175– 186.

Shardanand U. and Maes P., 1995, Social Information Filtering: Algorithms for Automating "Word of Mouth, in *Proceedings of the SIGCHI*, pp. 210-217.

Smith, D., Menon, S. and Sivakumar K., 2005, Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of Interactive Marketing*, 19:3, pp. 15-37.