Towards Understanding the Challenges Facing Effective Trust-Aware Recommendation

Yue Shi, Martha Larson, Alan Hanjalic
Multimedia Information Retrieval Lab
Delft University of Technology, Delft, The Netherlands
{y.shi, m.a.larson, a.hanjalic}@tudelft.nl

ABSTRACT

We introduce a method for generating semi-synthetic social data collections, which we use to study trust-aware recommendation. Specifically, we examine the effects of social graph degree distribution on user-based collaborative filtering that substitutes trusted users for conventional neighbors. Our semi-synthetic data collections are created via a naïve pruning process that maps a user-item matrix onto various social graphs with the degree distributions of real-world Web-based social systems. Our goal is to extend our understanding of the challenges facing effective trust-aware recommendation beyond the current possibilities, which are limited by data set availability. The improvement offered by trust-aware recommendation is shown to have substantial dependence on the degree distribution of the social graph.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *Information Filtering*

General Terms

Algorithms, Performance, Experimentation

Keywords

Recommender systems, social trust networks, trust-aware recommendation, collaborative filtering, Kronecker graph

1. INTRODUCTION

Trust-aware recommendation, i.e., approaches that make use of the information in a social graph, has recently been receiving increased research attention [6][8][12][16][17][18][21][25]. One aim of this research is to exploit the potential benefits that social trust information can bring to conventional collaborative filtering (CF). User-based CF is a highly successful technique based on the concept that an item should be recommended to a user if users with similar preferences like it [1][10]. Under conventional user-based CF approaches, the similarity of preference between users is derived from the user-item matrix, and a similarity measure is used to calculate the match between two users' profiles. The closest matches form the user's neighborhood and the profiles of users in the neighborhood provide the basis for recommendation. The large number of user-user comparisons necessary to determine an appropriate neighborhood is the major source of computation expense for conventional CF and also makes it unsuitable for distributed settings. Social trust information has the potential to address these issues [23]. If trusted users replace the neighbors used by conventional user-based CF, reliable recommendation is possible without the need to compute similarity neighborhoods.

The goal of the research presented in this paper is to improve our understanding of how particular properties of social data collections impact the benefit that social trust information contributes to CF performance. In particular, we investigated how the degree distribution of a social graph impacts user-based CF that substitutes trusted users for conventional neighborhoods. We consider social trust to include any connection that one user explicitly establishes with another, regardless of the user's motivation for creating the social connection. Defined in this way, social trust relationships include "friends" on YouTube, "people" on Delicious, and "trusted reviewers" on Epinions. We define a social data collection to be a set of user profiles containing item ratings or comparable item information (i.e., a user-item matrix) together with social trust information, consisting of the trust set of each user. All other users to which a given user has explicitly established links form that user's trust set. Taken together, the trust sets of the users in a collection define a user graph (i.e., a social graph), in which the nodes are the users and the edges are the trust connections. The degree of a node is the number of members in the corresponding user's trust set.

The starting point of our investigation is recent work in social network modeling, which has revealed that social networks as they exist in Web-based systems are characterized by certain properties [4]. In particular, the degree distributions of nodes follow a power-law with a characteristic slope. We must necessarily assume that in real-world social data systems these characteristic distributions hold. In this paper we address the following question: *How does the degree distribution known to characterize real-world Web-based social systems impact trust-aware recommendation?* Our investigation consists of experiments on a series of semi-synthetic social data collections that are generated by mapping existing user-item matrices to social graphs with the degree distributions of known Web-based social systems. By using semi-synthetic graphs, we are able to experiment with multiple social data collections. In contrast, current studies of trust-aware recommendation limit themselves to a single data set. In most cases this set is derived from Epinions.com, e.g., [12][16][17][18][21], although a few exceptions exist, e.g., [6][11]. Our goal is to arrive at a better understanding of the potentials and limitations of trust-aware recommendation. Our study makes use of currently available public resources, but extends our present understanding of trust-aware recommendation by moving beyond the analysis of a single data set. Our work makes two key contributions: first, a method for generating semi-simulated social data collections that exploits recent advances in social network modeling [13] and, second, our finding that the extremely long tail that characterizes degree distributions of real-world Web-based social systems has serious consequences for trust-aware recommendation approaches.

In the next section, we cover related work. Then we present our method for generating semi-synthetic social data collections and the results of our experimental analysis of three such data collections. The last section presents a discussion and conclusion.

RSWeb'10, September 26, 2010, Barcelona, Spain

2. RELATED WORK

2.1 Social Trust and Recommendation

2.1.1 The Nature of Social Trust

The concept of trust is used differently in different contexts. Within computer science, trust can be a reflection of the quality of a peer in a peer-to-peer system or the reliability of an information source (cf. [7]). For the purpose of recommendation, trust is understood to be a relationship between users and the type of friends to whom they would turn in real life for recommendations [16]. A full understanding of how users assign trust relationships in social collections is not required in order for social trust information to aid recommendation. Minimally, social trust must only serve to capture shared tastes between users. As pointed out in [8], trust-based recommender systems generally assume a relationship between trust and similarity for this purpose. For instance, in the work that has been devoted to empirically investigating the correlation between social trust and interest similarity, a correlation was found for books [23] and then extended to books and film [24]. In [23], support was found for the hypothesis about a correlation between trust and user similarity when the trust network is closely associated with a particular application. In this work, we assume that connections between users in the user graph reflect a match in user taste as least as well as the user-user similarity calculations used in conventional user-based collaborative filtering.

2.1.2 Trust-aware recommendation

Various techniques have been proposed to improve CF by integration of social trust information. Closest to the work presented in this paper are approaches that involve replacing the neighbor set on which predictions are based in conventional user-based CF with a trust set of users found via the social network. In [17], an algorithm called MoleTrust carries out a depth-first walk of the trust graph. When the propagation horizon is set to one (i.e., only immediate neighbors are used) MoleTrust outperforms CF for rating prediction. Increasing the propagation horizon is shown to increase the coverage, but at the cost of prediction accuracy. Trust information becomes noisier and less useful as trust propagation moves beyond the first degree, i.e., draws on information outside of the users trust set. FilmTrust [6] uses a trust-flow-based method called TidalTrust and excels in cases where a user has an opinion on an item that diverges from average. In [11], social annotation and the social network are integrated in an approach that uses a random walk with restarts. TrustWalker presents a random walk model to combine traditional item-based collaborative recommendation and trust-based recommendation [12], which achieves promising improvement with respect to accuracy and coverage.

Another empirical study on trust-aware recommendation investigates the influence on recommendation performance in terms of both an individual social friend and a community-like ally in the social network [25], which also indicates the efficiency of recommendation based on socially selected users compared to traditional user-based CF. In [21], the question of whom users should trust is examined with particular attention to the role of key users. Ma et al. [16] have proposed a matrix factorization framework to fuse the user-item rating matrix and user social trust network. The framework makes use of the social network as a constraint for learning user latent features rather than basing the factorization solely on the user-item rating matrix.

Different approaches to social trust in recommendation aim to exploit different advantages. As mentioned above, our work is aimed at reducing computational complexity and making conventional CF more suitable for use in a decentralized system. Here, we mention some additional aspects. In view of the fact that CF suffers from data sparseness, which leads to noise since the similarity measure must be frequently calculated for users with little profile overlap, [18] aims to use social trust to mitigate this effect. In [8], the point is made that trust might be able to capture something beyond overall similarity between user profiles. A trust network could be exploited to extend the coverage of conventional CF [17]. Since users pick their own trustees, trust-based approaches could be more robust to attack than conventional CF [17]. Finally, trust-based recommendation systems may have the advantage over conventional systems because of greater transparency of the source of the recommendation [3]. Our work could possibly aid the development of trust-aware recommendation approaches that fulfill other potentials of social trust, a topic we leave for future investigation.

2.2 Social Network Modeling

Real-world social networks have been shown to be characterized by a specific set of properties including power-law degree distributions [4][20], small diameter [19][22], shrinking diameter and densification power law [15]. Those properties are used to guide the modeling of social networks, which has resulted in a number of network models, e.g., preferential attachment model [2][5] and the small-world model [22]. The recently introduced Kronecker graph model has been demonstrated to be capable of capturing multiple real social network properties simultaneously [13][14]. We choose this model to synthesize the social graphs that we use to create our semi-synthetic social data collections.

3. EXPERIMENTAL FRAMEWORK

3.1 Semi-synthetic social data collections

We generated semi-synthetic data collections using a naïve pruning process that maps an existing collection of user profiles (i.e., a user-item matrix) onto a social graph that has been generated such that its degree distribution corresponds to that of a real-world Web-based social system.

3.1.1 Kronecker graph model

In order to synthesize social graphs with real-world properties, we adopt the stochastic Kronecker graph (KG) model [13]. If we assume that P_1 is a $N_1 \times N_1$ Kronecker initiator matrix containing values $\theta_{ij} \in P_1$ denoting the probability of the existence of an edge, then, a social network with number of nodes N_1^k, denoted as P_k, can be synthesized by the kth Kronecker product of P_1, as shown below:

$$P_k = P_1^{[k]} = \underbrace{P_1 \otimes P_1 \otimes \ldots \otimes P_1}_{k\ times} \qquad (1)$$

Finally, a graph can be obtained by sampling an instance from the probability distribution defined by P_k. We adopt the 2×2 initiator matrices empirically obtained in [13] in order to simulate social networks, i.e., Epinions, Delicious and Flickr. Notice that although the KG model is designed to simulate a large range of properties of social networks, here, we only investigate degree distributions. We use a representation of the synthetic graph that consists of a list of nodes ordered by degree from large to small.

3.1.2 Idealized trust sets

Our naïve mapping from an existing user user-item matrix to the synthetic social graph builds on the basic assumption, mentioned above, that the trust set of a user reflects that user's interests at least as well as the set of neighbors computed via similarity under conventional user-based CF. Building on this assumption, we create an idealized trust set for each user consisting of all users more similar than a threshold, θ. User-user similarity is calculated using the Pearson correlation [10], conventional user-based CF:

$$sim(u,v) = \frac{\sum_{i \in C_{uv}} (r_{ui} - \overline{r_u})(r_{vi} - \overline{r_v})}{\sqrt{\sum_{i \in C_{uv}} (r_{ui} - \overline{r_u})^2 \sum_{i \in C_{uv}} (r_{vi} - \overline{r_v})^2}} \quad (2)$$

Here, u and v are two users represented by standard profile vectors consisting of the ratings that they have assigned to items in the collection. C_{uv} denotes the intersection of items rated by both user u and v and r_{ui} denotes the user u's rating on item i. The average rating value of the user u across all co-rated items in C_{uv} is denoted by $\overline{r_u}$.

3.1.3 Social pruning

Our semi-synthetic social data collection is then generated by a process of social pruning that results in a *socially pruned user graph* (SPUG) is a user graph in which connections have been removed in order to make its degree distribution respect that of the KG model, which represents a real-world social system. The mapping process between users in the idealized user graph and nodes in the synthesized social graph is straightforward. The nodes in the synthetic social network have been ranked by order of degree (i.e., number of connections). Also, the users in the data collection (i.e., the original user-item matrix) have been ranked by the number of neighbors. We then associate the most highly connected node with user with the most neighbors and move down the list, pairing users and nodes in order. In each case, we prune the users from the idealized trust set so that it matches the degree of the associated node in the synthetic social network. A judicious choice of θ ensures that in a majority of cases, the degree distribution of the idealized user graph can be forced into the degree distribution of the target Web-based social system only by removing links. The pruning process is facilitated by our choice of mapping the most highly connected users to most highly connected nodes. In theory, any mapping between users and nodes would be adequate for our purposes. In the very rare case that a link must be added, we do so by adding a random connection. The resulting SPUG follows exactly the same degree distribution as the network that the KG model simulates, which means that the semi-synthetic social data collection shares the property of degree distribution with a real-word Web-based social system.

3.2 Top-N recommendation

We experiment on the task of Top-N recommendation, i.e., produce a list of items for the user ranked in order of user preference.

3.2.1 User-based CF with Trust Sets

For the purpose of ranking items we use the Trust Set Average Rating (TrustSetAvRat). The TrustSetAvRat rating score for user u on item q is calculated as:

$$r_{uq} = \frac{\sum_{v \in T(u)} t_{uv} r_{vq}}{\sum_{v \in T(u)} t_{uv} I_{vq}} \quad (3)$$

where $T(u)$ denotes the trust set of user u and t_{uv} is a weight indicting how much user u trusts v. TrustSetAvRat is equivalent to trust-based recommendation in [6] and MT1 in [17]. Here, we only use binary trust, i.e., t_{uv}=1 if u trusts v, 0 otherwise. I_{vq} is an indication function where I_{vq}=1 if user v rated item q, 0 otherwise.

3.2.2 Experimental baselines

As a naïve baseline, we use ItemAvRat, the collection wide average rating assigned to an item. Under this approach, for every user the recommended list contains the Top-N items collection-wide, ranked in order of overall popularity. As a second, more sophisticated baseline, we use the idealized trust sets calculated using user-user similarity in Eq. 2. This baseline, which we designate IdTrustSet, is comparable to conventional user-based CF. As can be observed in Eq. 3., IdTrustSet lacks the factor weighting the contributions of the individual users by their similarities with the target user used in conventional user-based CF. Since we are interested in a comparison with the extremely long tails of characteristics of the degree distributions of real-world social graphs, we choose θ such that the degree distribution of the resulting idealized trust sets approaches a power law distribution.

3.3 Data sets

In our experiments, we used the publicly available Jester (JS) data collection (rating scale -10–10) [9]. Our approach constitutes an innovative new use of these data. Using the latest JS data set (which contains a total of 64K users and 150 items), we create two subsets: JS1, which contains 1024 (2^{10}) users and JS2, which contains 4096 (2^{12}) users. We select only users that have rated at least 20 of the 150 items. Note that since the node count of a graph synthesized by the KG model is always a power of two, our choice of set size ensures an exact match during pruning.

3.4 Experimental Protocol

We adopt the widely used 5-fold cross-validation protocol in our experiments. In each fold we use 80% ratings of each user for training and remaining 20% ratings for testing. For Top-N recommendation, we take items with rating equal or higher than 7.5 to be relevant items. The training set is used to generate the predictions on all other items, which then can be compared with ground truth in the test set for evaluation. We report results using the Mean Reciprocal Rank (MRR), the inverse position of the topmost relevant document in the Top-N recommendation list.

4. EVALUATION

We create socially pruned user graphs by pruning the JS1 and JS2 data sets to respect the degree distributions of the synthetic graphs using the process described in Section 3.1.3. We create three semi-synthetic data sets with degree distributions corresponding to those of Epinions, Delicious and Flickr. Our implementation of Kronecker graph modeling is based on the publicly available software SNAP (http://snap.stanford.edu/snap/index.html).

The data sets have a dramatically large number of users with very small trust sets (< 10), differing dramatically from idealized trust sets. In Table 1, CF performance in the case of socially pruned user graphs (indicated as SPUG) can be seen to be markedly lower than in the case of either IdTrustSet (i.e., the baseline comparable to conventional user-based collaborative filtering) or ItemAvRat, the naïve baseline.

Table 1. MRR performance comparison between the baselines and JS1 and JS2 data sets, the socially pruned data sets

	ItemAvRat	**IdTrustSet (~ user-based CF)**	**Epinions SPUG**	**Delicious SPUG**	**Flickr SPUG**
JS1	0.279	0.345	0.130	0.142	0.126
JS2	0.180	0.263	0.144	0.161	0.146

We conclude that the very long tail observed in real-world Web-based systems has serious consequences for trust-aware recommendation. When the user-graph is forced into the degree distributions known to exist in the wild, CF performance is not longer able to approach the simple baseline.

5. DISCUSSION AND CONCLUSIONS

This paper has made a novel contribution to the current understanding of the challenges facing the effective use of social trust information for collaborative filtering. Our goal has been to transcend some of the limitations affecting current research on trust-aware recommendation and to shed light on a specific area of dependency between a property of social data collections (degree distribution of the social graph) and the performance of trust-aware recommendation. Our contributions are twofold. First, we proposed a method for producing semi-synthetic social data collections from existing data collections (i.e., user-item matrices). Second, our experiments provide evidence that the very long tail characterizing the degree distributions of real-world social graphs can have a devastating impact on the performance of CF.

Our results suggest that if a Web-based community is intended to support trust-aware recommendation, it should be planned in such a way that it will foster the emergence of patterns of social relationships among users that are advantageous for trust-aware recommendation. Suitable incentives should be provided to users so that user graphs are encouraged to develop, contrary to their natural tendencies, to have a degree distribution with a relatively shorter long tail. Note that we do not recommend forcing users into certain behavior, but rather an approach that would simply encourage users who set up few connections to set up more by making clear to them the potential benefits. Our exploratory experiments with idealized trust sets suggested that a long-tail distribution does not necessarily stand in the way of effective trust-aware recommendation, rather it is the *very* long tail that is damaging. Recall that our technique for generating semi-synthetic social data collections built on the assumption that in the real-world users' trust sets reflect user interest at least as well as the similarity neighborhoods used in conventional user-based collaborative filtering. The picture brightens, if, in a particular social data collection, users connect to each in a way that results in trust sets providing a better basis for recommendation than similarity neighborhoods. However, in a social data collection in which users choose the "wrong friends", the negative impact of very long-tail degree distributions could be exacerbated to the point of making trust-aware recommendation useless. Web-based communities are well advised to provide incentives to users to chose the "right friends."

Our future research will involve developing techniques to synthesize graphs that model second-degree relationships among users in order to test recommendation approaches that exploit trust propagation.

6. ACKNOWLEDGEMENTS

The research leading to these results was carried out within the PetaMedia Network of Excellence and has received funding from the European Commission's 7th Framework Program under grant agreement n° 216444.

7. REFERENCES

[1] Adomavicius G., and Tuzhilin, A., 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE TKDE, 17, 6, 734-749.

[2] Barabasi A. L., and Albert R., 1999. Emergence of scaling in random networks. Science, 286, 509-512.

[3] Bonhard, P., Sasse, M., July 2006. 'Knowing me, knowing you' — using profiles and social networking to improve recommender systems. BT Technology Journal 24, 3, 84-98.

[4] Clauset, A., Shalizi, C. R., and Newman M., 2009. Power-law distributions in empirical data. SIAM Review, 51, 661-703.

[5] Flaxman A. D., Frieze A. M., and Vera J., 2007. A geometric preferential attachment model of networks II. WAW '07, 41-55.

[6] Golbeck, J., 2006. Generating predictive movie recommendations from trust in social networks. iTrust '06, 93-104.

[7] Golbeck, J. 2006. Trust on the world wide web: a survey. *Found. Trends Web Sci.* 1, 2, 131-197.

[8] Golbeck, J., 2009. Trust and nuanced profile similarity in online social networks. ACM Trans. Web, 3, 4, 1-33.

[9] Goldberg, K., Roeder, T., Gupta, D., Perkins, C., July 2001. Eigentaste: A constant time collaborative filtering algorithm. Information Retrieval 4, 2, 133-151.

[10] Herlocker, J., Konstan, J., Borchers, A., and Riedl, J., 1999. An algorithmic framework for performing collaborative filtering. SIGIR '99, 230-237.

[11] Konstas, I., Stathopoulos, V., and Jose J. M., 2009. On social network and collaborative recommendation. SIGIR '09, 195-202.

[12] Jamali, M., and Ester, M., 2009. TrustWalker: a random walk model for combining trust-based and item-based recommendation. KDD '09, 397-406.

[13] Leskovec, J., Chakrabarti, D., Kleinberg, J., Faloutsos, C., and Ghahramani, Z., 2010. Kronecker Graphs: An approach to modeling networks. JMLR, 11(Feb):985-1042.

[14] Leskovec, J., and Faloutsos, C., 2007. Scalable Modeling of Real Graphs using Kronecker Multiplication. ICML '07, 497-504.

[15] Leskovec, J., Kleinberg, J., and Faloutsos, C., 2005. Graphs over time: densification laws, shrinking diameters and possible explanations. KDD '05, 177-187.

[16] Ma, H., King, I., and Lyu, M. R. 2009. Learning to recommend with social trust ensemble. SIGIR '09, 203-210.

[17] Massa, P. and Avesani, P., 2007. Trust-aware recommender systems. RecSys '07, 17–24.

[18] Massa, P., and Bhattacharjee, B., Using trust in recommender systems: An experimental analysis. iTrust '04, 221-235.

[19] Milgram, S., 1967. The small-world problem. Psychology Today, 2, 60-67.

[20] Newman, M., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics, 46, 5, 323-351.

[21] Victor, P., Cornelis, C., De Cock, M., and Teredesai, A. M. 2008. Key figure impact in trust-enhanced recommender systems. *AI Commun.* 21, 2-3, 127-143

[22] Watts, D. J., and Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. Nature, 393, 440-442.

[23] Ziegler, C.-N., Lausen, G., 2004. Analyzing correlation between trust and user similarity in online communities. *Trust Management*, 251-265.

[24] Ziegler, C. and Golbeck, J. 2007. Investigating interactions of trust and interest similarity. *Decis. Support Syst.* 43, 2, 460-475.

[25] Zheng, R., Wilkinson, D., and Provost, F., 2008. Social network collaborative filtering. Working paper CeDER-8-08, Center for Digital Economy Research, New York University.