Exploiting Social Networks in Recommendation: a Multi-Domain Comparison

Alejandro Bellogín[a,b], Iván Cantador[a], Pablo Castells[a], Fernando Díez[a]
[a]Information Retrieval Group, Department of Computer Science, Universidad Autónoma de Madrid
[b]Information Access, Centrum Wiskunde & Informatica
{alejandro.bellogin, ivan.cantador, pablo.castells, fernando.diez }@uam.es[a],
alejandro.bellogin@cwi.nl[b]

ABSTRACT

Recommender Systems aim at automatically finding the most useful products or services for a particular user, providing a personalised list of items according to different input and attributes of users and items. State-of-the-art recommender systems are usually based on ratings and implicit feedback given by users about the items. Recently, due to the large number of social systems appearing in the so called Web 2.0, where friendship relations between people are explicit, *social contexts* exploitation has started to receive significant interest. In particular, social recommenders have started to be investigated that exploit social links between users in a community to suggest interesting items. In this paper we compare a series of experiments developed in recent years with different datasets where standard collaborative and social filtering techniques were analysed. We show that social filtering techniques achieve very high performance in the three domains discussed (bookmarks, music, and movies), although they may have lower coverage than traditional collaborative filtering algorithms.

Categories and Subject Descriptors

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

General Terms

Algorithms, Experimentation, Performance

Keywords

Recommender systems, Social Networks, Evaluation

1. INTRODUCTION

With the advent of the Social Web, a variety of new recommendation approaches have been proposed in the literature [1]. Most of these approaches are based on the exploitation of social tagging information and explicit friendship relations between users (social filtering recommenders) [5, 8]. Commonly, algorithms dealing with social context attempt to exploit the social connections of an active user. For example, Shepitsen et al. [10] employs a personalisation algorithm for recommendation in folksonomies that relies on hierarchical tag clusters, which are used to recommend the most similar items to the user's closest cluster, by using the cosine similarity measure. Other works focus on graph-based techniques for finding the most relevant items for a particular user, inspired by algorithms from quite different areas, successfully bringing them to social recommendation [6].

In this paper, we compare the performance of social filtering methods with standard collaborative filtering (CF) baselines using four different datasets on three domains (bookmarks, music, and movies). With this goal in mind, in the next section we present the methods evaluated in this paper, then, in Section 3 we discuss the datasets used. After that, in Section 4 we present the results obtained.

2. SOCIAL FILTERING RECOMMENDERS

Inspired by the approach presented in Liu & Lee [8], we analyse a pure social recommender that incorporates social information into the user-based CF model, named as **friends-based** (FB). Standard user-based CF typically computes predictions by performing a weighted sum over a set of similar users (usually called neighbours) as follows [1]: $s(u,i) = C \sum_{v \in N(u)} sim(u,v) r(v,i)$, where $r(v,i)$ denotes the rating given by user v to item i, and $sim(u,v)$ is the similarity between the two users. In this context, FB makes use of the same formula as the user-based CF technique, but replaces the set of nearest neighbours ($N(u)$) with the active user's (explicit) friends.

In [3] we propose a **social popularity** recommender (SocPop), where the algorithm suggests those items that are more popular among the set of the active user's friends. A third social recommender is evaluated where explicit distances between users in the social graph are integrated in the prediction formula: $s(u,i) = \sum_{v \in X(u,L)} K^{-d(u,v)} r(v,i)$. This approach was originally proposed in [5] and named as **personal-social** (PerSoc), where the authors use the Breadth-First Search algorithm in order to build a social tree for each user (denoted as $X(u,L)$), where L is the maximum number of levels taken into consideration in the algorithm, and K is an attenuation coefficient of the social network that determines the extent of the effect of distance $d(u,v)$ (we use Dijkstra's algorithm, $K = 2$ and $L = 6$).

Besides these pure social recommenders, hybrid social recommenders are useful not only for exploiting the social context of a user, but for providing higher coverage in extreme situations (such as the social or rating cold start, where no social context or ratings are available for a particular user). In this paper we analyse the performance of a com-

bination between the friends-based method described above and the classic user-based CF method, where all the active user's friends along with the set of most similar nearest neighbours are used to produce recommendations. We name this method **user-and-friends-based** (UFB). Alternatively, more complex hybrid recommenders can be defined based on random walks [6] and linear combinations of the predictions from several recommenders [4], but we leave the comparison of these methods across several domains as future work (some initial insights can be found in [3]).

3. A MULTI-DOMAIN PERSPECTIVE

We report results using four different datasets on three domains. The first one was gathered from the social music website *Last.fm*. As described in [2], we built our dataset aiming to obtain a representative set of users, covering all music genres, and forming a dense social network. This dataset contains 1.9K users, 17.6K artists (17.0K of them tagged), 186.5K tag assignments (98.6 per user), and 25.4K friend relations (13.4 per user).

The second dataset was obtained from *Delicious*, a social bookmarking site for Web pages. Also described in [2], we built this dataset with the same goal in mind as the one stated for Last.fm dataset: to cover a broad range of document's topics, and obtain a dense social network. In this case, the dataset contains 1.9K users, 69.2K bookmarked Web pages, 437.6K tag assignments, and 15.3K friend relations. On average, each user profile has 56.1 bookmarks, 234.4 tag assignments, and 8.2 friends.

The third dataset used was provided in the social track of the CAMRa Challenge [9]. This dataset was gathered by the Filmtipset community, and contains social links between users, movie ratings, movie comments, and other attributes of users and movies. However, in such dataset every test user has a social network, which is not a realistic scenario, since in many social media applications such as Delicious or Last.fm the social network coverage is only partial. Because of this, we create a fourth dataset where we incorporate a number of users with no friends in the new test set used in our experiments, more specifically, such number corresponds to the number of test users contained in the original test set (439 users). We denote the former dataset as CAMRa-Social (CAMRa-S) and the latter as CAMRa-Collaborative (CAMRa-C).

4. PERFORMANCE COMPARISON

Table 1 shows the performance results of the four social filtering recommenders presented before on the four datasets already described. We also use a standard user-based CF method with 15 neighbours and Pearson's similarity [1] (UB) and a matrix factorisation approach in which the rating matrix is factorised into 50 dimensions [7] (MF) as baselines.

We observe that the best performing approach is the PerSoc strategy, which adapts the well-known CF formula by weighting the similarity between the user's and her neighbours' rating-based profiles with the users' distances in the social graph. These results thus provide empiric evidence that combining CF and social networking information produces better recommendations than CF alone. Very interestingly, the FB strategy, which recommends items liked by explicit friends, obtains acceptable precision values. As concluded by Konstas and colleagues [6] for Last.fm, recommendations generated from the users' social networks represent a good alternative to rating-based methods; here, we extend such conclusion to other domains like bookmarks and movies. Merging this strategy with CF (UFB), nonetheless, does not improve the results obtained by the approaches separately except in the movie domain, where the CF algorithm shows better performance than in the other contexts.

Table 1: Obtained performance values for different datasets (reported metric is P@10). Best value for each dataset in bold.

Method	Last.fm	Delicious	CAMRa-S	CAMRa-C
UB	0.009	0.008	0.072	0.052
MF	0.025	0.003	0.038	0.026
FB	0.043	0.023	0.057	0.050
SocPop	0.021	0.011	0.001	0.001
PerSoc	**0.085**	**0.054**	**0.344**	**0.342**
UFB	0.014	0.008	0.077	0.053

Additionally, when considering alternative evaluation metrics, we found in [2] that social filtering methods have lower coverage and novelty than traditional CF and content-based recommenders; however, their diversity is higher, as measured using α-nDCG. These negative aspects could be improved by building hybrid recommenders, where the performance accuracy is slightly degraded at the expenses of better coverage and novelty [3, 2].

5. ACKNOWLEDGMENTS

This work is supported by the Spanish Government (TIN2011-28538-C02-01) and the Government of Madrid (S2009TIC-1542).

6. REFERENCES

[1] Adomavicius, G., and Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. Knowl. Data Eng. 17*, 6 (2005), 734–749.

[2] Bellogín, A., Cantador, I., and Castells, P. A comparative study of heterogeneous item recommendations in social systems. *Inf. Sci. 221* (2013), 142–169.

[3] Bellogín, A., Cantador, I., Díez, F., Castells, P., and Chavarriaga, E. An empirical comparison of social, collaborative filtering, and hybrid recommenders. *ACM TIST 4*, 1 (2013), 14.

[4] Bellogín, A., Castells, P., and Cantador, I. Self-adjusting hybrid recommenders based on social network analysis. In *SIGIR* (2011), W.-Y. Ma, J.-Y. Nie, R. A. Baeza-Yates, T.-S. Chua, and W. B. Croft, Eds., ACM, pp. 1147–1148.

[5] Ben-Shimon, D., Tsikinovsky, A., Rokach, L., Meisels, A., Shani, G., and Naamani, L. Recommender system from personal social networks. In *AWIC* (2007), K. Wegrzyn-Wolska and P. S. Szczepaniak, Eds., vol. 43 of *Advances in Soft Computing*, Springer, pp. 47–55.

[6] Konstas, I., Stathopoulos, V., and Jose, J. M. On social networks and collaborative recommendation. In *SIGIR* (2009), J. Allan, J. A. Aslam, M. Sanderson, C. Zhai, and J. Zobel, Eds., ACM, pp. 195–202.

[7] Koren, Y., Bell, R. M., and Volinsky, C. Matrix factorization techniques for recommender systems. *IEEE Computer 42*, 8 (2009), 30–37.

[8] Liu, F., and Lee, H. J. Use of social network information to enhance collaborative filtering performance. *Expert Syst. Appl. 37*, 7 (2010), 4772–4778.

[9] Said, A., Berkovsky, S., and Luca, E. W. D. Introduction to special section on camra2010: Movie recommendation in context. *ACM TIST 4*, 1 (2013), 13.

[10] Shepitsen, A., Gemmell, J., Mobasher, B., and Burke, R. D. Personalized recommendation in social tagging systems using hierarchical clustering. In *RecSys* (2008), P. Pu, D. G. Bridge, B. Mobasher, and F. Ricci, Eds., ACM, pp. 259–266.