Using Social and Pseudo-Social Networks for Improved Recommendation Quality

Alan Said, Ernesto W. De Luca, Sahin Albayrak
DAI-Lab
TU-Berlin
{alan.said,ernesto.deluca,sahin.albayrak}@dai-lab.de

Abstract

Recommender systems attempt to find relevant data for their users. As the body of data available in the Web sphere becomes larger, this task becomes increasingly harder. In this paper we present a comparison of recommendation results when using different social and pseudo-social features commonly available in online movie recommendation communities. Social relations, whether inferred or not, hold implicit information about users' taste and interests. We present results of a simple method that extends standard collaborative filtering algorithms to include a social network and show that this explicit and implicit information (i.e. direct friendship, and indirect co-commenting etc.) can be used to improve the quality of recommendations.

1 Introduction

Estimates say that the currently accumulated amount of data in the digital universe reached 1.2 zettabytes (1 billion terabytes) in 2010, which corresponds to a 50% increase during the two last years [Gantz and Reinsel, 2010]. A body of data of this size presents substantial challenges for current information retrieval systems. Independent of whether the task is search-, classification- or recommendation-oriented, processing and personalizing results from these systems becomes one of the most important tasks in order to identify relevant information. Granted, most systems do not face data amounts of this size, it is however implied that this accumulated amount is reflected in many websites which have seen considerable increase of users during the same time, e.g. Netflix [Siedler, 2010].

In personalized recommender systems, the de facto standard *Collaborative Filtering* (CF) approach, is becoming an insufficient means to produce relevant results due to the information overload which follows from the rapid data growth [Montebello, 1998]. However, the significant increase in data brings benefits as well, benefits in the form of *richer* meta data, i.e. more information related to every transaction, consumption, movie rating, etc. Using this rich data to extend regular collaborative filtering approaches can result in better information management systems, no matter if they are retrieval or recommendation based.

In movie recommendation systems, recommender systems research has mostly been focused on algorithmic approaches to better use the available data. The two most popular movie recommendation datasets, from the Netflix Prize[1] and the Movielens[2] community, do not include any social or pseudo-social structures. However, this data is commonly available in other online recommendation communities.

1.1 Problem Statement and Contribution

In this paper, we evaluate how different social and pseudo-social relations can be employed in order to improve the quality of recommendations in a movie scenario. Our model presents how user-item interaction can be used to infer relations between users. We present early stage results where these relations, no matter if inferred or explicit, increase the performance of our collaborative filtering-based movie recommender

The main contribution of this paper is the evaluation of different types of social networks in order to improve recommendation quality.

1.2 Outline

In this paper, we limit ourselves to the domain of movie recommendation, using a dataset from the Moviepilot[3] online movie recommendation community, and present a simple extension of standard collaborative filtering which uses regular and inferred social networks similar to the method presented by Guy et al. [Guy *et al.*, 2009].

Our approach infers ties between users based on their history of *comments*, whether they have stated they are *fans* of the same people, whether they have stated they *like* the same news articles, and if they have an explicitly stated *friendship* relation.

The experiments performed in this paper show that when using these networks, we can improve recommendation results compared to regular collaborative filtering. The full details of our approach are presented in Section 3.

[1] http://www.netflixprize.com/
[2] http://www.movielens.org/
[3] http://www.moviepilot.de

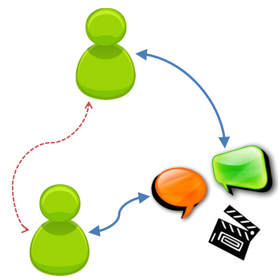

Figure 1: An inferred social tie (the red dotted line) is, in this case, created if two users have commented on the same movie, person or news article. The same principle is applied for users who are fans of the same actors, directors, or like the same news articles and comments.

2 Familiarity vs. Similarity

In standard collaborative filtering-based recommender systems, user similarities are calculated based on the user-movie relations (i.e. *similarity*), we use user-user relations in addition (i.e. *familiarity*). In our analysis and experiments we use a snapshot of the explicit friendship graph found in Moviepilot, as well as the implicit networks created when users interact with the same content (as shown in Figure 2), in order to improve the quality of our recommendations. The assumption is that a user's so-called *familiarity* networks hold implicit information about the user's so-called *similarity* (CF-based) network [Said *et al.*, 2010b]. We also present some statistical data on the dataset and its features.

3 Dataset and Experiments

Moviepilot is Germany's largest online movie recommendation community with more than one million users, over fifty thousand movies, and in excess of 10 million ratings.

3.1 The Dataset

Datasets provided by Moviepilot have been analyzed and researched previously [Said *et al.*, 2010a]. However, the dataset used in our evaluation differs from the ones used in prior publications. This dataset is a subset of the full, unfiltered data that creates the basis for the Moviepilot website. The dataset contains ratings by $10,000$ randomly selected users who have rated at least one movie. In addition to the ratings, the dataset also contains information on each user's friendship network within Moviepilot, as well as the comments posted by each user, the declarations of being a fan by each user (i.e. explicit statements saying a user is a fan of an actor, director, etc.) and the "diggs" of each user (i.e. users can "digg" different items such as comments, news articles etc.). The total number of ratings in our subset is $1,539,393$ spread over a period of four years (2006 to 2010). Table 1 shows the number of entities in the dataset and the approximate percentages of the full snapshot . The ratings are stored on a 0 to 100 scale with 0 being the lowest and 100 being the highest. The scale shown to the users is however 0.0 to 10.0. The networks used in this paper were either explicitly stated in the data (i.e. friendships) or were inferred from users' interactions with information available, i.e.:

Relation	Testset	%
Friendships	$3,764$	10%
Comments	$50,960$	30%
Fans	$170,092$	25%
Diggs	$25,259$	25%
Ratings	$1,539,393$	20%
Users$_{30+ratings}$	$10,000$	25%

Table 1: Dataset statistics for the snapshot we use and the (rough) percentage of the full dataset they represent. It should be noted that each Friendship relation is between two users, whereas each Comment-, Fan- and Digg-relation is a link between a user and the entity.

Type	Nodes	Edges
Friendship	$1,595$	$3,764$
Comments	$2,137$	$1,524,476$
Fans	$3,950$	$2,129,330$
Diggs	680	$20,028$

Table 2: The number of nodes and edges in every network. Similarly to the data in Table 1, each Friendship edge is between two users, whereas the other edges are between one user and the entity they interact with.

- the friendship graph - explicitly stated friendship relation between users
- the comments graph - an implicit network created when users comment on movies, actors, etc.
- the fan graph - an implicit network created when user are fans of the same people.
- the digg graph - an implicit network created when users "digg" the same news articles, comments, etc.

The sizes of the networks differ as the randomly selected users have diverse profiles, i.e. those with many friends and those with few, those who comment often and those who never comment, etc. The number of nodes and edges in each network is shown in Table 2, the number of ratings assigned by users in each of the networks is shown in Table 3.

3.2 Experimental Setup

For the experiments, 50 training and evaluation sets each for all networks were created. The evaluation sets consisted of

Type	Number	%
Friendship	$584,578$	38%
Comments	$697,012$	45%
Fans	$1,188,051$	77%
Diggs	$439,268$	29%

Table 3: The number and percentages of ratings assigned by users in the different networks (out of the 1.5 million ratings in our dataset). The sets are not necessarily overlapping.

circa 5000 ratings for 500 randomly selected users. In order to avoid problems related to cold start (when users have none or too few items for CF to generate good results) [Said *et al.*, 2009], for both users and items, we limit our evaluation to users who have rated at least 30 movies. For each of these users, 10 movies having been rated with a value above the user's average rating were extracted into the evaluation set (i.e. the set of true positive recommendations). The rest of the ratings were used for training. The recommendation algorithm was run twice for the 50 pairs of datasets, once taking the networks into consideration, and once neglecting the additional data. The results presented in this paper are averaged over all runs.

The recommendation algorithm used in our experiments was a slightly modified version of *K-Nearest Neighbor* using the Pearson Correlation Coefficient as the neighbor similarity measure. The pearson similarity of two users who were connected in the networks was multiplied by a factor of $10,000$ (the number of users in our dataset) in order to significantly affect the similarity measure. Experiments were performed with K set to 200. Additionally, a random recommender was used as a baseline for comparison. It should be noted that the algorithm itself is not the focus of our evaluation, rather the effects of using this additional information for recommendation.

3.3 Results

We evaluate our recommendations with the Mean Average Precision (MAP) and Precision at 10 (P@10) measures. These measures where chosen since they are well-known and widely-used in the field of Recommender Systems and Information Retrieval, providing a statistically sound estimate of the recommendation quality [Herlocker *et al.*, 2004].

Table 3(a) shows the precision levels obtained in our experiments. As the training and test splits for each network type have been created separately (due to the sets not necessarily being overlapping), they can thus not be compared to each other directly. Therefore, the table also shows the result of a standard Pearson-based KNN recommender on the same training and test split compared to the values of social recommendations. Table 3(b) shows the MAP values in a similar fashion.

Our resulting recommendations using social and pseudo-social networks perform between 0.2% and 5.4% better (in MAP values) than a regular KNN recommender and similarly in terms of P@10. We find that the pseudo-social network created from fan relations does not add much to the recommendation quality. Our belief is that this is related to the large number of edges in the network and the fact that people can be fans for different reasons. The other networks have larger impacts, with the explicitly stated social network performing better than the rest. We believe this is due to the relations expressing a type of "common ground" or agreement between the two parties.

4 Related Work

Recommender systems research originated in the late 1980's - early 1990's [Resnick *et al.*, 1994] and has since then become a ubiquitous topic found at almost every machine learning or information retrieval related conference.

(a) P@10

Type	P@10 10K	P@10	%
Friendship	$1.993E-3$	$1.847E-3$	7.9%
Comments	$3.551E-4$	$3.383E-4$	5.0%
Fans	$6.365E-4$	$6.342E-4$	0.4%
Diggs	$3.093E-4$	$2.845E-4$	8.7%

(b) MAP

Type	MAP 10K	MAP	%
Friendship	$5.154E-3$	$4.890E-3$	5.4%
Comments	$4.519E-3$	$4.417E-3$	2.3%
Fans	$5.208E-3$	$5.198E-3$	0.2%
Diggs	$4.493E-3$	$4.310E-3$	4.2%

Table 4: The Precision at 10 and Mean Average Precision values for our approach and for regular Collaborative Filtering for the same training and test datasets and the percental improvement.

More recently, much of the focus of the recommender systems community was on the Netflix Prize. Pilászy and Tikk [Pilászy and Tikk, 2009], presented provocative results showing that meta data related to movies is of little value when it comes to predicting movie ratings. Kirmenis and Birturk [Kirmenis and Birturk, 2008], on the other hand, show that a similar approach that utilizes user related meta data generates better recommendations than a metadata ignorant approach. A similar hybrid approach is evaluated by Lekakos and Caravelas [Lekakos and Caravelas, 2006], where similarity-based data is combined with its content-based counterpart to improve recommendations, with good results.

Similarly to the Netflix Prize dataset, the Movielens dataset, provided by the GroupLens[4] research lab, has been frequently used in recommender systems research. For instance, Herlocker et al. [Herlocker *et al.*, 2002] evaluated neighborhood-based recommendation using Movielens in order to create design guidelines for collaborative filtering-based recommenders. Rashid et al. [Rashid *et al.*, 2002] researched the problem every system encounters when a new user starts using the service. Which items to recommend, or to decide which few items will give the system the most information about the user.

Amatriain et al. [Amatriain *et al.*, 2009], pose that re-rating movies is of significantly higher value than rating new ones. They show how the amount of time that has passed since the original rating affects the users' new rating, and thus the quality of the recommendations.

Guy et al. [Guy *et al.*, 2009] create a system for recommending items based on a users' aggregated *familiarity* network. In this work, the familiarity network is created by assigning relations between users based on sources such as

[4] `http://www.grouplens.org/`

co-authorship of wiki pages within an organization's internal network, similar to the implicit networks studied in this paper. The results show that the familiarity network produces better recommendations than classical similarity based approaches. A similar approach is presented by Bonhard and Sasse [Bonhard and Sasse, 2006].

Another approach related to familiarity networks is the concept of trust-based recommendation. Golbeck and Hendler's [Golbeck and Hendler, 2006] present an approach based on explicitly defined trust gathered through the *FilmTrust*[5] movie recommendation website. FilmTrust asks its users to assign trust values to their peers, thus stating whose taste to follow and whose not to follow. They conclude that trust does add to the quality of the recommendations.

5 Conclusion and Future Work

In this paper we presented early stage results which indicate that the networks that users are part of contain latent information not present in the data found through ordinary user-based collaborative filtering methods. We showed, in a movie recommendation scenario, that the actions of users as well as their social networks are implicitly reflected in their rating behavior.

The work presented shows that there is much to gain by simple extensions of current standard algorithms. However, the approach needs to be extended and further researched in order to gain more insight into the different types of networks users can be part of, and how they affect the quality of recommendations. Also, combinations of networks, which we did not touch upon should be taken into consideration. Similarly, extending this research outside of the movie domain could provide a deeper understanding of network types and the users in them. Our current work focuses on combinations of several network types as well as the integration of demographic data, i.e. age, gender, etc.

The main contribution of our paper is an evaluation of different user-related (pseudo-) social networks, explicit and implicit. We have shown that, in a movie recommendation scenario, these types of networks appear to have an effect on the quality of recommender algorithms, even when implemented by very simple means.

6 Acknowledgments

The authors would like to express their gratitude to the Moviepilot team who contributed to this work with dataset, relevant insights and support. The work in this paper was conducted in the scope of the KMulE project which was sponsored by the German Federal Ministry of Economics and Technology (BMWi).

References

[Amatriain *et al.*, 2009] X. Amatriain, J. M. Pujol, N. Tintarev, and N. Oliver. Rate it again: increasing recommendation accuracy by user re-rating. In *RecSys'09*, 2009.

[Bonhard and Sasse, 2006] P. Bonhard and M. Sasse. Knowing me, knowing you using profiles and social networking to improve recommender systems. *BT Technology Journal*, 24(3), 2006.

[Gantz and Reinsel, 2010] J. Gantz and D. Reinsel. The digital universe decade are you ready?, 2010.

[Golbeck and Hendler, 2006] J. Golbeck and J. Hendler. FilmTrust: movie recommendations using trust in web-based social networks. In *CCNC'06*, volume 1, 2006.

[Guy *et al.*, 2009] I. Guy, N. Zwerdling, D. Carmel, I. Ronen, E. Uziel, S. Yogev, and S. Ofek-Koifman. Personalized recommendation of social software items based on social relations. In *RecSys'09*, 2009.

[Herlocker *et al.*, 2002] J. Herlocker, J. Konstan, and J. Riedl. Empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Informational Retrieval*, 5, 2002.

[Herlocker *et al.*, 2004] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22, 2004.

[Kirmenis and Birturk, 2008] O. Kirmenis and A. Birturk. A Content-Based user model generation and optimization approach for movie recommendation. In *Workshop on ITWP*. AAAI Press, 2008.

[Lekakos and Caravelas, 2006] G. Lekakos and P. Caravelas. A hybrid approach for movie recommendation. *Multimedia Tools and Applications*, 36(1-2), 2006.

[Montebello, 1998] M. Montebello. Information overload–an IR problem? *String Processing and Information Retrieval, International Symposium on*, 1998.

[Pilászy and Tikk, 2009] I. Pilászy and D. Tikk. Recommending new movies: even a few ratings are more valuable than metadata. In *RecSys'09*, 2009.

[Rashid *et al.*, 2002] A.M. Rashid, I. Albert, D. Cosley, S.K. Lam, S. McNee, J.A. Konstan, and J. Riedl. Getting to know you: Learning new user preferences in recommender systems. In *IUI'02*. ACM, 2002.

[Resnick *et al.*, 1994] P. Resnick, N. Iacovou, M. Sushak, P. Bergstrom, and J. Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *CSCW'94*, Chapel Hill, NC, 1994. ACM.

[Said *et al.*, 2009] A. Said, R. Wetzker, W.d Umbrath, and L. Hennig. A hybrid PLSA approach for warmer cold start in folksonomy recommendation. In *RecSys Workshop on RSWeb*, 2009.

[Said *et al.*, 2010a] A. Said, S. Berkovsky, and E. W. De Luca. Putting things in context: Challenge on context-aware movie recommendation. In *CAMRa'10*, 2010.

[Said *et al.*, 2010b] A. Said, E. W. De Luca, and S. Albayrak. How social relationships affect user similarities. In *IUI Workshop on SRS*, 2010.

[Siedler, 2010] MG Siedler. Netflix now 15 million users strong with over 60 percent of them streaming content. `http://techcrunch.com/2010/07/21/netflix-users/` (retrieved April, 2011), 2010.

[5] `http://trust.mindswap.org/FilmTrust/`