

An Experimental Study in Cross-Representation Mediation of User Models

Federica Cena
Università degli Studi di Torino
Dipartimento di Informatica
Torino - Italy
federica.cena@unito.it

Cristina Gena
Università degli Studi di Torino
Dipartimento di Informatica
Torino - Italy
cristina.gena@unito.it

Claudia Picardi
Università degli Studi di Torino
Dipartimento di Informatica
Torino - Italy
claudia.picardi@unito.it

ABSTRACT

The paper presents the result on cross-representation mediation of user models in the context of movie recommendation. We analyze the possibility of initializing the user models for a content-based recommender starting from movie ratings provided by users in other social applications. We focus in particular on (i) an approach for inferring user model preferences from rating and (ii) the experimentation of several methods to solve the missing value problem exploiting community-based ratings. We tested different variations of the proposed approach exploiting a subset of the MovieLens 10M Dataset, computing rating predictions, and MAE.

Keywords

Cross-representation mediation of user models, Movie recommendation, Content-based recommender systems

1. INTRODUCTION

Mediation of user models refers to the possibility of providing a user model U in a system S by exploiting other user model(s) the user may have in the same or other systems, which differ from U in one or several respects (e.g. representation, domain, or others). We focus on CF (collaborative-filtering) to CB (content-based) cross-representation mediation [1] where a CF user model (consisting in a set of user ratings) is used to extract a CB user model (expressing the user interest for the features characterizing the domain items).

This approach may offer a potential solution to mitigate the cold-start and sparsity problems in recommender systems, as well as the “paradox of the active user”, who is so eager to use the system that she does not want to spend time giving the system information about herself (in our case, her interests), which could improve the system performance.

User model mediation in general, and cross-representation mediation in particular, does not *per-se* solve the problem of *missing values*: for a content-based user model as we have described, this problem consists in not knowing the user interest for a certain feature of an item to be recommended.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

This may be due to the user’s reticence in providing information about her, or – in case the user model the target of a mediation – to an incompleteness in the source user model, or more generally to the fact that there are too many features for the user having expressed an interest in all of them, either directly or indirectly.

One possible solution is to fill the value with community-based preferences from another source outside the target system [4]. This is in line with the idea of social information access [3], i.e., methods for exploiting users past interaction within an information system, in order to provide better access to information to the future users. Similarly, in CF recommenders, the missing ratings are filled with default values, such as the middle value of the rating range, or the average user or item rating [2, 5].

In this article, we experiment with a combination of cross-representation mediation (user model transfer from CF to CB) as a way to extract both the user model, and a default interest or preference value to cover for missing values.

2. USER MODEL EXTRACTION

We propose to extract the content-based user model from a set of movie ratings. Each movie m , is described by an *id*, a *title*, and a set of features $desc(m) = \{F_1, \dots, F_n\}$ where each feature F_i is a pair (*category*, *value*). Categories are:

$$FCat = \{genre, directors, actors, production_country, tags\}.$$

The user model $UM(u)$ we compute consists of an **interest** function $int_u : Features \rightarrow [0, 1]$ (the interest of user u in a certain feature F) and an **action count** function $act_u : Features \rightarrow \mathbb{N}$ (the number of actions by u involving a feature F). $act_u(F)$ is extracted as the number of movies rated by u which have F in their description ($mov(u|F)$). $int_u(F)$ is extracted as the average (normalized) rating given by u to such movies.

The interest function we extract in this way tends to be under-defined for “sparse” feature categories (such as the *actor* and *director* categories), where there are a lot of possible different feature values. We propose to tackle this problem by providing a default interest function \overline{int} , computed from a set of “external” ratings, provided by users not otherwise involved in the experiment. Such function is meant to be used in place of int_u whenever it is undefined.

We experimented with three ways of computing the default function.

- **Middle value:** $\overline{int} = int_{mid} =$ the global average external rating (a constant).

- **Average user:** $\overline{int} = int_{ave} =$ the average external rating for each given feature (as if the community providing the external ratings was a single user with its own user model).
- **Notoriety-based:** $\overline{int} = int_{ntr} = int_{mid} * ntr$ where ntr represent the notoriety of a feature F , computed as the (normalized) number of ratings given to movies with F in it.

3. EVALUATION

We evaluated our approach using the MovieLens 10M Dataset [6] and, in order to compare our results with existing work, we replicated the experimental settings by Berkovsky et al. [1]. The main difference with their settings relies on the dataset. In fact Berkovsky et al. exploited the EachMovie dataset, storing 2,811,983 ratings of 72,916 users on 1,628 movies, which is no longer available. However MovieLens was originally based on this dataset.

We first randomly selected 5000 users from the MovieLens DB, whose 701017 samples were used as “external” ratings for computing the three default interest functions int_{mid} , int_{ave} , and int_{ntr} .

We then created 10 groups of 325 users each, according to the number of ratings available for each of them.

For each group we then performed a 10-fold cross evaluation, selecting 90% of their samples as training set and the remaining 10% as evaluation set.

For each user in each group, we extracted a user model from the training set, and then run a basic content-based recommender on the evaluation set, to obtain *rating predictions*. To evaluate the approaches we computed the mean absolute error (MAE) between such predictions and the *actual* ratings provided by the user herself.

Rating predictions (for each user u and movie m in her evaluation set) according to the following formula:

$$pr_{u,m} = \frac{\sum_{c \in FCat} score_c(u, m)}{\sum_{c \in FCat} w_c}, \quad (1)$$

$score_c(u, m)$ is the score obtained by movie m according to the user model of u with respect to category c . It is computed as the average interest the user has for the features in m associated with category c .

We experimented with two variants of formula 1: (i) including all categories (EQ variant), and (ii) excluding *production_country* (FS), taking in this latter case into account the *feature selection* analysis presented in [1], which suggests to exclude this category. We tested each combination of a default interest method (*mid*, *ave*, *ntr*) and a weight set (EQ, FS) with a ten-fold cross-evaluation. We checked the resulting MAEs for each pair of combinations in order to verify the statistical significance of their differences. We obtain the following results: (i) for both EQ and FS variants the int_{ave} default function provides the best results, and int_{int} outperforms int_{ntr} (significance $\geq 99\%$); (ii) excluding the production country category does not result in significant changes in MAEs, confirming this category as irrelevant.

We also compared our best-performing combination (FS+*ave*) with the results in [1], which in turn compared their cross-representation mediation approach (CBFS) with a standard collaborative filtering technique (CF) applied to the same data.

In the work by Berkovsky and colleagues CBFS obtains a

lower MAE than CF (about 0.16 to 0.185) for users with less than 75 samples; for more than 75 samples, CBFS worsens, stabilizing approximately at 0.20, while CF improves, decreasing to 0.17. Our approach tends to improve as the number of samples increases: its MAE is 0.17 (worse than CBFS, better than CF) for users with 1-25 samples, but then decreases to about 0.15. For more than 25 samples our approach obtains a better MAE than both CBFS and CF.

4. CONCLUSION

The results presented in this paper open the door to several further investigations. A first line of inquiry goes in the direction of combining our approach with [1], in order to benefit from the advantages of both. In fact, the two algorithms for user model inference take into account quite different aspects, and this is reflected in the different behavior they exhibit when used in recommendation. As future work, we will implement their approach using the MovieLens dataset, in order to have the two studies completely comparable.

A second study may investigate the possibility of using community ratings for fine tuning the weight of each feature category, rather than adopting the coarser approach of feature selection, where each category basically weights either 1 (selected) or 0 (discarded).

5. REFERENCES

- [1] S. Berkovsky, T. Kuflik, and F. Ricci. Cross-representation mediation of user models. *User Model. User-Adapt. Interact.*, 19(1-2):35–63, 2009.
- [2] J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, UAI’98, pages 43–52, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers Inc.
- [3] P. Brusilovsky. Social information access: The other side of the social web. In V. Geffert, J. Karhumäki, A. Bertoni, B. Preneel, P. Návrát, and M. Bieliková, editors, *SOFSEM 2008: Theory and Practice of Computer Science: 34th Conference on Current Trends in Theory and Practice of Computer Science, Nový Smokovec, Slovakia, January 19-25, 2008. Proceedings*, pages 5–22, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.
- [4] I. Cantador, I. Fernández-Tobías, S. Berkovsky, and P. Cremonesi. *Recommender Systems Handbook*, chapter Cross-Domain Recommender Systems, pages 919–959. Springer US, Boston, MA, 2015.
- [5] M. Deshpande and G. Karypis. Item-based top-n recommendation algorithms. *ACM Trans. Inf. Syst.*, 22(1):143–177, Jan. 2004.
- [6] F. M. Harper and J. A. Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4):19, 2016.