Mining the Social Semantic Web for making cross-domain recommendations

Ignacio Fernández-Tobías
Universidad Autónoma de Madrid
28049 Madrid, Spain
i.fernandez@uam.es

Cross-domain recommender systems filter and suggest items in a target domain by exploiting user preferences and/or domain knowledge available in a (likely related) source domain. In our research we are developing a framework for cross-domain recommendation capable of mining heterogeneous sources of information available in the so-called Social Semantic Web, such as semantically annotated data, user generated contents, and contextual signals.

Cross-domain information, recommender systems, semantic networks, social tagging, contextualization.

1. INTRODUCTION

Recommender systems filter and suggest items (e.g. Web pages, books, movies, music compositions) that best suit the user's preferences –tastes, interests and priorities– without requiring launching explicit search queries, as is usually done in information retrieval systems.

The vast majority of current recommender systems focus on a single domain. Netflix makes personalized recommendations of movies and TV series, and Last.fm suggests music compositions and artists. E-commerce sites like Amazon, however, may take benefit from exploiting the user's preferences on diverse types of items to provide recommendations in different but somehow related domains. Recommendation across domains could mitigate the cold-start problem when little information about the user's preferences is available in a target domain, and are potentially more diverse and serendipitous than single-domain recommendations.

The goal of cross-domain recommender systems is thus to suggest items in a target domain by exploiting user preferences and/or domain knowledge available in a (likely related) source domain. Hence, a major challenge for making cross-domain recommendations is how to transfer knowledge across different domains, i.e., how to build bridges between domains, in which usually there is little or no overlap between user preferences or item attributes.

Motivated by this fact, we propose a number of methods that exploit different information sources available in the so-called Social Semantic Web to establish links between domains, and develop novel recommendation algorithms that make use of such links. Specifically, we exploit semantically annotated data, user-generated contents (social tags, in particular), and contextual information.

2. RELATED WORK

Winoto and Tang (2008) identify three issues to investigate in cross-domain recommendation: the existence of global correlations between user preferences for items in different domains, the method to exploit data on a source domain for predicting preferences on a target domain, and the methodology and metrics to evaluate cross-domain recommendations. Li (2011) surveys works on cross-domain collaborative filtering, classifying existing approaches according to the type of knowledge transferred: rating patterns, latent features, and user/item inter-domain correlations. Tiroshi and Kuflik (2012) evaluate the influence of the involved domains in the recommendation using a kNN approach in which the neighborhood of a user in the target domain is selected among the most similar users from a source domain.

Besides collaborative filtering there have been some attempts to establish semantic relations between items of different domains. Loizou (2009) annotates and links items by means of concepts and relations extracted from Wikipedia. Then, using such relations, users and items are incorporated into a graph upon which a probabilistic recommendation model is built. Social tags have also been used to establish relations between items of different domains. Kaminskas and Ricci (2011) show that emotional tags can be used to effectively select

Proceedings of the 5th BCS IRSG Symposium on Future Directions in Information Access

music that fits places of interest. Shi et al. (2011) utilize tags to build inter-domain user-to-user and item-to-item similarities. These similarities are proportional to the numbers of tags shared by profiles from different domains, and are incorporated as constraints into a probabilistic matrix factorization model for collaborative filtering.

3. PROPOSED APPROACHES

3.1 Crossing domains with semantic relations

In (Kaminskas et al., 2013) we present an approach that exploits the multi-domain DBpedia ontology (http://dbpedia.org) to build semantic networks linking items from different domains. The approach consists of two main components: a class-level network of classes and relations belonging to certain domains of interest, and instance-level networks that are automatically generated instantiations of the class-level network for particular input items, by extracting and filtering information from the above ontology. On the built instance-level networks, a graph-based spreading activation algorithm ranks items in the target domain according to their semantic relatedness with the input items. We have evaluated our approach instantiating the framework for suggesting musicians and music compositions semantically related to places of interest. Through two user studies we show that our approach is able to effectively identify music items relevant for the users and suited to the target places.

3.2 Crossing domains with social tags

As shown by Shi et al. (2011), and Kaminskas and Ricci (2011), tags can play the role of a common feature space for user/item profiles from distinct domains. Kaminskas and Ricci (2011) show that a limited set of emotional tags assigned to both music and places can reveal latent similarities between both types of items. Shi et al. (2011) use the tags common to two folksonomies for computing similarities between users/items in the folksonomies' domains. We hypothesize that generic lexicons can be built with social tags, and that these lexicons can be used to generate profiles which would act as bridges to compute cross-domain user/item similarities. In Fernández-Tobías et al. (2013) we present an approach that automatically builds a lexicon that describes core emotions (e.g. *happy*, *sad*), by mining synonyms and antonyms extracted from the online Dictionary.com thesaurus. We show that our emotion representation is in accordance with the well-known psychological circumplex model of affect. Moreover, using tag co-occurrences, we link the lexicon with domain-specific emotion folksonomies extracted from social tagging systems, which let us transform domain-specific tag-based profiles into generic emotion-oriented profiles. For instance, a movie annotated with tags associated to the emotional category *suspenseful* strongly overlaps with the *tense* and *nervous* core emotions of the circumplex model. In a user study we show our method accurately infers emotions evoked by items in the movie and music domains.

3.3 Crossing domains with contextual data

We believe that a recommender system could exploit relations established through contextual signals, such as the time (e.g. movies and music compositions usually consumed on Christmas), and the user's mood and emotions (e.g. movies and music compositions that usually yield nostalgic feelings). We have explored a context modeling approach, formulating the recommendation task as a classification problem, and investigating if using or not those features in addition to content attributes actually leads to better accuracy. Preliminary empirical results have indicated that time context has preference predictive power in the movie and music domains, and it is improved by including social context in the movie domain.

4. REFERENCES

Fernández-Tobías, I., Cantador, I., Plaza, L. (2013) Modeling Emotions with Social Tags. 21st Conf. on User modeling, Adaption, and Personalization.

Kaminskas, M., Fernández-Tobías, I., Ricci, F., Cantador, I. (2013) Ontology-based identification of music for places. 13th Conf. on Information and Communication Technologies in Tourism.

Kaminskas, M., Ricci, F. (2011) Location-Adapted music recommendation using tags. 19th Conference on User Modeling, Adaptation, and Personalization, pp. 183-194.

Li, B.: Cross-domain collaborative filtering: A brief survey (2011) 23rd Conference on Tools with Artificial Intelligence, pp. 1085–1086.

Loizou, A. (2009) How to recommend music to film buffs: Enabling the provision of recommendations from multiple domains. PhD dissertation, University of Southampton.

Shi, Y., Larson, M., Hanjalic, A. (2011) Tags as bridges between domains: Improving recommendation with tag-induced cross-domain collaborative filtering. 19th Conf. on User modeling, Adaption, and Personalization, pp. 305-316.

Tiroshi, A., Kuflik, T. (2012) Domain ranking for cross domain collaborative filtering. 20th Conference on User Modeling, Adaptation, and Personalization, pp. 328-333.

Winoto, P., Ya Tang, T. (2008) If you like The devil wears Prada the book, will you also enjoy The devil wears Prada the movie? A study of cross-domain recommendations. New Generation Computing 26(3), pp. 209-225.