Tutorial on Cross-domain Recommender Systems

Iván Cantador
Departamento de Ingeniería Informática
Universidad Autónoma de Madrid
28049 Madrid, Spain
ivan.cantador@uam.es

Paolo Cremonesi
Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
20133 Milan, Italy
paolo.cremonesi@polimi.it

ABSTRACT
Cross-domain recommender systems aim to generate or enhance personalized recommendations in a target domain by exploiting knowledge (mainly user preferences) form other source domains. This may beneficial for generating better recommendations, e.g. mitigating the cold-start and sparsity problems in a target domain, and enabling personalized cross-selling for items from multiple domains. In this tutorial, we formalize the cross-domain recommendation problem, categorize and survey state of the art cross-domain recommender systems, discuss related evaluation issues, and outline future research directions on the topic.

Categories and Subject Descriptors
H.1.2 [**Models and Principles**]: User/Machine Systems – *human information processing*. H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *information filtering*. H.4.2. [**Information Systems Applications**]: Types of Systems – *decision support*.

General Terms
Algorithms, Experimentation, Human Factors, Performance.

Keywords
Recommender Systems, Cross-Domain Recommendation, Cross-Selling, Knowledge Transfer.

1. INTRODUCTION
The majority of recommender systems offer recommendations for items belonging to a single domain. For instance, Netflix recommends movies and TV programs, and Last.fm recommends songs and music albums. These domain-specific systems have been successfully deployed by numerous websites, and the single-domain recommendation functionality is not perceived as a limitation but rather pitched as a focus on a certain market.

Large e-commerce sites like Amazon and eBay, however, often store user feedback for items from multiple domains, and in social networks users often provide feedback for a variety of topics. It may, therefore, be beneficial to leverage all the available user data provided in various domains and systems, in order to generate more encompassing user models and better recommendations. Instead of treating each domain or item type independently, knowledge acquired in a *source* domain could be exploited in another, *target* domain. The research challenge of transferring the knowledge, and the business potential of delivering recommendations spanning across multiple domains, have triggered an increasing interest in the so-called cross-domain recommender systems.

RecSys '14, Oct 06-10 2014, San Jose or vicinity, CA, USA.
ACM 978-1-4503-2668-1/14/10.
http://dx.doi.org/10.1145/2645710.2645777

2. OUTLINE AND TOPICS
2.1 Formulation of the Cross-Domain Recommendation Problem
In the literature authors have considered several notions of domains. For instance, some have treated items like *movies* and *books* as belonging to different domains, while others have considered items such as *action movies* and *comedy movies* as distinct domains. In the tutorial, we discuss different definitions of domain according to the attributes and types of recommended items, distinguishing between four levels, namely (i) item attribute, (ii) item type, (iii) item, and (iv) system level. We also consider the type of user feedback, evaluation scale, and context used by the recommender to explain the notion of domain.

The research on cross-domain recommendation generally aims to exploit knowledge from a source domain $\mathcal{D}_S$ to perform or improve recommendations in a target domain $\mathcal{D}_T$. Let $\mathcal{U}_S$ and $\mathcal{U}_T$ be their sets of users, and let $\mathcal{I}_S$ and $\mathcal{I}_T$ be their sets of items. The users of a domain are those who expressed preferences (ratings, reviews, tags, or consumption logs) for the domain items. The items of a domain do not necessarily have preferences from users of the domain, but may have content-based features that establish their belonging to the domain. Analyzing the literature, we observe that the addressed tasks are diverse, and a consensual definition of the cross-domain recommendation problem has not been formulated yet. Sorted in increasing order of complexity, we distinguish between the following three recommendation tasks:

- **Multi-domain recommendation**: recommend items in both the source and target domains, i.e., recommend items in $\mathcal{I}_S \cup \mathcal{I}_T$ to users in $\mathcal{U}_S$ (or, equivalently, in $\mathcal{U}_T$ or $\mathcal{U}_S \cup \mathcal{U}_T$).
- **Linked-domain recommendation**: recommend items in the target domain by exploiting knowledge from the source and target domains, i.e., recommend items in $\mathcal{I}_T$ to users in $\mathcal{U}_S$ by exploiting knowledge about $\mathcal{U}_S \cup \mathcal{U}_T$ and/or $\mathcal{I}_S \cup \mathcal{I}_T$.
- **Cross-domain recommendation**: recommend items in the target domain by exploiting knowledge from the source domain, i.e., recommend items in $\mathcal{I}_T$ to users in $\mathcal{U}_S$ by exploiting knowledge about $\mathcal{U}_S$ and/or $\mathcal{I}_S$.

In the tutorial, we detail the above recommendation tasks.

2.2 Categorization of Cross-domain Recommender Systems
The cross-domain recommendation problem has been addressed from various perspectives in different research areas. It has been faced by means of user preference aggregation and mediation strategies for the cross-system personalization problem in *user modeling* [1][2][8], as a potential solution to mitigate the cold-start and sparsity problems in *recommender systems* [3][9][10], and as a practical application of knowledge transfer in *machine learning* [5][6][7]. This has entailed the development of a wide array of approaches, which in many cases are difficult to compare due to

e.g., the user preferences they use, the cross-domain scenarios they deal with, and the algorithms on which they are based.

In the tutorial, we discuss different categories of cross-domain recommendation techniques proposed in the literature. Then, aiming to reconcile these categories in a way that captures and unifies their core ideas, we present a categorization focused on the exploitation of knowledge in cross-domain recommendation, which dictates the following two-level taxonomy:

Aggregating knowledge. Knowledge from various source domains is aggregated to perform recommendations in a target domain. Three use cases are considered:

- ***Merging user preferences***: the aggregated knowledge consists of user preferences, e.g., ratings, tags, transaction logs, and click-through data.
- ***Mediating user modeling data***: the aggregated knowledge comes from user modeling data exploited by various recommender systems, e.g., user similarities and user neighborhoods.
- ***Combining recommendations***: the aggregated knowledge is composed of single-domain recommendations, e.g. rating estimations and rating probability distributions.

Linking and transferring knowledge. Knowledge linkage or transfer between domains is established to support recommendations. Three variants are considered:

- ***Linking domains***: linking domains by a common knowledge, e.g., item attributes, association rules, semantic networks, and inter-domain correlations.
- ***Sharing latent features***: the source and target domains are related by means of implicit latent features.
- ***Transferring rating patterns***: explicit or implicit rating patterns from source domains are exploited in the target domain.

2.3 Evaluation of Cross-domain Recommender Systems

Whether a cross-domain recommender system is good or bad cannot be evaluated without taking into account for what it is intended. The nature of the evaluation must be connected to the purpose for which the recommendations are required. Thus, in the tutorial. we first describe the cross-domain **recommendation goals** addressed in the literature:

- **Addressing the new user problem**. This may be solved by exploiting the user's preferences collected in a different source domain.
- **Addressing the new item problem**. This is particularly evident when cross-selling new products from different domains.
- **Improving accuracy**. Data collected outside the target domain can increase the rating density and upgrade the recommendation quality.
- **Improving diversity**. The diversity of recommendations can be improved by considering multiple domains, as this may provide a better coverage of the range of user preferences.
- **Enhancing user models**. The main goal of cross-domain user modeling applications is to enhance user models. Achieving this goal may have personalization-oriented benefits such as discovering new user interests for the target domain, enhancing similarities between users and items, and measuring vulnerability in social networks.

Based on these goals, in the tutorial we discuss **evaluation scenarios**, **methodologies** and **metrics**.

2.4 Future Research on Cross-domain Recommendation

One interesting issue that deserves more attention in future work is the synergy between cross-domain and **context-aware recommendations**: different contexts (e.g. location, time, and mood) can be treated as different domains, and context can be treated as a bridge between different domains.

Another important issue concerns the **evaluation metrics** adopted. A common practice with cross-domain recommender systems is to evaluate their relevance through prediction metrics such as MAE and RMSE. However, the true advantage of cross-domain recommendations is not necessarily in their accuracy, but rather in their novelty and diversity, which may lead to a higher satisfaction and utility for the user.

A third open research issue refers to the usage of cross-domain recommender systems for reducing the **user model elicitation** effort. Cross-domain recommender systems could be used as alternative elicitation tools able to build detailed user profiles without or mitigating the need to collect explicit user preferences in cold-start situations.

3. REFERENCES

[1] Abel, F., Helder, E., Houben, G.-J., Henze, N., Krause, D. 2013. Cross-system User Modeling and Personalization on the Social Web. *User Modeling and User-Adapted Interaction*, 23(2-3), pp. 169-209.

[2] Berkovsky, S., Kuflik, T., Ricci, F. 2008. Mediation of User Models for Enhanced Personalization in Recommender Systems. *User Modeling and User-Adapted Interaction*, 18(3), pp. 245-286.

[3] Cremonesi, P., Tripodi, A., Turrin, R. 2011. Cross-domain Recommender Systems. *Proc. of the 11th IEEE International Conference on Data Mining Workshops*, pp. 496-503.

[4] Fernández-Tobíias, I., Cantador, I., Kaminskas, M., Ricci, F. 2012. Cross-domain Recommender Systems: A Survey of the State of the Art. *Proc. of the 2nd Spanish Conference on Information Retrieval*, pp. 187-198.

[5] Gao, S., Luo, H., Chen, D., Li, S., Gallinari, P., Guo, J. 2013. Cross-Domain Recommendation via Cluster-Level Latent Factor Model. *Proc. of the 17th and 24th European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 161-176.

[6] Li, B., Yang, Q., Xue, X. 2009. Can Movies and Books Collaborate? Cross-domain Collaborative Filtering for Sparsity Reduction. *Proc. of the 21st International Joint conference on Artificial Intelligence*, pp. 2052-2057.

[7] Pan, W., Xiang, E. W., Liu, N. N., Yang, Q. 2010. Transfer Learning in Collaborative Filtering for Sparsity Reduction. *Proc. of the 24th AAAI Conf. on Artificial Intelligence*, pp. 210-235.

[8] Shapira, B., Rokach, L., Freilikhman, S. 2013. Facebook Single and Cross Domain Data for Recommendation Systems. *User Modeling and User-Adapted Interaction*, 23(2-3), pp. 211-247.

[9] Shi, Y., Larson, M., Hanjalic, A. 2011. Tags as Bridges between Domains: Improving Recommendation with Tag-induced Cross-domain Collaborative Filtering. *Proc. of the 19th International Conference on User Modeling, Adaption, and Personalization*, pp. 305-316.

[10] Tiroshi, A., Berkovsky, S., Kaafar, M. A., Chen, T., Kuflik, T. 2013. Cross Social Networks Interests Predictions Based on Graph Features. *Proc. of the 7th ACM Conference on Recommender Systems*, pp. 319-322.