Chapter 1

Introduction

Nowadays we are living in an era that is overloaded with information. For example, if you enter the word "book" as the keyword for a Google search, you will get about 1,500,000,000 hits! Decision-making in this environment can sometimes become a nightmare. There are too many choices and we simply cannot explore them all. Therefore, it would be really helpful to have a system to help us to find the right choice. Such a system can help us only if it knows what we like and what we do not like. Then, the system can filter irrelevant choices and recommend options that are likely to be interesting for us. Such systems, which learn user preferences and provide personalized recommendations to them are called *Recommender Systems* (Adomavicius & Tuzhilin, 2005; Goldberg *et al.*, 1992; Konstan *et al.*, 1997a).

Today all major web companies are equipped with recommender systems. The recommended item can be anything. To name just a few examples: a book in Amazon, a movie in Netflix, a friend on Facebook, and so on. Table 1.1 summarizes the most important recommender systems in different areas. Recommender systems can also be used for physical environments, such as museums. Visitors to physical museums are often overwhelmed by the vast amount of information available in the space they are exploring, making it difficult for them to select the content that is likely to be interesting to them personally. Recommender systems can recommend specific exhibits that are likely to be interesting to individual visitors (Karimi *et al.*, 2011d).

Recommendation algorithms fall into two main categories: content-based and collaborative filtering (Adomavicius & Tuzhilin, 2005). In the content-based approach, there is a set of attributes for each item. Items that have similar attributes to those that have been liked by the target user are recommended to them. For example, if a user likes a book from an author, another book from the same author is recommended to that user because both books have the same attributes. The main drawback of content-based recommendation is that it requires complete profiles for items (Resnick & Varian, 1997). On the other hand, collaborative filtering does not need such profiles. It predicts the interests of users by reusing taste information from similar users (Burke, 2002a; Konstan *et al.*, 1997a). Also, there are some methods that combine content-

Table 1.1: Examples of recommender systems and their applications

Company	Category	Recommended Item
Amazon, eBay, Netflix	e-Commerce	Product
Facebook, LinkedIn	Social Network	Friend, Group
YouTube	Social Network	Video Clip
CNN, Routers	News Agency	News, Article
TiVo	Public media channel	TV Channel
Pandaro	Public media channel	Radio Channel
Yahoo! Answer	Q & A Community	Questions
Delicious	Social bookmarking	Web page
Bibsonomy	Social bookmarking	Tags

based and collaborative filtering and call it *hybrid* recommendation algorithms (Burke, 2002b; Pennock *et al.*, 2000; Schein *et al.*, 2002).

The collaborative filtering method falls into two categories: memory-based algorithms and model-based algorithms. In memory-based techniques, the value of the unknown rating is computed as an aggregate of the ratings of some other (usually, the N most similar) users for the same item (Konstan *et al.*, 1997a). Model-based collaborative techniques provide recommendations by estimating parameters of statistical models for user ratings. Nevertheless, recent research (especially as has been demonstrated during the Netflix challenge[1]) indicates that Matrix Factorization (MF) (Koren *et al.*, 2009) is a superior predictive model compared to other approaches (Koren *et al.*, 2009).

1.1 Motivation

Evidently, the performance of collaborative filtering depends on the amount of information that users provide regarding items, most often in the form of ratings. This problem is amplified for new users because they have not provided any rating, which impacts negatively on the quality of generated recommendations. This problem is called *new user problem* or *cold-start problem*. Different approaches can be taken to deal with the cold-start problem. We can switch to content-based algorithms until enough ratings are given by new users. However, this is not possible if items do not have complete attributes. Another possibility is to use non-personalized methods, such as most popular. But such methods do not perform well and it is risky to use them for new users. Note that new users of a recommender system are sensitive, they leave the system if they get a bad first impression and go to competitor systems. Needless to say, new users are vital for the revenue of companies.

A simple and effective way to overcome this problem, is by posing queries to new users so that they express their preferences about selected items, e.g. by rating them.

[1] www.netflixprize.com

Nevertheless, the selection of items must take into consideration that users are not willing to answer a lot of such queries. To address this problem, *active learning* methods have been proposed to acquire the most informative ratings, i.e ratings from users that will help most in determining their interests (Harpale & Yang, 2008; Jin & Si, 2004). In this thesis, we will study the application of active learning for the cold-start problem in recommender systems.

1.2 Contribution

The aim of this thesis is to take inspiration from the literature of active learning for machine learning and develop new methods for the new user problem in recommender systems. The thesis consists of two parts. In the first part, to be consistent with the settings of active learning in machine learning and the related works on the new user problem in recommender system, it is assumed that the new user is always able to rate the queried items. Next, this constraint is relaxed and new users are allowed not to rate the items. Specifically, our contributions are as follows:

- **Formal Definition of Active Learning for Recommender Systems.** We formalize the problem of active learning for recommender systems. To the best of our knowledge, this is the first formal definition of active learning for recommender systems.
- **Active Learning for Aspect Model.** We develop a new active learning method for the aspect model (Hofmann, 2003; Hofmann & Puzicha, 1999). It takes into account the learning algorithm of the aspect model. While this method competes in terms of accuracy with a complicated Bayesian method, it is in the order of magnitude faster than it. Moreover, we compare the aspect model to matrix factorization (Koren, 2008) from the perspective of active learning and show that matrix factorization is more accurate and faster than the aspect model.
- **Active Learning for Matrix Factorization.** We develop five active learning algorithms for the cold-start problem based on matrix factorization. Most of these methods capitalize explicitly on the characteristics of matrix factorization, such as the learning algorithm and similarity in the latent space.
- **Factorized Decision Trees (FDT).** FDT combines the power of matrix factorization with the easy-to-understand structure of decision trees to build adaptive questionnaires. Furthermore, we propose a sampling method to speed up the tree construction algorithm. It is called Most Popular Sampling (MPS).
- **Learning Active Learning** We introduce a new approach for active learning in recommender systems that is called *learning active learning.* The main idea is to consider past users as (artificial) new users in order to learn the right queries to be asked to new users for active-learning purposes.

- **Active Learning for Museum Recommender Systems** We discuss how active learning can be used for recommender systems in museums. As a museum recommender system, we take our own project called RFID-Enhanced Museum for Interactive Experience (REMIX) (Karimi *et al.*, 2011d).

- **Empirical Evaluation and Analysis.** All algorithms proposed in this thesis are thoroughly evaluated on the basis of several real-life datasets and compared to state-of-the-art algorithms.

After mentioning the contributions, we would like to clarify what is not covered in this thesis. The cold-start problem has three meanings: new user, new item and new system. New system means a system that has just been inaugurated, so there are no ratings from any user for any item. Active learning has been applied for the new system problem (Boutilier *et al.*, 2003; Rish & Tesauro, 2008; Rubens & Sugiyama, 2007; Sutherland *et al.*, 2013), the new item problem (Deodhar *et al.*, 2009; Huang, 2007; Park & Chu, 2009). Also, some works combine the new system and new user problem in the sense that they suppose when new users enter the recommender system, there are not yet many active users because the system has not been running for a very long time (Elahi *et al.*, 2012, 2013, 2014). In this thesis, we address only the new user problem, assuming that there are already enough active users in the system. It should be mentioned that in general new user and new item problems are symmetric. However, we will not evaluate the developed methods for the new item problem.

In this thesis, we address the new user problem by posing queries to new users in order to get ratings from them. There are other approaches to deal with this problem that are not the concern of this thesis:

- **Implicit Feedback.** (Zhang *et al.*, 2009; Zigoris, 2006) leverage implicit feedback, such as search keywords or user clicks to learn new user preferences.

- **Content-based recommendation** (Gantner *et al.*, 2010; Gunawardana & Meek, 2008) combine content-based attributes with collaborative filtering.

- **Demographic information.** (Safoury & Salah, 2013) use demographic information on new users.

1.3 Published Work

Most of the work reported in this thesis has already been published at peer-reviewed international conferences:

- Rasoul Karimi, Martin Wistuba, Alexandros Nanopoulos, Lars Schmidt-Thieme (2013): Factorized Decision Trees for Active Learning in Recommender Systems, in Proceedings of the IEEE International Conference on Tools with Artificial Intelligence (ICTAI), Washington D.C, USA.

- Rasoul Karimi, Christoph Freudenthaler, Alexandros Nanopoulos, Lars Schmidt-Thieme (2013): Towards Optimal Active Learning for Matrix Factorization in Recommender Systems, in Workshop on Knowledge Discovery, Data Mining and Machine Learning (KDML-2013), Bamberg, Germany*(resubmission)*
- Rasoul Karimi, Christoph Freudenthaler, Alexandros Nanopoulos, Lars Schmidt-Thieme (2012): Exploiting the Characteristics of Matrix Factorization for Active Learning in Recommender Systems, in Doctoral Symposium of the 6th Annual ACM Conference on Recommender Systems (RecSys), Dublin, Irelan, pp. 317-320
- Rasoul Karimi, Christoph Freudenthaler, Alexandros Nanopoulos, Lars Schmidt-Thieme (2011): Towards Optimal Active Learning for Matrix Factorization in Recommender Systems, in 23th IEEE International Conference on Tools With Artificial Intelligence (ICTAI), Florida, USA.
- Rasoul Karimi, Christoph Freudenthaler, Alexandros Nanopoulos, Lars Schmidt-Thieme (2011): Non-myopic Active Learning for Recommender Systems based on Matrix Factorization, in 12th IEEE International Conference on Information Reuse and Integration (IRI), Las Vegas, USA.
- Rasoul Karimi, Christoph Freudenthaler, Alexandros Nanopoulos, Lars Schmidt-Thieme (2011): Active Learning for Aspect Model in Recommender Systems, in IEEE Symposium on Computational Intelligence and Data Mining (CIDM).
- Rasoul Karimi, Alexandros Nanopoulos, Lars Schmidt-Thieme (2011): RFID-Enhanced Museum for Interactive Experience, in MultiMedia for Cultural Heritage (MM4CH), Modena, Italy

1.4 Chapter Overview

Besides the introduction, this thesis is organized as follows:

- In chapter 2 we formalize the problem of active learning for recommender systems. We introduce notations, which will be used in the rest of the thesis. Furthermore, we describe datasets and the evaluation protocol.
- In chapter 3, we explain the background techniques that are needed to understand this thesis. This consists of two sections: in the first section, different aspects of active learning are discussed and in the second section, the state-of-the-art recommendation methods are illustrated.
- Related work is reviewed in chapter 4. First, the related work on active learning is reviewed, especially those which have influenced cold-start recommendation. Then the related work on active learning for the new user problem in recommender systems is reviewed.