Electronic Commerce Research and Applications 13 (2014) 440–455

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

# Leveraging prior ratings for recommender systems in e-commerce

Guibing Guo [a,*], Jie Zhang [a], Daniel Thalmann [a], Neil Yorke-Smith [b,c]

[a] School of Computer Engineering, Nanyang Technological University, Singapore
[b] American University of Beirut, Beirut, Lebanon
[c] University of Cambridge, Cambridge, UK

ARTICLE INFO

*Article history:*
Received 16 August 2013
Received in revised form 26 June 2014
Accepted 14 October 2014
Available online 30 October 2014

*Keywords:*
Prior ratings
Recommender systems
Rating confidence
Similarity measure
Data sparsity
Cold start

ABSTRACT

User ratings are the essence of recommender systems in e-commerce. Lack of motivation to provide ratings and eligibility to rate generally only after purchase restrain the effectiveness of such systems and contribute to the well-known data sparsity and cold start problems. This article proposes a new information source for recommender systems, called *prior ratings*. Prior ratings are based on users' experiences of virtual products in a mediated environment, and they can be submitted prior to purchase. A conceptual model of prior ratings is proposed, integrating the environmental factor *presence* whose effects on product evaluation have not been studied previously. A user study conducted in website and virtual store modalities demonstrates the validity of the conceptual model, in that users are more willing and confident to provide prior ratings in virtual environments. A method is proposed to show how to leverage prior ratings in collaborative filtering. Experimental results indicate the effectiveness of prior ratings in improving predictive performance.



## 1. Introduction

User ratings are crucial for recommender systems in e-commerce in order to provide quality personalized product recommendations. However, users can lack motivation to provide ratings (why should I bother to report my experience of an item?), and ratings can generally be given only after purchase (how can I share my experience of an item I have not tried?). Without sufficient rating information for preference modelling, the effectiveness of recommender systems is hindered—as seen in well-known problems such as *data sparsity* and *cold start* (Su and Khoshgoftaar 2009).

The former problem, data sparsity, refers to the difficulty in finding a sufficient number of reliable users, since users in general only rate a small portion of items, while the latter problem, cold start, refers to the difficulty in providing accurate recommendations for those users who have rated a few items, e.g., less than five items. Cold start is an extreme case of the data sparsity problem. The key issue is that only limited rating information is available for preference modelling, whereby inherently and severely hindering the recommendation performance.

Although many approaches have been proposed to address these problems either by furthering the use of existent ratings (Ahn 2008; Guo et al. 2013b), or by including to additional information (Massa and Avesani 2007; Konstas et al. 2009; Guy et al. 2009; Jamali and Ester 2011; Guo et al. 2012, 2014a), few researchers have attempted to elicit more user ratings from the perspective of user interfaces, so as to inherently mitigate the severity of these problems. On the other hand, Virtual Reality (VR) environments (e.g., Second Life (Rymaszewski 2007)), have received considerable attention because of their ability to provide users with immersive virtual user experiences. Users can experience media more richly and can interact in real time with *virtual products*—the 'second existence' of real products in a mediated environment (Hemp 2006). Although these environments offer potentially useful information for preference modelling, research on e-commerce in VR is still in its infancy.

This article proposes a new information source, called *prior ratings*, built upon *virtual product experiences* (Li et al. 2003). Prior ratings can be issued prior to purchase by interacting with virtual products represented in a mediated environment. The aim of this article is to study (1) the concept and nature of prior ratings with respect to product attributes and environmental factors; and (2) the usefulness of prior ratings in coping with the data sparsity and cold start problems of recommender systems.

In particular, first, we propose a conceptual model of prior ratings to provide a principled foundation, integrating the environ-

* Corresponding author at: Institute for Media Innovation, 50 Nanyang Drive, Research Techno Plaza, XFrontiers Block, Level 03-01, Singapore 637553, Singapore. Tel.: +65 84005322.

*E-mail addresses:* gguo1@ntu.edu.sg (G. Guo), zhangj@ntu.edu.sg (J. Zhang), danielthalmann@ntu.edu.sg (D. Thalmann), nysmith@aub.edu.lb (N. Yorke-Smith).

http://dx.doi.org/10.1016/j.elerap.2014.10.003
1567-4223/

mental factor *presence* whose effects on product evaluation have not been studied previously. Five hypotheses and two research questions are proposed to verify the validity of the conceptual model. We recruited volunteers and performed user studies in both 2D (website) and 3D (virtual store) user interface modalities. The results demonstrate the validity of the conceptual model under our experimental settings, and indicate that users are more willing and confident to give prior ratings in a VR store (due to a stronger sense of presence) than in a website.

Then, second, by integrating the prior rating and confidence data collected from the user studies into a novel adapted collaborative filtering technique that we develop, we empirically demonstrate the usefulness of prior ratings in improving recommendation performance in terms of accuracy and coverage.

Our work sheds light on inherently alleviating the data sparsity and cold start problems by the design of user interfaces with rich media and interactions that elicit confident prior ratings from users.

**Contribution.** Summarized, the major contributions of this article are in three-fold: (1) we introduce a new information source (and its conceptual model) called *prior ratings*, which holds potential to benefit recommender systems in e-commerce; (2) we design a user study to validate the conceptual model of prior ratings; and (3) we propose and evaluate a collaborative filtering technique to demonstrate how to leverage prior ratings in predicting the ratings of products. A preliminary version of our work is published in (Guo et al. 2013a).

**Outline.** Section 2 gives an overview of related research in the literature. Section 3 details the proposed conceptual model of prior ratings, and proposes five related hypotheses and two research questions. Section 4 reports on a user study designed to validate the conceptual model. Then, Section 5 discusses the relationship between prior ratings and other information sources for recommender systems, and the limitations and implications of the user study and results. Based on the rating and confidence data collected from the user study, Section 6 introduces a variant of traditional collaborative filtering technique and demonstrates the usefulness of prior ratings in improving predictive performance. Finally, Section 7 concludes our work and outlines the future research.

## 2. Related work

Many approaches have been proposed to resolve the data sparsity and cold start problems. From the perspective of information source, we classify them into two categories. The first category adopts rating information only. There are two kinds of approaches, memory-based and model-based. For memory-based methods, various authors have proposed new similarity measures to better model user correlation to resolve the concerned issues, given the inefficiency of traditional similarity measures (Lathia et al. 2008). Specifically, Lathia et al. (2008) propose a concordance-based measure based on the amount of concordant, discordant and tied pairs of ratings between two users. It measures the extent to which the two users agree with each other. Ahn (2008) develops the PIP measure by studying the semantics of ratings in terms of *Proximity, Impact* and *Popularity*. The basic idea is that users with semantic agreements should be more similar than those with semantic disagreements. Bobadilla et al. (2012) design the *singularities measure* from the perspective of item singularity. The intuition behind is that ratings agreed on high-singular items should be counted more than those agreed on low-singular items in computing user similarity. Guo et al. (2013b) propose a novel *Bayesian similarity* by taking into account both the direction and length of rating vectors. The weights of evidences (i.e., rating pairs) are carefully computed and integrated into the Bayesian inference. Experimental results show that better performance can be achieved than the other similarity measures.

However, memory-based approaches do not scale well to large-scale data sets. In contrast, model-based methods possess better scalability and often perform better than memory-based ones (Koren et al. 2009). The reason is that not only ratings of two users but also ratings of other users are adopted to learn the features of users and items, and thus better handle the data sparsity and cold start issues. Gunawardana and Meek (2008) report to capture pairwise item interactions by using a Boltzmann machine, whose parameters are associated with item contents. They show that better performance is achieved in the case of cold-start situations. Gantner et al. (2010) attempt to learn a function mapping user/item attributes to latent features of a matrix factorization model. With such mappings, the latent factors learned by a matrix factorization can be applied to new users or new items. Liu et al. (2011) propose a representative-based matrix factorization method that aims to find out the most representative users and items in the system. Then, for the cold-start users, their preferences can be elicited by asking them to rate the most representative items; the same holds for the cold-start items. To combat the data sparsity problem, Ahmed et al. (2013) propose a method to learn user preferences over item attributes by applying a personalized Bayesian hierarchical model, which combines both globally and locally learned user preferences.

In summary, all these approaches, both memory-based and model-based, attempt to integrate user/item attributes into a certain recommendation model in order to handle the concerned issues. However, the attributes of users/items may not be available for a recommender system due to the concern of, e.g., privacy.

The second category adopts additional information other than ratings. For example, Konstas et al. (2009) take into consideration both the social annotation (tag) and friendships inherently established among users in a music track recommender system. By leveraging data from multiple channels including memberships in a project, Guy et al. (2009) build a system for recommending people of interest to active users. Ma et al. (2011) propose a matrix factorization model regularized by users' social friendships. The intuition is that a user-specific vector should be close to that of his friends. A stronger relationship than friendship is social trust, based on which Massa and Avesani (2007) develop a trust-aware recommender system. They show that data sparsity can be better handled without a significant decrement in predictive accuracy. Guo et al. (2014b) define trust in recommender systems as one's belief in the other's ability in providing accurate ratings. Guo et al. (2012, 2014a) merge the ratings of trusted neighbours to form a new rating profile for the active users by which the concerned problems are shown to be alleviated.

Other than these memory-based approaches, trust is also integrated into matrix factorization models for better recommendation performance. Ma et al. (2009) propose the social trust ensemble that forms a linear combination of a matrix factorization model and a trust-based neighbourhood model. Jamali and Ester (2010) propose a matrix factorization model where a user's user-specific vector is influenced by the average of her trusted neighbours. Tang et al. (2013) take into account both the global and local trust in the recommendation model, and show that predictive performance can be improved to some extent. Yang et al. (2013) report that the active user's ratings will be influenced by the ratings of users who trust her and those who she trusts. Experimental results show that their approach works the best among all other trust-based approaches. One of the problems of matrix factorization models lies in the difficulty of explaining how recommendations are generated, as these patterns are based on latent features. Another problem is that users' social information may not exist,