Overlaying social information: The effects on users' search and information-selection behavior

Valeria Orso [a,*], Tuukka Ruotsalo [b], Jukka Leino [b], Luciano Gamberini [a,c], Giulio Jacucci [b]

[a] *Department of General Psychology, University of Padova, via Venezia 8, Padova, Italy*
[b] *Helsinki Institute for Information Technology HIIT, Department of Computer Science University of Helsinki, Gustaf Hällströmin katu 2B, Helsinki, Finland*
[c] *Human Inspired Technologies Research Centre, University of Padova, via Venezia 8, Padova, Italy*

ARTICLE INFO

ABSTRACT

Previous research investigated how to leverage the new type of social data available on the web, e.g., tags, ratings and reviews, in recommending and personalizing information. However, previous works mainly focused on predicting ratings using collaborative filtering or quantifying personalized ranking quality in simulations. As a consequence, the effect of social information in user's information search and information-selection behavior remains elusive. The objective of our research is to investigate the effects of social information on users' interactive search and information-selection behavior. We present a computational method and a system implementation combining different graph overlays: social, personal and search-time user input that are visualized for the user to support interactive information search. We report on a controlled laboratory experiment, in which 24 users performed search tasks using three system variants with different graphs as overlays composed from the largest publicly available social content and review data from Yelp: personal preferences, tags combined with personal preferences, and tags and social ratings combined with personal preferences. Data comprising search logs, questionnaires, simulations, and eye-tracking recordings show that: 1) the search effectiveness is improved by using and visualizing the social rating information and the personal preference information as compared to content-based ranking. 2) The need to consult external information before selecting information is reduced by the presentation of the effects of different overlays on the search results. Search effectiveness improvements can be attributed to the use of social rating and personal preference overlays, which was also confirmed in a follow-up simulation study. With the proposed method we demonstrate that social information can be incorporated to the interactive search process by overlaying graphs representing different information sources. We show that the combination of social rating information and personal preference information improves search effectiveness and reduce the need to consult external information. Our method and findings can inform the design of interactive search systems that leverage the information available on the social web.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

* Corresponding author.
E-mail address: valeria.orso@unipd.it (V. Orso).

1. Introduction

The era of social computing has kindled massive amounts of new types of data on the web, including tags, ratings, and reviews of a variety of content. A wealth of work before the rise of social review and rating sites has investigated ways to utilize social data in recommending and personalizing information. The early approaches successfully used social data in recommender systems by predicting ratings via collaborative filtering (Breese et al., 1998; Goldberg et al., 1992; Konstas et al., 2009; Resnick et al., 1994) and utilizing social information to enhance Web search ranking (Agrawal et al., 2015; Amitay et al., 2009; Carmel et al., 2009). Previous research has also demonstrated that personalization using user's behavioral data can lead to improved search effectiveness (Pitkow et al., 2002; Teevan et al., 2005) and that social information can be successfully incorporated in retrieval methods (Bao et al., 2007; Guy et al., 2010).

The motivation for further investigating approaches in this direction is twofold. Firstly, *blurring of search and recommendation* harnessing the variety of social data, such as ratings, tags, reviews, search time queries, and personal preferences, calls for new approaches to investigate how different social and content information can be simultaneously incorporated in an online information-seeking process to empower information access.

Secondly, *user benefits in information selection* quantifying the effect of social data for user behavior and utility in selecting information to enhance decision-making requires empirical approaches that extend beyond simulating rating predictions or rankings to reveal the behavioral and experimental benefits for the users.

Conversely, recommender systems or personalization research focuses on predicting static ratings or a ranking for content items (Konstas et al., 2009; Teevan et al., 2005). Instead, we propose studying how social information embedded in ranking and result presentation affects users' behavior when making decisions to select useful information.

For example, consider the case of a user looking for a place to have dinner in an unfamiliar city and seeking advice from the social web. The user may have search content preferences that he or she expresses as a query, e.g., "Asian restaurants". The user would also be likely to choose a restaurant matching her long-term personal preferences, say "Thai food," rather than relying solely on the query "Asian restaurant". At the same time, the user would be likely to rely on the common opinion of other like-minded users in the social web on the quality of a particular restaurant offering Asian food.

The scenario exemplifies different layers of social search and personalization that could be potentially relevant for a typical search task: search time content preferences (Asian restaurant), personal content preferences (Thai food), and social preferences (opinions of other users who are likeminded and have preferences regarding Asian restaurants, Thai food, or similar restaurants).

To make the decision about where to have dinner, the user could benefit from a ranking that takes into account all of this information. Moreover, the user could need information on why and how the restaurants suggested to him/her were ranked: based on their match to the query or her long-term preferences, the ratings that other like-minded people had given for the restaurants, or a combination of these.

To this end, we investigate user search behavior using socially created information layers in presence of various data: content tags, social ratings, personal preferences, and search time queries.

1.1. Research objective

The main objective of the present study is to investigate whether the combined deployment of different layers of social information affects the users' search and information selection behavior. More specifically, we aim at identifying which of the information layers can improve the search effectiveness, thus benefiting the user's information-selection behavior. In addition, we aim at testing the effectiveness of the proposed computational method in an interactive online setting.

1.2. Contributions

The contributions of this article are as follows:

- *Empirical evidence of the benefits of social overlay information for search effectiveness and information-selection behavior*. The effects are demonstrated by data comprising logs, eye-tracking recordings, simulation data, and questionnaires from a task-based user study in which 24 participants planned their activities in a point-of-interest search and recommendation scenario on the largest publicly available social web dataset from Yelp (Blomo et al., 2013) consisting of hundreds of thousands of ratings, tags, and items.
- *Computational method utilizing overlaid social graphs*. This method combines social, personal, and search-time user input. Such pieces of information are modeled as overlaid graphs, and relevance is computed by computing random walks with restarts on the overlaid graph. User's search-time preferences, such as queries or other feedback, are incorporated as prior information to the restart computation and arbitrary amount of overlays can be processed in the same graph computation framework.