

Review of RecSys'14 Joint Workshop on Interfaces and Human Decision Making for Recommender Systems

John O'Donovan

Dept. of Computer Science
University of California
Santa Barbara, CA
jod@cs.ucsb.edu

Nava Tintarev

Dept. of Computer Science,
University of Aberdeen,
United Kingdom
n.tintarev@abdn.ac.uk

Introduction and Background

Since the emergence of recommender systems, a large majority of research focuses on objective accuracy criteria and less attention has been paid to how users interact with the system and the efficacy of interface designs from the end-user perspective. The field has reached a point where it is ready to look beyond algorithms, into users interactions, decision making processes and overall experience.

Accordingly, the IntRS@RecSys workshop was held in 2014 to explore the human aspect of recommender systems, with a particular focus on the impact of interfaces and interaction design on decision-making and user experiences with recommender systems, and to explore methodologies to evaluate these human aspects of the recommendation process that go beyond traditional automated approaches. The workshop was a successful amalgamation of the “Decisions” (Felfernig et al.) and “InterfaceRS” (Tintarev et al.) workshops from previous RecSys conferences.

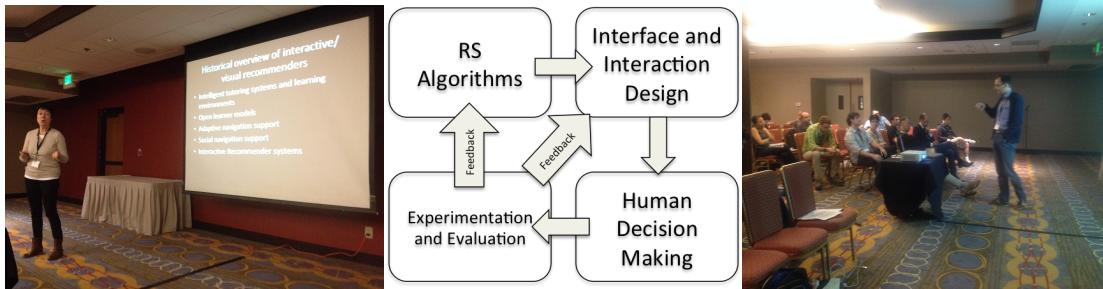


Figure 1. Left: Julita Vassileva's Keynote on Visualization for Recommender Systems; Center: High level workshop themes and interactions.; Right: Wolfgang Wörndl talking about Mobile Interfaces and rating mechanisms.

Review of Program and Presentations

The workshop was held at the main conference venue in the Crowne Plaza hotel, Foster City, California. Over the full day there were approximately 40 attendees with academia and industry both well represented, making for a lively and interesting discussion throughout.

After a brief opening discussion by John O'Donovan, the day began with an excellent keynote by Julita Vassileva on the state of the art in visual interactive recommender systems. Julita described research in the area to date in a nice summary and then introduced some of her own labs recent works in visualization for recommenders. This gave a good background to the research area and set the stage well for the sessions that followed.

The remainder of the day was divided into three sessions: Ratings (chaired by Alexander Felfernig), Explanations (chaired by John O'Donovan) and Visualization (chaired by Pasquale Lops), and the day closed with a discussion led by John and Pasquale. The ratings session had two long presentations (20+15mins) and one short (10+10). Wolfgang Wörndl presented a study of rating behavior based on different types of interaction methods on mobile devices, such as touch, drag and tilt. Gediminas Adomavicius discussed a novel method for automatically identifying bias in user preference ratings, and adjusting an algorithm accordingly. Martin Stettinger described the Choicla system, which provides decision support for groups of users in making personnel decisions. Martin demoed the tool's interactive UI and showed how different users could enter their preferences about potential new hire for a company. This session ended with a discussion about the benefits and potential dangers in falsely representing (boosting) average user ratings on items. The industry side argued that there is a potential sales benefit. A few people in the audience counter-argued that this would have a negative long-term impact on user trust in the recommender system. As a researcher interested in trust issues, I strongly agree with the latter point.

Lunch was held in the hotel and many discussions carried over to the restaurant. The first afternoon session was on Explanations for recommender systems, chaired by John O'Donovan. The session opened with Béatrice Lamche presenting a study on interactive explanations in mobile shopping recommenders. This was followed by Derek Bridge from UCC, Ireland talking about their interestingly titled work "If you liked Herlocker et al.'s paper, then you might like this paper too". Derek gave a great description of their experiment, which aimed to capture the influence of different types of recommendation explanation with respect to a decision-making task. The session closed with a short presentation by Markus Zanker on the persuasiveness of fact-based explanations for recommender systems. Specifically, this study looked at three explanation styles (sentences, facts or argument style) and measured their impact in a live user experiment, finding that fact-based explanations are more robust over time than sentence-based explanations. Discussion at the end of the session included that explanations are critical for recommender systems, but the right degree/level of explanation is context sensitive and can impact user acceptance of a recommendation in a significant way.

Participants refueled with a coffee break before the last session on Visualization for recommender systems, chaired by Pasquale Lops. This was a shorter 2-paper session. Julita Vassileva took the stage again to present work by her student Wesley Waldner on visualizing timeline activity on Twitter. The final paper of the day was

presented by Denis Parra (PUC Chile) on the effect of different set-based visualizations on user exploration of recommendations. In particular, Denis described a user study on cluster-based and Venn-diagram-based representations of recommendations. The study found that for a typical audience, Venn-diagrams are better received than the more complex cluster-map visualizations.

Review of Discussions and Closing Comments

Overall it was a great day and there was no shortage of questions and discussion. Every talk had more questions and discussions than time allowed for, indicating that there is good interest in the general workshop topics from the community. There were many notable names in the audience, including Alfred Kobsa (UCI), Barry Smyth (UCD) and Martijn Willemsen (TUE). John O'Donovan and Pasquale Lops led the wrap-up session and summarized the key points of the day. Key points included that by adding a human into the recommendation pipeline through a visual and potentially interactive user interface, we are introducing many new variables which complicate the system and makes evaluation very difficult. During this session there was a call for more benchmarking for systems that support visualization of and interaction with recommender algorithms, perhaps in the form of a dataset or platform that allowed fair comparison between systems. This discussion met with a unanimous agreement from the audience. Vijai Mohan (Amazon) suggested the possibility of a challenge competition along the lines of the Netflix prize competition but for visual interactive recommenders. He theorized that Amazon could potentially provide a platform to plug in a recommender interface to real products. Participants could then submit interactive interfaces, which can then be evaluated through real purchases and other interaction metrics. This suggestion received a lot of positive feedback for the audience, but it was agreed that it would take significant effort to make such an event a reality. This will be a key discussion topic for IntRS'15, which has just been accepted to the 2015 conference in Vienna as a full day workshop.