

Visualizing Recommendations to Support Exploration, Transparency and Controllability

Katrien Verbert^{1,4}, Denis Parra^{2,3}, Peter Brusilovsky², Erik Duval⁴

¹Department of Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands
{k.verbert@tue.nl}

²School of Information Sciences, University of Pittsburgh, 135 North Bellefield Avenue,
Pittsburgh, PA 15260, USA {dap89, peterb}@pitt.edu

³Department of Computer Science, PUC Chile, Santiago, Chile

⁴Departement Computerwetenschappen, KU Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium
{erik.duval@cs.kuleuven.be}

ABSTRACT

Research on recommender systems has traditionally focused on the development of algorithms to improve accuracy of recommendations. So far, little research has been done to enable user interaction with such systems as a basis to support exploration and control by end users. In this paper, we present our research on the use of information visualization techniques to interact with recommender systems. We investigated how information visualization can improve user understanding of the typically black-box rationale behind recommendations in order to increase their perceived relevance and meaning and to support exploration and user involvement in the recommendation process. Our study has been performed using TalkExplorer, an interactive visualization tool developed for attendees of academic conferences. The results of user studies performed at two conferences allowed us to obtain interesting insights to enhance user interfaces that integrate recommendation technology. More specifically, effectiveness and probability of item selection both increase when users are able to explore and interrelate multiple entities – i.e. items bookmarked by users, recommendations and tags.

Author Keywords

User interfaces for recommender systems; information visualization; user studies.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User interfaces. H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors; Design; Experimentation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

UI'13, March 19–22, 2013, Santa Monica, CA, USA.

Copyright © 2013 ACM 978-1-4503-1965-2/13/03...\$15.00.

INTRODUCTION

Interactive information visualization and recommendation techniques have both been explored as ways to help people deal with abundance of information. The main advantage of interactive visualization is that a multi-dimensional representation allows the user to more easily see multiple aspects of data while being in control when exploring information. The main advantage of the traditional recommendation approach is that offering a clear list of items ranked by perceived interest reduces cognitive overload associated with exploring a rich set of items.

In this paper, we present our research on the combination of both approaches. We investigated how graphical representations and the ability to combine a personalized prospect offered by a recommender engine with other valuable prospects can improve user trust in the results offered by the black-box recommender engines and increase user ability to find interesting items.

Our work has been motivated by the presence of multiple *relevance prospects* in modern personalized social tagging systems. An important feature pioneered by social tagging systems and later used in other kinds of social systems is the ability to explore different *community relevance prospects* by examining items bookmarked by a specific user or items associated by various users with a specific tag. Items bookmarked by a specific user offer a *social relevance prospect*: if this user is known and trustable or appears to be like-minded (bookmarked a number of items known as interesting) a collection of his or her bookmarks is perceived as an interesting and relevant set that is worth to explore for more useful items. Similarly, items marked by a specific tag offer a *content relevance prospect*. Items related to a tag of interest or a tag that was used to mark many known interesting items are also perceived as potentially relevant and worth to explore. In this context, a ranked list of recommended items offered by a specific recommender engine can be considered as yet another *personalized relevance prospect*.

The problem that we address is that existing personalized social systems do not allow their users to explore and combine multiple relevance prospects. Only one prospect can be explored at any given time – a list of recommended items, a list of items bookmarked by a specific user or a list of items marked with a specific tag. We believe that exploring a single prospect is not sufficient since none of the prospects could be fully reliable and trustable by the users (that includes recommendations generated by black-box engines). In this context, the ability to combine prospects might offer a more reliable and attractive way to explore information. For example, knowing that a specific item has been not only recommended by a recommender engine, but also bookmarked by two trustable users can remarkably increase user trust in the quality of this item.

To solve the aforementioned problem and to offer users an approach to explore and combine multiple relevance prospects, we suggest a specific interactive visualization approach. This visualization embodies suggestions offered by various recommender systems *as recommender agents* that can be perceived as being analogous to human users and as a result, exportable in parallel with the relevance prospects offered by users and tags in social systems. We believe that using interactive visualization can increase the transparency of the recommendation process and allow the users to be in control of their exploration.

A special issue on interfaces for recommender systems [25] illustrates the interest and importance of intelligent interfaces for recommender systems. Such interfaces are investigated to provide new capabilities to the users of the recommender system to search, browse, and understand the results of the recommendations. In our work, we focus on the use of *information visualization techniques* to support such new capabilities. The research contribution of this work is threefold:

- 1) First, we present a novel and synergetic approach to combine multiple relevance prospects that include *personalized relevance* as offered by different recommenders and *social relevance* as offered in social bookmarking systems. In this approach, recommender systems are presented as agents and their interrelationship can be explored (i.e. a user can explore which items are suggested by multiple recommender agents). In parallel, real users and their bookmarks are shown and users can explore both interrelationships between users as well as interrelationships between agents and users. Third, a user can explore tags and interrelationships between tags, users and agents. To our knowledge, this combination has not yet been investigated as a way to support exploration and controllability, and to increase trust and acceptance of recommendations.
- 2) Second, we present a user interface which serves to both explain the provenance of recommendations in a transparent way and to support exploration. Users can

browse bookmarks of other users, tags and suggestions of recommender agents as a basis to find relevant items.

- 3) Third, we present the results of a study that evaluated the usefulness of this interactive interface with 21 users. In this study we attempted to assess how the ability to control and combine entities involved in the recommendation process can influence user performance and satisfaction.

This paper is organized as follows: first we present related work in the area of user interfaces for recommender systems. Then, we introduce TalkExplorer, an interactive visualization of users, tags, talks at a conference and recommendations for conference attendees. The evaluation of this visualization at conferences is presented next. Finally, we discuss the results of this case study, lessons learnt and future research opportunities.

BACKGROUND AND RELATED WORK

Recommender systems

Recommender algorithms can be broadly categorized in three areas:

1. *Collaborative filtering* (CF) recognizes commonalities between users or between items on the basis of explicit or implicit relevance indications [13] such as ratings [3] and tags [25]. Implicit data used by recommender systems include actions like reading [21] or watching TV series [14]. Starting from the active user, user-based CF first identifies a neighborhood of similar users to then recommend items based on the nearest neighbors' top preferences. On the other side, item-based CF identifies similarly rated items, and then uses these similar items to identify the recommendations. CF is the most widely implemented and most mature technology [6].
2. *Content-based filtering* matches descriptions of items to descriptions of users [26]. This approach bases predictions on information about individual users and items, and ignores contributions from other users.
3. *Hybrid recommender systems* combine recommendation techniques, to gain better performance with fewer drawbacks [6].

Although these algorithms have been implemented and validated on a large scale in several application areas [22], there are important challenges that need to be addressed before recommendation can realize its full potential.

1. Collaborative recommendation techniques often suffer from *cold start* issues, i.e. they cannot make effective recommendations for new users or for new items that have no explicit or implicit relevance indicators yet [7].
2. It is *difficult to explain* the rationale behind recommendations to end users [12]: the complexity of recommendation algorithms often prevents users from comprehending recommended results and can lead to *trust* issues when recommendations fail.

3. Allowing users to *control* the way they can sort lists of recommendations [17] or the neighbors' contribution in a social recommender [18] has shown a positive effect in user satisfaction. However, there are several ways that users can control elements in the interface: which are the most effective for the user experience?

The design and development of user interfaces for recommender systems has gained increased interest. Such interfaces are researched to provide new capabilities to search, browse, and understand the results of the recommendations [27]. Among others, explaining recommendations to provide transparency and to increase trust has been extensively investigated [28]. In most cases, such explanations are presented in plain text and indicate why a specific item is suggested to a user – such as “*Because you have selected or highly rated: Movie A*”.

In addition to supporting transparency, we are particularly interested to enable interaction with recommender systems as a basis to support exploration and controllability. In recent years, some research has been done to visualize recommendations to enable such new capabilities. We elaborate on existing work in this area in the next section.

Visualizing recommendations

Most existing work in the area of visualizing recommendations focuses on interaction with collaborative filtering recommender systems. PeerChooser [23] is a visual interactive recommender that uses a graph-based representation to show relations between users and recommended items of a collaborative filtering recommender system. Similarly, SmallWorlds [10] allows exploration of relationships between recommended items and similar friends, in multiple layers of similarity. These systems enable users to explore such relationships as a basis to provide transparency and to support the user to find new relevant items. Pharos [31] is a social map-based recommender system that visualizes a summary of social network activity of different communities. The system uses topic modeling [4] to provide new users with an overview of the site, alleviating the cold start problem.

Some systems focus specifically on tags that are used by social recommenders. SFViz (Social Friends Visualization) [11] visualizes social connections among users and user interests in order to increase awareness in a social network and to help people find potential friends with similar interests. This system uses a Radial Space-Filling (RSF) technique [9] to visualize a tag tree and a circle layout with edge bundling to show a social network.

FaceTag [29] is a tool that helps users see the relationship between user-generated tags and recommended facets in a classification scheme. Tagsplanations [30] presents both tag relevance (the degree to which a tag describes an item) and tag preference (the user's sentiment toward a tag) as a basis to explain recommendations. Kammerer et al. [15] designed a tag-based search browser to recommend relevant tags for

further search. Research on this stream focuses on information and meta-information concerning items, and ignores the users who contributed such information and relationships among those users [11].

More recently, TasteWeights [5] has been introduced as a system that allows users to control the importance of friends and peers in social systems to obtain recommendations. Similar to our work, TasteWeights introduces the concept of an interface for hybrid recommender systems. The system elicits preference data and relevance feedback from users at run-time and uses these data to adapt recommendations to the current needs of the user. To our knowledge, this is one of the first systems that enables interaction with a hybrid recommender system and that can be adjusted by end users to control the output of the systems. In our work, we extend this concept of visualizing and combining the output of multiple recommenders as a basis to support exploration and controllability. While our work also enables end users to explore intersections of multiple recommenders (and hence in this respect we build on concepts that have been introduced by TasteWeights), the major difference and innovation of our work is that we allow end users to combine multiple *relevance prospects* in order to increase the perceived relevance and meaning of items. Our visualization embodies suggestions offered by various recommender systems *as recommender agents*. Items bookmarked by a specific user offer a *social relevance prospect*: if this user is known or appears to be like-minded a collection of her bookmarks is perceived as an interesting and relevant set that is worth to explore. Similarly, items marked by a specific tag offer a *content relevance prospect*. To our knowledge, this combination of multiple relevance prospects (agents, tags and real users) has not yet been explored as a way to support exploration and controllability, and to increase trust and acceptance of recommendations. Secondly, to our knowledge, our work is the first attempt to evaluate whether enabling end users to combine multiple relevance prospects increases the effectiveness and probability of item selection.

TALKEXPLORER

TalkExplorer is an interactive visualization tool that enables users to explore and bookmark research papers and talks at a conference using recommender agents and social data (tags, bookmarks, and connections to other users). The visualization was built as a component of the conference support system *Conference Navigator 3*. We first present Conference Navigator 3 (hereinafter CN3) and the recommendation functionalities that it provides. Then, we present the objectives of visualizing recommendations and details on the design and development of TalkExplorer. Evaluation results are presented in the next section.

Conference Navigator

Conference Navigator 3 is a social personalized system that supports attendees to academic conferences [24].

Study: UMAP 2012 User Survey Conference Navigator All Conferences FAQ Blog Search Denis Parra

UMAP UMAP 2012 Home Papers People My CN3 TalkExplorer

Developing a Scale for Assessing Instructor Attitudes Towards Open Learner Models

Tweet 0 Add to schedule Share to Group Search

Authors: Carrie Demmans Epp
 Type: Poster
 Your Tags: You have not added tags to this paper
 Track: UMAP 2012
 Scheduled: 2012-07-18 17:15-17:45
 DOI: /Vol-872/umap2012_poster_01.pdf

Abstract: The views of students with respect to open learner models have been extensively explored, whereas the views of instructors have not been fully explored. We have, therefore, begun the development of a scale to assess this. We describe the initial item development and revision which was based on cognitive interviews. Following this, we describe a pilot study that was performed to further refine the scale before attempting to validate it. Once the scale has been validated, it could be used to assess individual instructors' views towards specific open learner models or open learner model use within specified contexts.

Links:
 No link available.
 Add a link

Rate the relevance of this talk to your interests: (not relevant at all) 1 2 3 4 5 (strongly relevant)

People who scheduled this presentation, also scheduled:

- Modeling the Learner in 4-D (5)
Keynote
Ryan Baker
- Contextualizing Useful Recommendations (4)
Keynote
Francesco Ricci
- Towards Adaptive Information Visualization: On the Influence of User Characteristics (3)
Long Research Paper
Derek Toker, Cristina Conati, Giuseppe Carenini, Mona Haraty

Attending (5)
 Bahram Amini Julita Junjie W. Stephan Weibelzahl Xidao Wen

Shared in Groups (0)

Tags
 open_learner_models(1) instructor_attitudes(1) development(1) learner_model(1)

Feedback

Figure 1: Talk details page in Conference Navigator 3 (CN3)

At the time of writing, 18 conferences have been supported by CN3. Among different features, Conference Navigator 3 provides a conference schedule, a list of the conference talks and details of each talk (Figure 1). It also provides information about people related to the conference, such as the list of attendees and the list of authors. Users can add papers to create a personal schedule, they can add tags to each talk, and they can also connect with other CN3 users by following them (unidirectional relationship) or connecting with them (bidirectional relationship). Social information collected by CN3 is extensively used to help users find interesting papers. For example, in the page called “Top Items”, CN3 summarizes the most popular articles, the most active people, their institutions, and also the most popular tags associated to the papers.

When visiting talk information, as shown in Figure 1, users can also see who *scheduled* each talk during the conference and which tags were assigned to this talk. This social information is also used to provide links to similar papers (“People who scheduled this presentation, also scheduled:”) mimicking the well-known Amazon.com’s suggestions [19]. Similarly, when visiting a user page, other users can see which talks she is planning to attend (given that personal privacy settings provide access to this information). Finally, talks marked with a specific tag by the community of users can be explored.

In addition to social information access, CN3 offers personal recommendation of talks and papers. CN3 supports two kinds of recommendations: content-based and tag-based. The content-based engine constructs the user interest profile as a vector of terms with weights based on TF-IDF [1] using the content of the papers that the user has scheduled. Then, it recommends papers that match the profile of interests. The tag-based recommender engine makes use of the tags that users associate to conference talks. By matching user tags (tags applied by a target user) with item tags (tags assigned to different talks by the community of users) using the Okapi BM25 algorithm [20], the engine identifies relevant talks and suggests them to the active user.

Visualizing recommendations, tags and users

In the original CN3, ranked links produced by the content-based and tag-based recommenders are presented in separate pages and can be used by users to find new talks. In addition, users can explore bookmarks and tags of other users as a basis to explore new items. In this paper, we are particularly interested in assessing the potential influence of perceived relevancy and meaning of recommended items when we enable end users to explore and combine these multiple relevance prospects.

TalkExplorer is an interactive visualization developed on top of data collected by CN3.

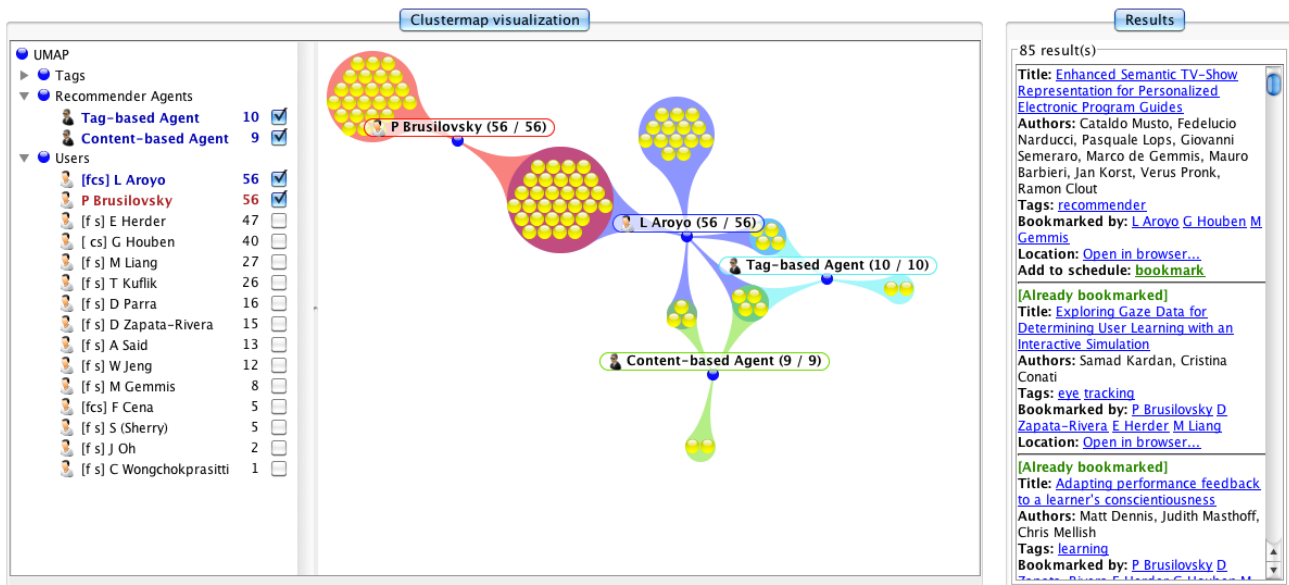


Figure 2: TalkExplorer

The interface serves to both explain the provenance of recommendations in a transparent way and to support exploration and control by end users. More specifically, users can browse and interrelate bookmarks of other users, tags and suggestions of recommender agents in order to find relevant items.

The visualization is implemented as a Java applet and uses the Aduna clustermap visualization library [1]. This software library visualizes sets of categorized objects and their interrelationships. The library has been used in related research to explore the interaction of users, resources and tags in social tagging systems [16]. In this past research, the library was deployed on top of delicious.com data to explore bookmarks of users and to support exploratory search. We adapted this initial version to visualize the interactions of users, tags, and *recommender agents* in terms of conference talks that they have in common – as illustrated in Figure 2. The objective was to explore new ways to enable end users to interact with recommender engines and items that are suggested to them.

Recommender systems are presented as talk-collecting agents so that interrelationship can be explored. In parallel, real users and their bookmarks are shown, and users can explore both interrelationships between users as well as interrelationships between agents and users (i.e. which other users have bookmarked talks that are recommended to them by one or more agents). In addition, relationships with tags can be explored to identify relevant items. We are hypothesizing that visualizing these relationships can help users to find relevant talks to attend at a conference, and that this visualization can provide transparency and increase trust.

As shown in Figure 2, fifteen users from the neighborhood of the active user are shown and users can explore items

that these other users have bookmarked. The selection of the neighborhood users is based on two types of explicit CN3 links - i.e. users that the active user follows (user name preceded by [f] in Figure 2) and connections (user name preceded by [c]), respectively. In addition, users that have similar interests based on common bookmarks are shown in this neighborhood (preceded by [s]). The output of two CN3 recommendation algorithms is shown as items selected by the content-based and the tag-based recommender agents in parallel to bookmarks of real users and to bookmarks associated with specific tags.

TalkExplorer allows users to explore the three different relevance prospect mentioned above using a three-component interface shown in Figure 2. On the left side, the entity selection panel allows users to select tags, users and recommender agents to be added and displayed in the canvas area. This canvas area, at the center of the applet, shows a clustermap visualization - i.e., different clusters of talks linked by connected components. The labeled circles in this canvas area represent entities: users, recommender agents or tags. Yellow circles represent individual talks, and the bubbles that involve them represent clusters of talks.

In Figure 2, two users are shown (P Brusilovsky and L Aroyo), as well as suggestions of the tag-based and content-based recommender agents. The cluster map visualization enables users to explore relationships between items that are associated with different entities (i.e., recommended by an agent, bookmarked by a user, tagged with a tag). For instance, a user can see which users have bookmarked a talk that is suggested to them by a recommender agent by exploring the intersection of the agent and a specific user. Users can arrange the different entities displayed in the canvas by dragging them with the mouse.

Finally, the rightmost panel shows the detailed list of talks. This can be a list of all the talks presented in the canvas area, or a subset of them related to the selected entity. For example, if a user clicks on a specific user name in the canvas area, the papers that the selected user has bookmarked are presented in the list. If a user clicks on a cluster (for example, the cluster showing talks that were bookmarked by a user and a specific agent) the list of these talks is presented.

EVALUATION

We evaluated TalkExplorer at two conferences where CN3 was used as the main conference assistance system: ACM Hypertext 2012 conference in June 2012 (HT 2012) and User Modeling, Adaptation, and Personalization conference in July 2012 (UMAP 2012). Both evaluations were performed with attendees of respective conferences using real conference data (i.e., using actual talks and schedules and bookmarks, tags and ratings of the conference participants). Users were asked to explore conference talks using the visualization provided by TalkExplorer. As explained in the previous section, the visualization provided access to the content-based and tag-based recommender agents and allowed to explore talks bookmarked by users or tagged with community tags.

Participants

In the HT 2012 evaluation, fourteen users participated in a controlled experiment at the conference. We inquired about the number of Hypertext conferences participants have attended, as well as their knowledge of recommendation and visualization techniques, respectively. The latter were rated on a five point Likert scale. On average, participants attended 1.5 conferences in the past (std. deviation 0.732). Most of the participants have knowledge about or expertise with visualization techniques (average 4.285, std. deviation 0.7). In addition, familiarity with recommendation techniques is high – although less extensive than with visualization techniques (average 3.7, std. dev. 0.8).

Seven participants of the UMAP 2012 conference participated in the second study of our visualization. They had a high familiarity with visualization techniques (mean 4.2, std. deviation 0.76) and a relatively high familiarity with recommendation techniques (mean 3.7, std. deviation 0.95). On average, participants attended 2 UMAP conferences (std. deviation 1.5).

Tasks

We asked users to complete three tasks:

1. In the first task, they were asked to find a new relevant talk to attend by exploring talks that users in their neighborhood bookmarked (Task 1 – T1)
2. In the second task, subjects had to find a new relevant talk by exploring the content-based and tag-based recommender agents (Task 2 – T2).
3. In the third task, they were asked to find a new relevant talk by exploring the tags (Task 3 – T3).

Data collection

Data was collected in two ways. The think aloud protocol was used during the experiment to facilitate the collection of relevant feedback from participants. We recorded the screen and voice of participants using Camtasia Studio [8]. Afterwards, participants were asked to fill out a survey inquiring about their needs at a conference and the usefulness of the visualization to address these needs.

Results

To assess the value of interactive multi-prospect visualization offered by TalkExplorer, we have analyzed the way in which users *explore* and *use* the visualization. In the remainder of this section, we refer to selectable users, agents and tags as *entities* in the visualization. Papers or talks associated with these entities are referred to as *items*. We refer to *intersections* of entities when multiple entities were selected at the same time and their common items were explored.

We measured the *effectiveness* and *yield* of different combinations of entities to gain insight in the relative success rate of different combinations of entities to find relevant items.

Effectiveness measures how frequently a specific combination type produced a display that was used to bookmark at least one interesting item. It is calculated as the number of cases where the exploration of this combination type resulted in a bookmark, divided by the total number of times this combination type was explored. For instance, the set of items of a related user was explored 75 times by all participants. 23 of these sets were used to bookmark a new item. Thus, the effectiveness of exploring the set of items of a specific user is $23/75=31\%$. The number of item sets explored and the item sets used to bookmark a relevant talk, as well as the *effectiveness*, are presented in Figure 5.

In addition, we counted the number of items in the sets where the selection was made to check *yield* (productivity) of different kinds of sets. The yield of a specific combination type was measured by summing up the total number of selections made from each combination type, divided by the total number of items that were shown in the combinations where the selection was made. In other words, yield measures a chance of a random item shown in a specific combination type to be useful (i.e. bookmarked by the user). Yield results are presented in Figure 6.

Figure 3 presents an example where the active user, E Duval, used the intersection of two other users as a basis to find a relevant item. In this example, E Duval used the set of 8 items in the intersection of two other users (P Brusilovsky and D Parra) to find an item. The yield indicates the number of selections made from a specific set of entities divided by the sum of the number of items in this set (8 in the example presented in Figure 3).

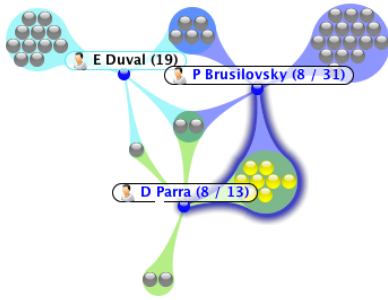


Figure 3: Exploring an intersection: items bookmarked by 2 users but not yet bookmarked by the active user

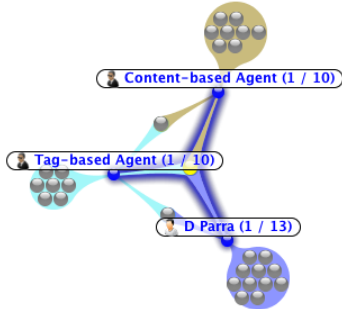


Figure 4: Talk in intersection of agents and another user

For the intersection of 2 users, twice an item was selected out of 2 items and twice an item was selected out of a set containing 1 item. We have yield or probability of selection of $(1+1+1)/(2+2+1+1)=0.66$.

Task 1

In the first task (T1), users were asked to find a relevant talk by exploring bookmarks of users in their neighborhood. Results are presented in Figure 5 and Figure 6. The set of items of one specific user with whom the active user is related was explored 75 times by all participants. 23 of these sets were used to bookmark a new item. The number of items in these 23 sets is 276. Thus, the effectiveness of $23/75 = 31\%$ (first top bars in Figure 5) and the yield is $23/276 = 8\%$ (first bar in Figure 6).

Fifteen users explored intersections of two related users focusing on talks that they have not yet bookmarked (as illustrated in Figure 3). This kind of set was used to bookmark a talk 4 times (effectiveness = 27%) and the sum of items in the used sets was 6 (yield = 66%).

Talks in the intersection of three or four other users were explored 12 and 6 times and used 5 and 3 times, respectively (effectiveness of 42% and 50%). The number of items in the selection set was 14 and 7, respectively (yield of 37% and 43%). As we can see from this data, the general trend is clear: the sets that allow users to explore the overlap of several prospects are both more effective and have higher yield. Moreover, there is a general tendency for effectiveness and yield to increase when more entities are used in the selection process. Small fluctuations within the general trend can be explained by the small sample.

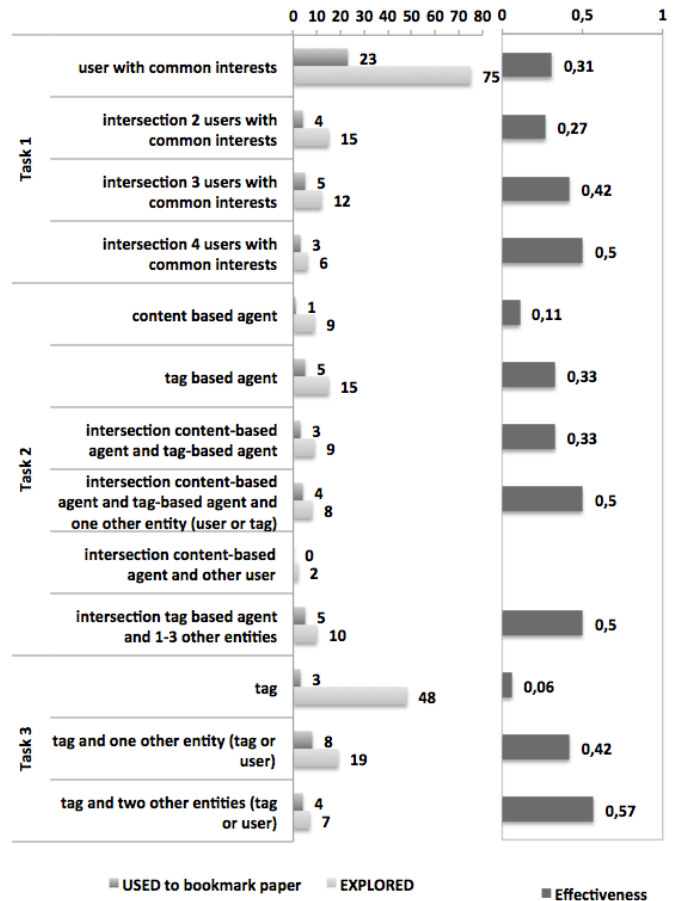


Figure 5: Summary of actions explored, used to bookmark a paper, and effectiveness of those actions.

Task 2

In the second task (T2), users were asked to find a relevant talk by exploring the output of recommender agents (a content-based and a tag-based agent). Results are presented in Figure 5 and 6, the middle set with 6 possible actions.

One out of nine users found a relevant talk by exploring suggestions by the content-based agent that were not related to any other entities on the screen (effectiveness=11%, yield=11%). Five out of 15 users found a relevant talk by exploring suggestions of the tag-based agent. Three out of nine users found relevant items by exploring the intersection of agents (i.e. talks that were suggested to them by both the content-based and the tag-based agent). Four out of eight users found relevant items by exploring the intersection of the agents with another entity. Figure 4 presents such an example.

The set of items of the content-based agent in combination with another user was explored twice, but not used to find a relevant item. The tag-based agent in combination with one or more entities was successful in 50% of the cases. The results presented in Figure 5 indicate the same trend: a higher number of entities (prospects) involved in the intersection increases the effectiveness and the yield of the resulting set.

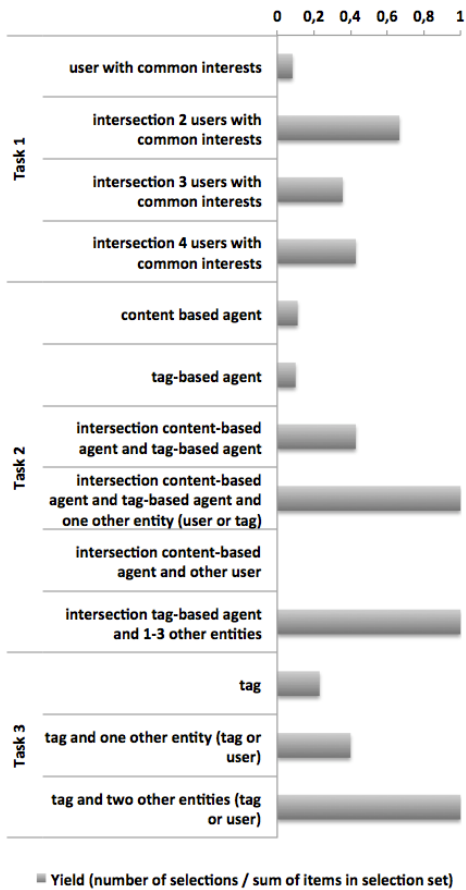


Figure 6: Summary of yield results

Task 3

In the third task (T3), we asked users to find interesting talks by exploring tags that were added by users to talks. Results are presented in the bottom set of 3 actions in Figure 5 and Figure 6.

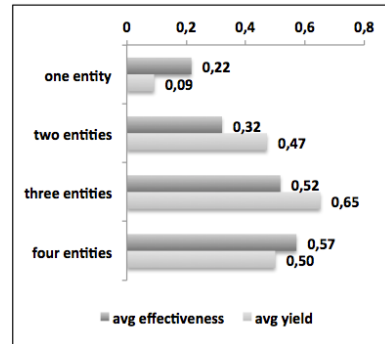
As the data shows, using a single tag prospect (i.e., exploring items related to one selected tag) results in the lowest effectiveness registered in the study – as only three users were able to find a relevant item (effectiveness 6%). The sum of the number of items in the set when a selection was made was 13 (yield 3/13=23%).

Combining a tag prospect with one or more additional entities was more effective. 19 users explored the combination of a tag with another entity and 8 users used this intersection to bookmark an item (effectiveness=42%, yield=40%). A tag in relation to two other entities was effective in 57% of the explored cases. Some users indicated that they particularly liked this functionality – as this allows them to retrieve specifically items of their topic of interest from users they know or who have a high reputation in the field.

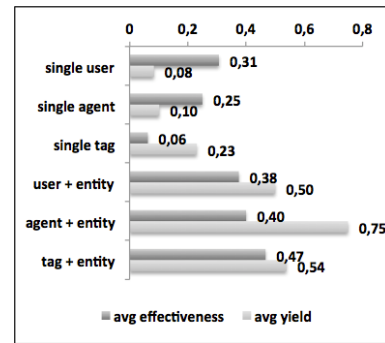
Summary of results

Effectiveness and yield results are summarized in Figure 7. Overall, these results indicate that effectiveness of an explored set increases once more entities are integrated.

Similar trends are observed when we look at yield. These results indicate that the probability of selecting an item from a set increases if more entities are overlapped.



(a)



(b)

| | Significance effectiveness | Significance yield |
|----------------------------|----------------------------|--------------------|
| multiple vs. one entity | 0.003469 | 0 |
| user vs. (user + entity) | 0.5934 | 0 |
| agent vs. (agent + entity) | 0.3417 | 0 |
| tag v/s (tag + entity) | 0 | 0.1 |

(c)

Figure 7: Summary of effectiveness and yield results

In Figure 7.a, we can see that the effectiveness and yield increases from 22% and 9% when a single entity is used to 57% and 50% when 4 entities are used. In Figure 7.b, we can see that yield increases from 8% when the set of items of one user is used, 10% when the set of items of one recommender agent is used and 23% when the set of items of one tag is used to 50% when the intersection of a user and another entity used, 75% when the intersection of an agent with another entity is used and 54% when a tag is used in combination with another entity. There is a significant difference in effectiveness when a tag is used in combination with another entity (see Figure 7.c). Effectiveness differences between users and agents in combinations with other entities are not significant. In general, effectiveness and probability of item selection is significantly higher when multiple entities are used (first

row Figure 7.c). These results illustrate that enabling end users to explore interrelationships between prospects (sets of items in the overlap of entities) increases the probability of finding a relevant item.

The results also allow us to make several interesting observations. First, it is interesting to note that the least effective kind of set is the set of items related to exactly one tag (6% effectiveness). Incidentally, this is the only option to use tags for item exploration offered in many tag-based system. As shown by our data, the systems that do not allow exploring items related to combinations of several tags are not doing a good service to their users.

In contrast, exploring a prospect of a single related user in relation to the target user is a relatively effective approach – almost 1/3 of explored combinations produced bookmarks. It shows that the prospects of human users are more valuable (and trustable) than prospects offered by tags.

Overall, these results illustrate the added value of enabling users to combine social and personalization features of recommenders as a basis to increase the relevance and value of suggested items.

Questionnaire results

To collect user subjective feedback, we used a questionnaire. Its goal was to inquire about needs and issues that users have when attending a conference and the extent to which our visualization addresses those needs and issues. This questionnaire was used to collect some preliminary feedback and included only a few questions we considered important at this stage. We elaborate in the next section on the use of standardized questionnaires.

Figure 8 presents the results of our questionnaire on a five point Likert scale. The first column of bar charts presents answers to questions that inquire about the importance of issues and needs at a conference. The second column presents to which extent TalkExplorer addresses these needs and issues. The results indicate that conference attendees perceive ‘finding relevant talks at a conference’ as an important need (median 5) and that TalkExplorer addresses this need in a good way (median 4). ‘Being aware which talks my friends are attending’ is also perceived as important (first column) and most participants indicate that this need is addressed by TalkExplorer (second column). ‘Being aware which talks people I follow are attending’ is considered less important and is also not so well addressed by our tool (median 3). ‘Finding other users who have similar interests’ is important for many participants (median 4) and is addressed well by TalkExplorer (median 4). ‘Gaining insight into why talks are recommended’ is important for many, but not all, participants. This need is addressed well by TalkExplorer according to most, but again not all, participants (median 4).

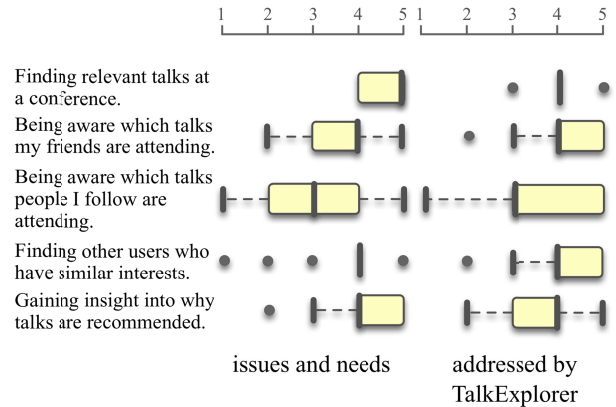


Figure 8: questionnaire issues and needs

In addition, we inquired about perceived usefulness of the visualization to support transparency and exploration. These results are presented in Figure 9 and indicate that participants perceive the visualization as useful because it gives more insight than a plain list of talks (median 4). In addition, most participants liked our idea of adding agents in parallel to real users as a means to find interesting talks (median 4) – among others to compare relevancy of recommendations (median 5).

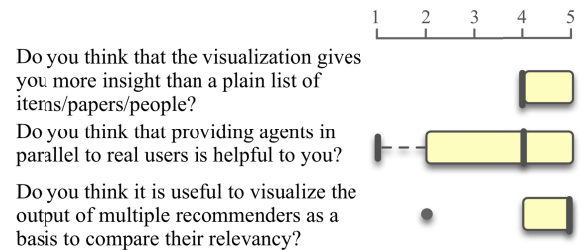


Figure 9: perceived usefulness

Think-aloud data analysis

An analysis of think-aloud data revealed some usability issues with the visualization and some additional insights into the usefulness of combining multiple relevance prospects. This data indicates that it was not clear why 15 users are shown in the neighborhood of a user. Participants asked why some users are shown by default and how these users are selected. Explaining the rationale of the selection of these users, such as the similarity match based on common bookmarks, can potentially resolve this issue. In addition, it would be interesting to explore how users can navigate through interest data of all users – while still distinguishing potentially more relevant bookmarks from users with common interests. People recommendation as opposed to recommendation of papers and talks was another suggestion from a participant.

Moreover, when users removed themselves from the visualization and explored bookmarks of other users only, it was not clear for them which of these items were already in their schedule. We resolved this issue for the UMAP evaluation by indicating in the right side panel whether a

talk is already bookmarked (see Figure 2). Using different colors in the canvas could be a second solution.

Finally, participants remarked that they particularly liked combining tag prospects with agents or user prospects – as this functionality enables them to identify talks of a specific topic of interest from a set of recommended items or items from a user. These remarks illustrate the perceived usefulness of combining multiple prospects and are also reflected in the data presented in Figure 7. These data indicate that effectiveness of individual tags is significantly lower than a tag in combination with another entity.

DISCUSSION

Evaluation results presented in the previous section indicate that ability to combine visually multiple relevance prospects (i.e., bookmarks of users, suggestions of recommender agents and talks marked by specific tags) is perceived as useful to increase the relevance and value of recommendations. Results from our questionnaire are generally positive and indicate that participants value our visualization as a way to gain insight into why talks are recommended. In addition, they indicate that such a visualization gives more insight than a typical ranked list of recommendations.

Interaction patterns indicate that users often explore relationships between entities to find relevant items. Although items of one specific user were explored most often to find a relevant item (see Figure 5, first row), the effectiveness and probability of selecting an item is significantly lower than with intersections of multiple entities. Moreover, interrelating tags with other entities increase their effectiveness significantly (as summarized in Figure 7).

While these results illustrate the usefulness of visualizing and combining recommendations, tags and users, there are several limitations to this study that should be articulated and addressed in follow up studies. First, we asked users to explicitly explore users in their neighborhood, recommender agents and tags in three separate tasks. While results of these tasks give some interesting insights in the usefulness of these entities and the way users interacted with additional entities during these tasks, we cannot draw strong conclusions about the relative effectiveness of tags, users and agents in this way. The order of the tasks may have had an influence on the effectiveness of these entities. Moreover, we explicitly asked to explore these entities. In a follow up study, we are capturing interactions of users with the visualization in an open setting where users are free to explore various entities. Such a study and analysis of interaction patterns will yield more accurate data with respect to relative effectiveness of tags, users and agents. Second, participants at both UMAP 2012 and HT 2012 are highly knowledgeable on topics of recommendation and visualization and so they might not be representative of the general conference audience. Third, the questionnaire we used was a preliminary set of questions we assembled to

gather initial feedback. In a follow up study, we plan to conduct more elaborate surveys based on standardized questionnaires to assess accuracy, diversity and novelty. Such studies are necessary to gain insight in other potential benefits that this interface can offer to end users.

CONCLUSION AND FUTURE WORK

In this paper we have presented and discussed the results of two studies involving conference attendees that select relevant talks by making use of the TalkExplorer visualization tool, embedded in CN3. TalkExplorer allows users to explore items (talks) by combining different entities (users, tags, recommender agents). After analyzing the users' behavior and their answers in a survey, we highlight three results.

First, to the best of our knowledge, the recommender systems literature offers neither studies that represent recommender algorithms as agents nor systems that let users interact with and combine the output of recommendation algorithms with social prospects. Our results indicate that this can be a significant contribution to the area of user interaction in recommender systems, and we plan to expand our research in this area to other domains beyond paper suggestions. We also plan to investigate different ways that users can apply to control and combine the output of several recommendation methods.

Second, results of our questionnaire indicate that end users perceive our interactive visualization. However, it is still an open question what is the impact of *personal characteristics* (such as expertise, and the users' visual processing fit) on the usefulness of our approach, and how different ways to manipulate the interface that the system offers its users might influence user performance and satisfaction.

Finally, users show a better performance finding relevant items – in terms of the number of actions needed to discover relevant items – by foraging for additional evidence. In TalkExplorer, users accomplish that by intersecting the preferred items of several users, items associated with tags, and agent recommendations. This opens research opportunities for studying further the benefits of interactive recommender interfaces in terms of user trust in the systems, transparency and satisfaction.

ACKNOWLEDGMENTS

The authors would like to thank Joris Klerkx for making available the source code of his framework code. In addition, we thank all the participants of the user studies for their participation and useful feedback. Research of Katrien Verbert was supported by postdoctoral fellowship grant of the Research Foundation – Flanders (FWO). The authors would like to thank the anonymous reviewers for their valuable feedback and suggestions.

REFERENCES

1. Aduna clustermap. <http://www.aduna-software.com/technology/clustermap>

2. Baeza-Yates, R.A. and Ribeiro-Neto, B. Modern Information Retrieval. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1999
3. Bennett, J. and Lanning, S. The Netflix Prize. *Communications of the ACM*, 52, 10 (2007), 8 pages. Citeseer.
4. Blei, D.M., Ng, A.Y. and Jordan, M.I. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, (2003), 993-1022
5. Bostandjiev, S., O'Donovan, J. and Höllerer, T. TasteWeights: a visual interactive hybrid recommender system. In *Proceedings of the sixth ACM conference on Recommender systems (RecSys '12)*. ACM, New York, NY, USA (2012), 35-42.
6. Burke, R. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User Adapted Interaction*, 12, 4 (2002), 331-370.
7. Burke, R. Evaluating the dynamic properties of recommendation algorithms. In *Proceedings of RecSys10: fourth ACM conference on Recommender systems*, ACM Press New York, NY, USA (2010), 225-228.
8. Camtasia Studio:
<http://www.techsmith.com/camtasia.html>
9. Chuah M. Dynamic aggregation with circular visual designs. In *Proceedings of the IEEE Symposium on Information Visualization (1998)*, 35-43
10. Gretarsson, B., O'Donovan, J., Bostandjiev, S., Hall, C. and Höllerer, T. SmallWorlds: Visualizing Social Recommendations. *Comput. Graph. Forum*, 29, 3 (2010), 833-842.
11. Gou, L., You, F., Guo, J., Wu, L. and Zhang, X. SFViz: interest-based friends exploration and recommendation in social networks. In *Proceedings of the 2011 Visual Information Communication - International Symposium (VINCI '11)*. ACM, New York, NY, USA (2011), 10 pages.
12. Herlocker, J.L., Konstan, J.A. and Riedl, J. Explaining collaborative filtering recommendations. In *Proceedings of CSCW00: ACM conference on Computer supported cooperative work*, ACM Press (2000), 241-250
13. Herlocker, J.L., Konstan, J.A., Terveen, L.G. and Riedl, J.T. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22, 1 (2004), 5-53.
14. Hu, Y., Koren, Y. and Volinsky, C. Collaborative filtering for implicit feedback datasets. In *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*. IEEE Computer Society, Washington, DC (2008), 263-272.
15. Kammerer, Y., Nairn, R., Pirolli, P. and Chi, E. H. (2009). Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA (2009) 625-634.
16. Klerkx, J. and Duval, E. Visualising social bookmarks. *Journal Of Digital Information*, 10, 2 (2009), 40 pages. Multimedia Research Group, U. of Southampton.
17. Knijnenburg, B.P., Reijmer, N.J.M. and Willemsen, M.C. Each to his own: how different users call for different interaction methods in recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems (RecSys '11)*. ACM (2011), 141-148.
18. Knijnenburg, B.P., Bostandjiev, S., O'Donovan, J. and Kobsa, A. Inspectability and control in social recommenders. In *Proceedings of the sixth ACM conference on Recommender systems (RecSys '12)*. ACM, New York, NY, USA (2012), 43-50.
19. Linden, G., Smith, B. and York, J. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, 7, 1 (Jan/Feb 2003), 76-80.
20. Manning, C.D., Raghavan, P. and Schtze, H. Introduction to Information Retrieval, Cambridge University Press, New York, NY, 2008
21. Morita, M. and Shinoda, Y. Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval. In *SIGIR 1994: Proceedings of the 17th Annual International ACM SIGIR Conference*. Springer-Verlag New York, Inc., New York, NY, USA (1994), 272-281.
22. Nageswara, R.K. and Talwar, V.G. Application domain and functional classification of recommender systems a survey. *DESIDOC Journal of Library and Information Technology*, 28, 3 (2008), 17-35
23. O'Donovan, J., Smyth, B., Gretarsson, B., Bostandjiev, S., and Höllerer, T. PeerChooser: visual interactive recommendation. In *Proceedings of the twenty-sixth conference on Human factors in computing systems (CHI '08)*. ACM, NY, USA, 2008, 1085-1088.
24. Parra, D., Jeng, W., Brusilovsky, P., Lopez, C. and Sahebi, S. Conference Navigator 3: An Online Social Conference Support System. Poster at UMAP 2012. Montreal, Canada.
25. Parra, D. and Brusilovsky, P. Collaborative filtering for social tagging systems: an experiment with CiteULike. In *Proceedings of the third ACM conference on Recommender systems (RecSys '09)*. ACM, New York, NY, USA (2009), 237-240.
26. Pazzani, M.J. and Billsus, D. Content-Based Recommendation Systems. In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, *The Adaptive Web*, volume 4321 of LNCS (2007), 325-341. Springer-Verlag, Springer.
27. Riedl, J. and Dourish, P. Introduction to the special section on recommender systems. *ACM Trans. Comput.-Hum. Interact.* 12, 3 (Sept.2005), 371-373.

28. Tintarev, N. and Masthoff, J. Designing and Evaluating Explanations for Recommender Systems. *Recommender Systems Handbook*, (2011), 479-510
29. Tonkin, E., Corrado, E. M., Moulaison, H. L., Kipp, M. E. I., Resmini, A., Pfeiffer, H. D. and Zhang, Q. Collaborative and Social Tagging Networks. *Ariadne*, 54. (2008).
30. Vig, J., Sen, S. and Riedl, J. Tagsplanations: explaining recommendations using tags. In *Proceedings of the 14th international conference on Intelligent user interfaces (IUI '09)*. ACM, New York, NY, USA (2009), 47-56.
31. Zhao, S., Zhou, M.X., Yuan, Q., Zhang, X., Zheng, W., and Fu, R. Who is talking about what: social map-based recommendation for content-centric social websites. In *Proceedings of the fourth ACM conference on Recommender systems (RecSys '10)*. ACM, New York, NY, USA (2010), 143-150.