Providing Control and Transparency in a Social Recommender System for Academic Conferences

Chun-Hua Tsai
University of Pittsburgh
135 N Bellefield Ave
Pittsburgh, PA 15260
cht77@pitt.edu

Peter Brusilovsky
University of Pittsburgh
135 N Bellefield Ave
Pittsburgh, PA 15260
peterb@pitt.edu

ABSTRACT

A social recommender system aims to provide useful suggestion to the user and prevent social overload problem. Most of the research efforts are spent on push high relevant item on top of the ranked list, using a weight ensemble approach. However, we argue the "learned" static fusion is not enough to specific contexts. In this paper, we develop a series visual recommendation components and control panel for the user to interact with the recommendation result of an academic conference. The system offers a better recommendation transparency and user-driven fusion through recommended sources. The experiment result shows the user did fuse the different recommended sources and exploration patterns among tasks. The post-study survey is positively associated with the system and explanation function effectiveness. This finding shed light on the future research of design a recommender system with human intervention and the interface beyond the static ranked list.

KEYWORDS

Social Recommendation; Explanation; Transparency; User Control

1 INTRODUCTION

The ranked list is the most distinct and visible feature of information retrieval and recommender systems. A lot of research efforts have been spent to push relevant items as high as possible on this list, while several measures have been created to assess the effectiveness of such ranking systems. However, what could be done if a particular context offers more then one important aspect of relevance, with each aspect requiring a different ranking? For example, in a personalized information retrieval system, search results could be ranked by their relevance to the query or their similarity to the user profile [1]. In a social system for academic conferences [6], recommended attendees could be ranked by their social distance, the similarity of their past publications, or the similarity of their interests, as reflected by shared bookmarks. The current way to resolve this problem in the field of recommender systems is to use

ensembling; namely, a weighted combination of two or more ranking approaches. Optimal weights for such an ensemble could be found using some ensemble training method, such as Breiman's stacked regression [5], and then used to fuse the sources within each context.

The problem with the "learned" static fusion is that in different situations, a user might prefer a particular aspect of relevance or a specific combination, and as a result, an average "best" fusion will never be adequate. For example, when searching for unknown but like-minded conference attendees, a user could obtain the best results by fusing an inverted social distance ranking with a regular interest similarity. Research on retrieval and recommender system interfaces suggest resolving this problem by engaging the user in selecting the best approach or fusion of approaches. Several projects in both fields demonstrated that users could learn and efficiently use these selection and fusion interfaces to obtain superior results [1, 12, 16].

Our paper expands the current work on user-controlled multi-aspect recommendations in two directions. First, in contrast to earlier work focused on item recommendations, we want to explore controllable recommendations of *people* as social and academic contacts. Second, we want to concentrate on an unexplored aspect of this research: explanation and transparency. The need to offer better transparency and explanations of recommendations is now generally recognized in the field [20]. However, this explanation becomes especially important for interfaces with a user-driven fusion of recommended sources. To combine individual sources in a meaningful way, users need to have a solid understanding of why a recommended person has been ranked high or low, according to a particular aspect of relevance, as well as in the integrated ranked list.

In this paper, we present RelExplorer [1], a system for recommending and exploring co-attendees at an academic conference. The system uses three separate recommender engines that suggest the most relevant attendees in respect to social distance, the relevance of their past work, and the similarity of current interests. RelExplorer allows users to fuse rankings produced by these recommendation sources according to the current need, explore the obtained unified ranking, and receive an extensive explanation of ranking results. To assess the value of the user-driven fusion and the overall explanation functionality of RelExplorer, we conducted a user study at two international conferences. In the following sections, we present the design of RelExplorer, introduce our studies, and review the obtained results.

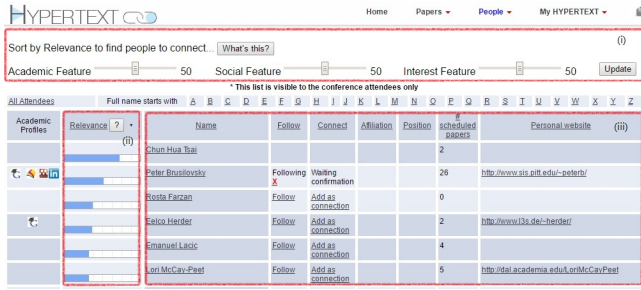[1]System Demo: http://halley.exp.sis.pitt.edu/cn3/portalindex.php

Figure 1: A screenshot of the RelExplorer system: (i) A control panel of three feature sliders. (ii) A ranking list of the personalized relevance score. (iii) The user profile information and social media functions from the Conference Navigator 3 System [6].

2 RELATED WORK

While early research on recommender systems mostly focused on ranking and prediction, it has been recently recognized that users will easily distrust even a perfect ranking if it lacks interpretability. To increase the overall level of user acceptance, [3, 16] have proposed the use of interactive recommendation interfaces with transparency and controllability in place of a static ranked list. Also, visualization techniques were explored to improve the comprehension of the recommendation result [8, 10]. Recent studies indicate the effectiveness of intelligent interfaces that support transparency, exploration, and controllability in various contexts, such as recommending conference talks [23], expert finding [9], people matching [7, 15, 17], and collaboration [2].

3 RELEXPLORER

3.1 System Design

Figure 1 shows a screenshot of the front-end view of the RelExplorer system. The view consists of three parts. 1) A control panel with three sliders that controls the fusion of three elementary recommender engines or *features*. The user can adjust the slider from 0-100 for Academic, Social, and Interest features, based on their current needs. The weight of each contributing engine in the fusion is determined by the selected weight, e.g. setting three features equally at 50 means that each component is weighted as 33.3%. 2) A set of bars show a fused relevance score, which is calculated as a linear combination of Academic, Social, and Interest features with selected weights. These features are discussed in more detail in section 3.2. 3) A basic user profile that includes name, affiliation and (for authors) titles of papers presented at the conference. Each name in the list is a link to the profile page that shows more personal information, along with visualized explanations of component rankings (see section 3.3 for details).

RelExplorer is embedded in the Conference Navigator System (CN3), a social support system for academic conferences [6]. The system has been used to support 37 conferences at the time of writing this paper. CN3 has 6,500 users, 6,398 articles, 11,939 authors, 28,590 bookmarks, and 1,336 social connections. To solve the cold start issue that occurs when users have no bookmarks

or social connections [21], we used the Aminer dataset [18]. This dataset includes 2,092,356 papers, 1,712,433 authors and 4,258,615 co-authorship. By combining the CN3 historical data and Aminer database, RelExplorer can produce necessary recommendations for CN3 users.

3.2 Recommendation Components

The RelExplorer uses three separate recommender engines that suggest co-attendees to meet on the basis of: 1) The similarity of past publications (Academic feature); 2) Social network distance (Social feature); and 3) Similarity of interests (Interest feature).

Academic Feature: The academic feature is determined by publication similarity between two attendees using cosine similarity [13, 22]. The function is defined as: $Sim_{Academic}(x, y) = (x \cdot y)/\|x\|\|y\|$, where x and y are word vectors for user x and y. For all attendees, we build documents assembled from titles and abstracts of their publications. We applied TF-IDF to create the document vector with a word frequency upper bound 0.5 and lower bound 0.01 to eliminate both common and rarely used words. We consider unigrams and bigrams to cover academic terms.

Social Feature: The social feature is calculated by collaborative network distance and neighbor similarity in the CN3 system and Aminer dataset. The goal is to generate the ranking of "most connected scholar" for the user. The function is defined as:
$Sim_{Social}(x, y) = (1 - \theta)(p) + (\theta)(cn)$, where p is the shortest path between user x and y; cn is the number of common neighbors of user x and y; and θ is the weighting ratio between two methods. We adopted the Depth-first search (DFS) method to calculate shortest path p [19] and common neighborhood (CN) [14] for neighbor overlapping similarity. The formula cn is $\Gamma(x) \cap \Gamma(y)$, where $\Gamma()$ indicates the neighbors of a given user x and y in two-hop degrees.

Interest Feature: The interest feature is determined by the data of co-bookmarked papers and co-connection authors in the CN3 system. The goal of this feature is recommending the attendees of a conference who share similar interest. The function is defined as $Sim_{Interest}(x, y) = (b_x) \cap (b_y) + (c_x) \cap (c_y)$, where b_x, b_y represent the paper bookmarking of user x and y; c_x, c_y represents the friend connection of user x and y.

3.3 Explanation Components

RelExplorer provides four explanation components to justify and explain ratings produced by recommendation components.

SocialViz: Social similarity is explained using topology-based visualization that lets users understand the connection of any conference attendees (Figure 2a). This tool uses the interactive force layout project of D3.js [4] to show the shortest path to connect two users generated with the DFS method.

SocialBubble: This component (Figure 2b), uses an interactive bubble menu from d3.js [4] to show the common coauthorship neighborhood between two users. The middle circle shows an author who has the highest coauthor overlapping rate. The system will pick randomly if there is no single one-degree coauthor between the two users.

Publications: The attendees of the conference are usually scholars with a list of publications. This list is a useful way to become quickly familiar with a user of interest. To show user publications within
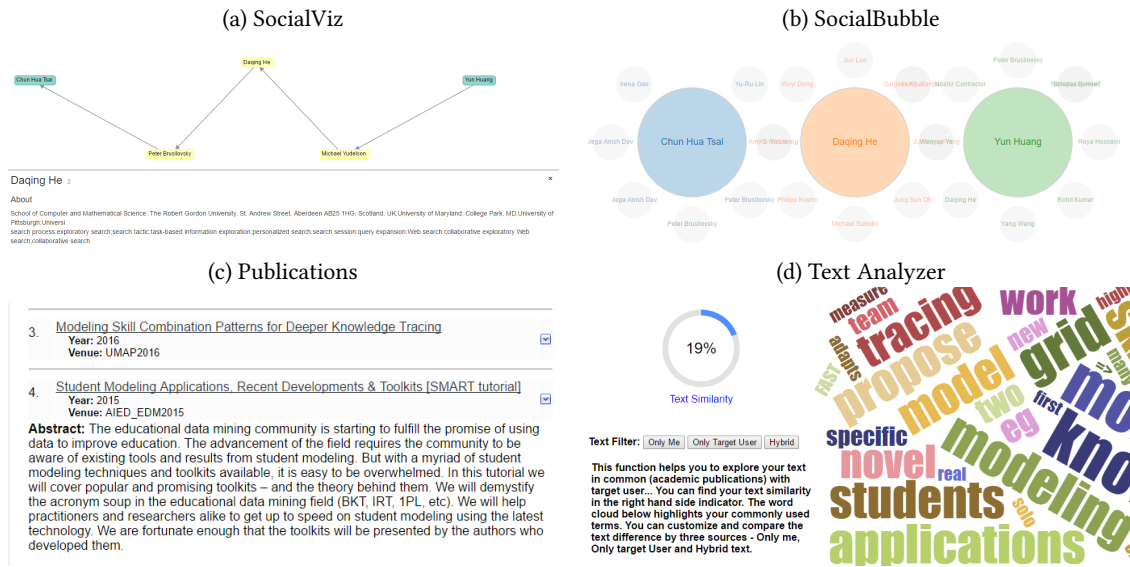
(a) SocialViz

(b) SocialBubble

(c) Publications

(d) Text Analyzer

Figure 2: (*a*)*SocialViz* : **a topology-based visualization based on an academic collaboration network. The green node represents the origin/target scholar and the yellow node is the path between them. This figure shows a reasonable path that goes through three scientists.** (*b*)*SocialBubble* : **the blue bubble and green bubble represent the origin and target user, respectively. The middle orange bubble is the co-author who has the highest number of co-authored paper with both the origin and the target scholar. The minor bubble represents coauthors of participating scholars. This tool helps to determine any indirect relationships between the origin and target users.** (*c*)*Publications* : **a global view of the publication list of the target scholar generated from both the conference data and the external Aminer dataset.** (*d*)*TextAnalyzer* : **this tool helps to understand the level of content similarity between the publications of the origin and target scholars. A document similarity percentage bar and word cloud are provided. The word cloud can be customized to individual/hybrid mode to compare the difference between the words.**

and beyond the current conference, we use historic CN3 data [6] and the Aminer dataset [18]. This publication tool (Figure 2c) shows the publication list of a user with highlighted papers that appears in the same conference series as the current conference.

TextAnalyzer: This tool visualizes text similarity between two attendees (Figure 2d). It is shown as a percentage of the text similarity between two users and a word cloud [4] to explore and compare the most popular words in their publications.

4 USER STUDY

To assess the value of user-driven fusion and explanation functionality of RelExplorer, we conducted a user study at two international conferences: ACM Hypertext (HT) 2016 and ACM Conference on User Modeling, Adaptation and Personalization (UMAP) 2016, both held in Halifax, Canada. There were 65 attendees at HT and 115 attendees at UMAP. All attendees at both conferences received a CN3 account by email before the event date. Conference participants and authors were encouraged to use the system before the meeting began through an official email from the conference organizers.

4.1 Setup

For the user study, which was a controlled experiment at the conference venue, we recruited 16 attendees (5 female and 11 male) from both conferences; twelve were from HT and four were from UMAP. Half of them were aged 20 to 29 and the other half were aged 30 to

39. Among participants, there were 12 Ph.D. students, two master's students, and two junior faculty members. At the beginning of the study, we asked the participants to report their relevant experience using a five-point scale. On average, most of the participants had high confidence in using recommender systems (average 3.81 with a standard deviation of 1.04). Their background knowledge about recommender systems was relatively high (average 4.37 with a standard deviation of 0.61). Following that, we asked the participants to complete three simple training tasks and two search testing tasks. These tasks are described below. User actions performed in CN3 while completing the tasks were logged and timed.

Training: 1) Set up a recommendation factor weighting, based on your preferences, by using the control panel on top of the author page; 2) Sort the authors by relevance from high to low; 3) Click the top ranking author and review the information at different tabs.

Testing: *Task 1: Find Known Attendees* (a) Find two conference attendees you already know; (b) decide whether you need to follow each of them or connect to them in the system; (c) examine information about these participants to find out how these two people can help to establish new connections at this conference. *Task 2: Explore Unknown Attendees* (a) Find two conference attendees who you don't yet know in person but whom you are interested to meet and talk with; (b) decide whether you need to follow each of them or connect to them in the system; (c) examine information about these attendees to find out who could introduce you to them, or how you could introduce yourself to attract interest.
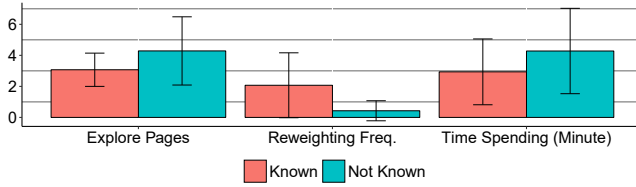
Figure 3: Experiment Measurement: a comparison of test value, included explore pages, re-weighting frequency and time spending between tasks.

4.2 Log Analysis

The log analysis (Figure 3) focused on comparing the number of explored user pages, the frequency of feature re-weighing in the attendee list, and the amount of time spent. We applied the Wilcoxon signed rank test with continuity correction to test the difference between tasks.

Re-weighting Frequency At average, the users applied 1.25 feature re-weighing (SD=1.73) while solving two user-finding tasks. This provides some evidence that weighing was useful in the context of given tasks. There was a significant effect for re-weighting frequency (p=0.008) between tasks. The user tends to change the control panel more frequently during the first task.

Explored Pages The test result indicates that finding and exploring known people requires fewer clicks (M=3.07, SD=1.07) than finding and exploring unknown people (M=4.28, SD=2.19). The difference approached the borderline of significance (p=0.06).

Time Spent It takes less time to finish the finding known-people task (M=2.93, SD=2.12) than to finish the finding unknown-people task (M=4.27, SD=2.75). The difference shows a trend toward significance (p=0.09). The tester engages with the system longer when exploring the new connections. This data correlates with the number of explored pages. When these findings are taken together, it hints that finding known and unknown people are reasonably different tasks that might need different types of interface support.

4.3 Questionnaire

We asked all study participants to fill in a questionnaire after the experiment. The questionnaire assessed their experience using a five-point Likert scale. Table 1 summarizes user perceptions about system and explanation effectiveness. According to the survey, the feedback of system usability (M=3.93, SD=0.85), satisfaction (M=3.93, SD=0.68), and reuse prospects (M=4.12, SD=0.80) is relatively high. The quality of people recommendation (M=3.87, SD=0.8) and information variety (M=4, SD=0.51) of the system also received positive feedback. The participants indicated that RelExplorer provided sufficient (M=4, SD=0.63) and easily understand (M=4, SD=0.73) explanations for the user to explore the people of interest at the conference.

Questions 11-14 show feedback on specific visualization components. The Publication List and Text Analyzer are positively assessed (M=3.81, SD=0.98) by the participants. SocialViz received the highest score (M=4. 25, SD=0. 57), which means that a topology-style display was useful to explore the social relations. Meanwhile, the SocialBubble received the lowest score (M=3.68, SD=0.94), which

Table 1: Post-Study Questionnaire Survey

Question	Score
Q1: I become familiar with the system very quickly.	4.00 (0.89)
Q2: The information provided in different tabs on a person's page was sufficient for me to find interested people.	3.93 (0.85)
Q3: I will frequently use this system in future conference.	4.12 (0.80)
Q4: I like the people recommendation result by the system.	3.87 (0.80)
Q5: The recommended result contained a lot of variety (of the control panel) for me to explore.	4.00 (0.51)
Q6: I have fun when I am using the system.	3.81 (0.91)
Q7: I have to invest a lot of effort to obtain a useful result from the system.	2.43 (0.96)
Q8: The system has no real benefit for me.	1.87 (1.08)
Q9: The information provided for the explanation was sufficient for me to explore the interested people.	4.00 (0.63)
Q10: I found the explanation is easy to understand.	4.00 (0.73)
Q11: ScholarViz Tab helps me to understand the people recommendation result.	4.25 (0.57)
Q12: ScholarBubble Tab helps me to understand the people recommendation result.	3.68 (0.94)
Q13: Publication List Tab helps me to understand the people recommendation result.	3.81 (0.98)
Q14: Text Analyzer Tab helps me to understand the people recommendation result.	3.81 (0.98)

indicates that this visualization was less useful in understanding the recommendation results.

5 DISCUSSION

In this paper, we presented a user-controlled social recommender system for exploring social contacts at academic conferences, which included several explanation components. The system offers a user-driven fusion of three recommender engines and a better level of transparency into the recommendation process. We conducted a user study to assess the value of user-driven fusion and explanation functionality. The experiment results show that attendees explored a range of different fusion settings when solving realistic attendee-exploration tasks. We also observed that more pages were explored and more time was spent when exploring the new social contacts. However, it is likely that when finding known attendees, participants rely on name search rather than on recommendation and explanations alone. The post-study survey shows positive feedback for system components and explanation effectiveness. The user feedback provides evidence that the system is useful for the user to explore the social contacts at a conference venue. Also, there is evidence that the explanation components helped the users to correctly interpret the recommendation results. In particular, the topology-style explanation received higher scores than both the text- and network-based explanations.

This study provided some evidence that the "best" fusion of recommended sources varies among users and tasks. It calls for a human-in-a-loop recommender system that combines user collaboration in helping and arranging information for solving a particular task [11]. However, in these systems, users need to have a good understanding of the recommended results and their component relevance aspects. It brings the challenges of designing an interface with transparency, explanation, and controls of the recommender system for user interaction. In future works, we plan to develop an interface that goes beyond the ranked list, and that will support diversity exploration and interaction among multiple aspects of relevance.

REFERENCES

[1] Jae-wook Ahn, Peter Brusilovsky, Daqing He, Jonathan Grady, and Qi Li. 2008. Personalized Web Exploration with Task Models. In *the 17th international conference on World Wide Web, WWW '08*. ACM, 1–10.

[2] Nesserine Benchettara, Rushed Kanawati, and Céline Rouveirol. 2010. A supervised machine learning link prediction approach for academic collaboration recommendation. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 253–256.

[3] Svetlin Bostandjiev, John O'Donovan, and Tobias Höllerer. 2012. TasteWeights: a visual interactive hybrid recommender system. In *Proceedings of the sixth ACM conference on Recommender systems*. ACM, 35–42.

[4] Mike Bostock. 2016. D3. js-Data-Driven Documents (2016). *URL: https://d3js. org* (2016).

[5] Leo Breiman. 1996. Stacked Regressions. *Machine Learning* 24, 1 (July 1996), 49–64. DOI:http://dx.doi.org/10.1023/a:1018046112532

[6] Peter Brusilovsky, Jung Sun Oh, Claudia López, Denis Parra, and Wei Jeng. 2016. Linking information and people in a social system for academic conferences. *New Review of Hypermedia and Multimedia* (2016), 1–31.

[7] Jilin Chen, Werner Geyer, Casey Dugan, Michael Muller, and Ido Guy. 2009. Make new friends, but keep the old: recommending people on social networking sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 201–210.

[8] Niklas Elmqvist and Jean-Daniel Fekete. 2010. Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE Transactions on Visualization and Computer Graphics* 16, 3 (2010), 439–454.

[9] Shuguang Han, Daqing He, Jiepu Jiang, and Zhen Yue. 2013. Supporting exploratory people search: a study of factor transparency and user control. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*. ACM, 449–458.

[10] Chen He, Denis Parra, and Katrien Verbert. 2016. Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications* 56 (2016), 9–27.

[11] Andreas Holzinger. 2016. Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics* 3, 2 (2016), 119–131.

[12] Bart P. Knijnenburg, Svetlin Bostandjiev, John O'Donovan, and Alfred Kobsa. 2012. Inspectability and Control in Social Recommenders. In *6th ACM Conference on Recommender System*. 43–50. http://dl.acm.org/citation.cfm?id=2365966

[13] Christopher D Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to information retrieval*. Vol. 1. Cambridge university press Cambridge.

[14] Mark EJ Newman. 2001. Clustering and preferential attachment in growing networks. *Physical Review E* 64, 2 (2001), 025102.

[15] John O'Donovan, Barry Smyth, Brynjar Gretarsson, Svetlin Bostandjiev, and Tobias Höllerer. 2008. PeerChooser: visual interactive recommendation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1085–1088.

[16] Denis Parra and Peter Brusilovsky. 2015. User-controllable personalization: A case study with SetFusion. *International Journal of Human-Computer Studies* 78 (2015), 43–67.

[17] Luiz Pizzato, Tomasz Rej, Joshua Akehurst, Irena Koprinska, Kalina Yacef, and Judy Kay. 2013. Recommending people to people: the nature of reciprocal recommenders with a case study in online dating. *User Modeling and User-Adapted Interaction* 23, 5 (2013), 447–488.

[18] Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. 2008. Arnetminer: extraction and mining of academic social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 990–998.

[19] Robert Tarjan. 1972. Depth-first search and linear graph algorithms. *SIAM journal on computing* 1, 2 (1972), 146–160.

[20] Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (1 Oct. 2012), 399–439. DOI:http://dx.doi.org/10.1007/s11257-011-9117-5

[21] Chun-Hua Tsai and Peter Brusilovsky. 2016. A personalized people recommender system using global search approach. *IConference 2016 Proceedings* (2016).

[22] Chun-Hua Tsai and Yu-Ru Lin. 2016. Tracing and Predicting Collaboration for Junior Scholars. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 375–380.

[23] Katrien Verbert, Denis Parra, Peter Brusilovsky, and Erik Duval. 2013. Visualizing recommendations to support exploration, transparency and controllability. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 351–362.