Beyond the Ranked List: User-Driven Exploration and Diversification of Social Recommendation

Chun-Hua Tsai
University of Pittsburgh
Pittsburgh, USA
cht77@pitt.edu

Peter Brusilovsky
University of Pittsburgh
Pittsburgh, USA
peterb@pitt.edu

ABSTRACT
The beyond-relevance objectives of recommender systems have been drawing more and more attention. For example, a diversity-enhanced interface has been shown to associate positively with overall levels of user satisfaction. However, little is known about how users adopt diversity-enhanced interfaces to accomplish various real-world tasks. In this paper, we present two attempts at creating a visual diversity-enhanced interface that presents recommendations beyond a simple ranked list. Our goal was to design a recommender system interface to help users explore the different relevance prospects of recommended items in parallel and to stress their diversity. Two within-subject user studies in the context of social recommendation at academic conferences were conducted to compare our visual interfaces. Results from our user study show that the visual interfaces significantly reduced the exploration efforts required for given tasks and helped users to perceive the recommendation diversity. We show that the users examined a diverse set of recommended items while experiencing an improvement in overall user satisfaction. Also, the users' subjective evaluations show significant improvement in many user-centric metrics. Experiences are discussed that shed light on avenues for future interface designs.

Author Keywords
Social Recommendation; User Interface; User-Driven Exploration; Diversification; Diversity

INTRODUCTION
Recommending *people* within a social system is a challenging task. A traditional approach for offering recommendations is generating a single ranked list of relevant people that is adapted to the profile of the user requesting information. However, users may look for other people for a range of reasons; for example, they may wish to re-connect with an acquaintance or to find new friends with similar interests [6]. This diversity in user needs makes it difficult to generate a static ranked list that fits all cases.

IUI'18, March 7–11, 2018, Tokyo, Japan

ISBN 978-1-4503-4945-1/18/03…$15.00

DOI: https://doi.org/10.1145/3172944.3172959

A particular instance in which a single ranked list might not work well is in a parallel hybrid-recommendation system that fuses several recommendation sources. In this case, different sources might be preferred for various needs (i.e., social similarity could work best for finding known friends while content-based similarity could be used to find people with similar interests [6]). Several authors have argued that the best approach in this situation is to offer users the ability to control the fusion by choosing the contribution of various algorithms [6, 10] or data sources [3]. Providing a visual interface that makes the process of fusion more transparent - for example, by showing recommended sources and their overlaps as set diagrams [29, 44] - could further address this problem. However, the set-based approach that has previously been applied to visualize the results of controlled source fusion has limited applicability, since it ignores the strength of relevance within each of the sources (which is a continuous variable). In this paper, we attempt to overcome the limitations of set-based visual fusion by exploring two visual fusion approaches that reflect the continuously varying strength of relevance within each source while keeping the fusion process transparent.

When selecting a visual design for the transparent fusion of recommendation sources, we focused on better-informing users about the *diversity* of the recommended results. It has been previously demonstrated that a proper user interface could promote diversity in information exploration. For example, a diversity-enhancing interface evaluated in [15] led to higher user satisfaction than a simple ranked list interface. Several attempts to design a diversity-focused interface using a dimension reduction technique to present opinion similarity by latent distance have been presented in [46, 11, 35]. However, the clustering distance was not easily interpreted, and as a result, users were unable to make personalized judgments. It remains challenging to design a recommender interface in which the user can both perceive the diversity and be able to control the interface to filter the recommended items.

The project presented in this paper reports our experience with two different visual recommender interfaces. First, we proposed a recommender interface that explores the value of a two-dimensional scatter-plot visualization to present recommendations with several dimensions of relevance. In our context, the scatter plot interface was used to help users combine different aspects of relevance for recommended items, while providing inspectability to the users. Second, we proposed a recommender interface that enhances the fusion control function within a ranked list with meaningful visual encoding for multiple dimensions of relevance. The users can adjust the relevance weightings to customize the recommendation results, which provides the user with a greater level of control over their results.

The two interfaces were designed to explore the value of user-controllable and diversity-aware interfaces in a social recommender system. Each of the interfaces has been evaluated in a controlled field study in the target context. The results show that the new visual interfaces reduce exploration efforts for a set of realistic tasks, and also make the users more aware of the diversity of recommended items. Also, the users' subjective evaluation shows a significant improvement in many user-centric metrics. We further discussed the effects of the proposed interfaces on the users' experience with a diversity-enhanced social recommender system.

The main contribution of this paper is threefold. 1) we propose two interfaces that support the continuously controlled fusion of several relevance aspects with inspectability and controllability. 2) we provide evidence that the diversity-aware interface not only helps the user to perceive diversity but also helps the user to improve usability in the real world beyond simple relevance tasks. 3) finally, we discuss the user experience effects on proposed interfaces through a structural equation model analysis.

BACKGROUND

Users access a hybrid social recommender system for different reasons. A static ranked list may not be suitable for all scenarios, which creates the challenge of increasing system's controllability as well as making diversity in the recommendation items more evident [40]. However, not every user equally values diversity [1]. The level of diversity-seeking is an existing individual difference. The findings of [23] indicated the individual differences in the information-seeking process and the needs of designing a customized interface to fit different users. [4] argues for considering a "diverse conceptions of democracy" when we develop a diversity-enhancing tool or application. Furthermore, to only present different information may not cause users to interact with diverse content. If a user feels threatened by unfamiliar information, a reinforcing effect may happen to cause users to avoid interacting with the various content [21, 11]. Hence, a diversity-aware recommender system should consider aspects of both item and user diversity, but not decrease the overall levels of user satisfaction and system usability [36].

Many scholars have suggested different explanation functions to increase the inspectability of recommender system. The function provides the transparency that let users realize how the system works [37, 14]. The exposure of the recommendation process through visual interfaces can also increase the inspectability of the system [20]. Many different types of research have been done on this subject. For example, [42] provides recommendation visualization to increase the transparency of the recommender system. [44] provides a set-based visualization to let the user explore the desired recommendation items. Other researchers further indicated that the value of explaining interfaces could enhance user experiences. The explanation interface was associated with the perception of recommendation quality [37], gaining trust in the system [7] and experiencing the competence of the system [45]. The studies of [13, 27, 26] have all mentioned that providing a controllable interface in the social recommender system can increase overall user satisfaction. The authors adopted an interactive graphical interface to present a social recommender that enables control on an item or user-level preference in a collaborative recommender system.

Some previous studies have been conducted to solve issues of recommendation diversity through interface design. For example, adopting a visual discovery interface can increase the click-through rate (CTR) across different item categories in an e-commerce website [35]. The user can explore new or relevant products without the need for search queries. The key factor of the interface is to provide the user with the ability to control the filtering of the recommendation contents. [32, 46] proposed a user-controllable interface for users to interactively change the ranking or feature weighting for a better-personalized ranking. [15, 11] proposed interfaces to show the various recommendation results that promote the users' perception of the diversity of recommendations. The study of [41] demonstrated a more diverse exploration pattern when the user was adopting a two-dimensional interface, versus a standard ranked list.

There are studies tried to enhance the recommendation diversity through exploratory search interfaces. For instance, [28] introduced an interface for interactive search, which uses overlaying graphs representing different information sources. The studies [18, 19] proposed interfaces which can present multi-faceted information on map or touch screen to diversify and explore the generated recommendations. The studies of [46] adopted dimensionality-reduction techniques to project the multidimensional data in two or three dimensions for the purposes of visualization. SciNet interface [31, 12] recommends keywords spatially in an interactive visualization to help diversification in exploratory search. However, the nature of each item is expressed in polar coordinates rather by two independent axes.

FIRST ATTEMPT: SCATTER VIZ

In a hybrid recommendation context with multiple types of relevance, the traditional ranked list makes it hard for the user to recognize how different relevance aspects are correlated. A typical example of this situation is recommending other attendees to meet at a research conference. Here a range of similarity functions (social, past publications, current interest, location) could indicate a person worth to meet. To help conference attendees in their conference networking, we propose a dual social recommender interface, **Scatter Viz**, which includes a ranked list and visual scatter plot components. The ranked list was selected as a traditional way of presenting recommended results in a single dimension, listed from high to low relevance. The scatter plot was chosen as an intuitive way to present multidimensional data [17]. We hoped that the ability to view recommended items in two dimensions could reveal the overall diversity of results and help to correlate multiple types of relevance among the social recommendations.

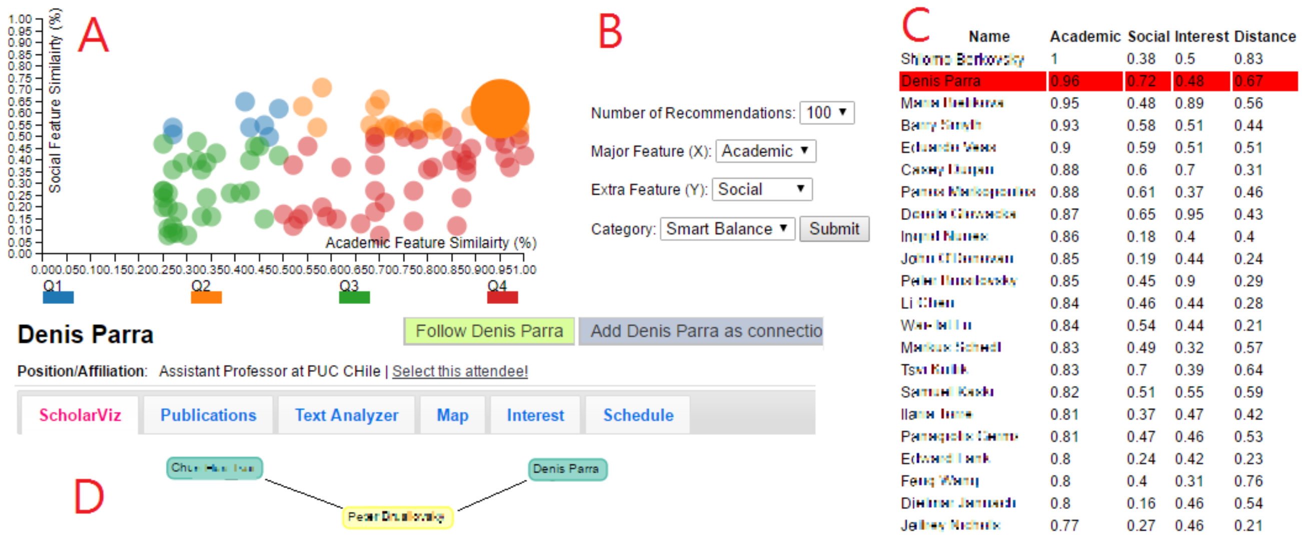

Figure 1. The design of Scatter Viz: (A) Scatter Plot; (B) Control Panel; (C) Ranked List; (D) User Profile Page. The interface supports exploration of recommended items in Section A or C and detail inspection in section D. (The scholar names have been pixelated for privacy protection.)

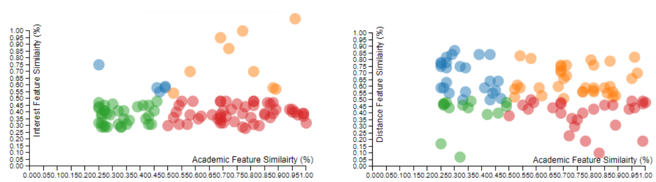

Figure 2. Scatter plot layouts: the layout would adjust, based on the selected *Major* and *Extra* features (Section B in Figure 1). Here is an example that presents the sane items presented in Academic/Interest and Academic/Distance feature coordinates. The nodes are colored using four quadrants if the Category "Smart Balance."

Visual Design

Figure 1 illustrates the design of the dual interface. **Section A** is the scatter plot. The interface presents each item (a conference attendee) as a circle on the canvas in two selected dimensions. The user can move the mouse over the circle to highlight the selection. **Section B** shows the control panel, with which the user can interact. The user can select the number of recommendations to display and choose the *major feature* and the *extra feature* to visualize the recommendations on the scatter plot. The major feature is used to rank the results along the X-axis and in the ranked list (section C), while the extra feature shows the diversity of results in the selected aspect along the Y-axis. To further investigate the diversity of the displayed recommendations, the user can also use another data aspect as a *category* to color-code the results. The default category was *Smart Balance*, which highlights four quadrants of the displayed data with a 0.5 ratio. Figure 2 presents a sample scatter plot layout combining "Academic" and "Interest/Distance" relevance features and color-coded using the Smart Balance Category.

Section C is the standard ranked list. More precisely, it is a combination of four ranked lists produced by four recommender engines, as explained below. To make the four dimensions more comprehensive, the model normalized relevance scores from 0-1 of each user to the target user, generated by each recommender engine. All the relevance scores are shown on the right side of the ranked list. The user can hover over each row to highlight the location in the scatter plot or click for a more detailed user profile. **Section D** presents more detailed information about the person who has been selected in either the visualization or the ranked list. Among other aspects, four of the six tabs visually explain how each recommender engine calculates the relevance of the selected user to the target user. Due to the page limitation, the details of each explanation tabs are omitted. The design detail of the explanation functions can be found in the work of [42].

The visual encoding affects the way users process the information. Pre-attentive processing let users absorb and precept the enormous amount of information in a short period [9]. The proposed interface helps to present the recommendation results in two kinds of visual encoding. First, the interface displays the recommendation relevance in two dimensions. The visual encoding helps the user to *spot* the item in different dimensions. It helps the user to make a decision beyond single or combined relevance, which is more realistic in many real-world scenarios. For example, a user may be interested in a scholar who whose research area is highly relevant to their research and who is also affiliated with nearby cities. The scatter interface helps to filter a group of recommended items with the two desired relevance features. Second, the node is color-coded in different categorical features; for example, in Smart Balance mode, the node is color-coded by the four quadrants between two dimensions of features. The user can perceive the tendencies of the recommendation item, based on their coloring, and the user can also update the layout with different Category features, including the meta-data of the recommended scholar's title, position, and home country. In addition, both the node and table row are highlighted synchronously while the user moves over the recommended items (see example in Figure 1). As a result, the scatter plot interface can be used for recommended item selection or just as a diversity-oriented recommendation explanation.

Personalized Relevance Model

To rank the recommended attendees by their relevance to the target user, the system uses four separate recommender engines that rank other attendees along four dimensions that we call *features*: text similarity of their *academic* publications, *social* similarity through the co-authorship network, similarity of current *interests* measured as similarity of their bookmarked talks, and the *distance* of their place of affiliation to the target user. Each of these features is defined below.

(1) The Academic feature is determined by the degree of publication similarity between two attendees using cosine similarity [43]. The function is defined as:

$$Sim_{Academic}(x,y) = (t_x \cdot t_y)/\|t_x\|\|t_y\| \tag{1}$$

where t is word vectors for user x and y. We used tf-idf to create the vector with a word frequency upper bound of 0.5 and lower bound of 0.01 to eliminate both common and rarely used words.

(2) The Social feature approximates the social similarity between the target and recommended users by combining co-authorship network distance and common neighbor similarity from publication data. We adopted the depth-first search (DFS) method to calculate the shortest path p [34] and common neighborhood (CN) [24] for the number n of coauthors overlapping in two degrees for user x and y.

$$Sim_{Social}(x,y) = p + n \tag{2}$$

(3) The Interest feature is determined by the the number of co-bookmarked papers and co-connected authors within the experimental social system [5]. The function is defined as

$$Sim_{Interest}(x,y) = (b_x) \cap (b_y) + (c_x) \cap (c_y) \tag{3}$$

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.

(4) The Distance feature is a measure of geographic distance between attendees. We retrieve longitude and latitude data based on attendees' affiliation information. We used the Haversine formula to compute the geographic distance between any pair of attendees [43].

$$Sim_{Distance}(x,y) = Haversine(Geo_x, Geo_y) \tag{4}$$

where *Geo* are pairs of latitude and longitude coordinates for user x and y.

Diversity Navigation Model

The Personalized Relevance Model determines the combined relevance score for each conference attendee, i.e., instead of ranking the recommended people using a static ensemble fusion of relevance aspects, the system allows the user to rank and visualize items using on different aspects of relevance through our proposed interface. We can measure the user's selection diversity with two diversification models.

(1) Feature Diversification: the user can select any two pairs of proposed features and spot the recommended items from the intersection of their relevance. All of the proposed features were calculated on a different scale. For example, the distance feature is the physical distance in miles, while the academic feature is calculated as a percentage. To enable the comparison of diverse features, we adopted a standard Z-score to normalize all the features to the same scale, from 0 to 1. The function was defined as:

$$ZScore = \frac{x_i - u_j}{\sigma_j} \tag{5}$$

where x_i is the ith recommended item and j represents the corresponding feature with its average u and variance σ. Then, we use a standard Z-table to convert the *ZScore* to the corresponding percentile p_{ij}. Hence, we can list all the features on the same scale for presentation in a ranked list and scatter plot diagram.

(2) Category Diversification: it is a model of diversifying the different categories [16]. In the scatter plot, we color-code the items from different categories, such as title, position, and country. In the ranked list, we listed the category as one column for a user to access.

(3) Selection Diversity: we can then measure the user selection/exploration diversity, based on the two diversification models. We observe the user's interaction with items from different "quadrants" (feature intersections) [38], such as high academic and high social features, or high academic and low social features. The extent of diversity is measured by *Shannon Entropy*:

$$Entropy : d_u = -\sum_{i=1}^{4} p_i log_4 p_i \tag{6}$$

where p_i is the probability for a particular quadrant (feature or category) and the proportion of all of the user's selections [22]. Based on the definition, we can measure the diversity in the different aspects of the relevance dimension. We can compare the combinations of all the proposed features. For example, in a recommendation system fused with $four$ features. We can measure the entropy difference among the $4*(4-1) = 12$ pair of dimensions.

STUDY 1: RANKED LIST V.S. SCATTER VIZ

Data and Participants

The recommendations produced by all four engines are mostly based on data collected by the Conference Navigator 3 (CN3) system [5]. The system has been used to support 38 conferences at the time of writing this paper and has data on approximately 6,398 articles presented at these conferences; 11,939 authors; 6,500 users (attendees of these conferences); 28,590 bookmarks; and 1,336 social connections. We used the Aminer dataset [33] to mediate the cold-start issue for academic and social engines that occurs when users have no publications or co-authorship information [39].

A total of 25 participants (13 female) were recruited for the user study. All of the participants were attendees at the 2017 Intelligent User Interfaces Conference (IUI 2017). Since the primary goal of our system was to help junior scholars connect with other people in the field, we specifically selected junior scholars, such as graduate students or research assistants. The participants came from 15 different countries; their ages ranged from 20 to 50 (M=37, SE=7.07). All of them could be considered as knowledgeable in the area of the intelligent interface for at least one academic publication from IUI 2017. To control for any prior experience with the recommender system, we included a question about in the background questionnaire. The average answer score was (M=3.28, SE=1.13) on a five-point scale.

Experiment Design and Procedure

To assess the value of the proposed interface, we compared the dual interface with the scatter plot and the ranked list (SCATTER) with a baseline interface using only a ranked list (RANK) with Section A (in Figure 1) removed. The study used a within-subjects design. All participants were asked to use each interface for three following tasks and to fill out a post-stage questionnaire at the end of their work with each interface. At the end of the study, participants were asked to compare interfaces regarding their explicit preference. The order of using interfaces was randomized to control for the effect of ordering. In other words, half of the participants started the study with the SCATTER interface. To minimize the learning effect (becoming familiar with data), we used data from two years of the same conference: the SCATTER interface used papers and attendees from IUI 2017, while the RANK interface used the corresponding data from IUI 2016.

Participants were given the same three tasks for each interface. The tasks were explicitly designed as diverse but realistic tasks that could be naturally pursued by attendees at research conferences. **Task 1:** Your Ph.D. adviser has asked you to find four Committee Member candidates for your dissertation defense. You need to find candidates with expertise close to your research field while trying to lower their travel cost to your defense. **Task 2:** Your adviser has asked you to meet four attending scholars, preferably from different regions across the world, who have a close connection to your research group. **Task 3:** You want to find four junior scholars (not yet faculty members) with reasonably similar interests among the conference attendees to establish networking.

The participants were asked to pick suitable candidates among conference attendees, based on their best judgment in each task. When designing the tasks, we attempted to make them realistic, yet focused on multiple aspects of relevance, as many real tasks are. We consider that task 1 is relevance-oriented and that tasks 2& 3 are diversity-oriented. For a relevance-oriented task, we expect to see if the proposed interface helps the user to coordinate different relevance aspects of the desired target efficiently. In contrast, for the diversity-oriented task, we expect the system to help to recognize the diversity of recommended items, as compared to the baseline interface.

Action Analysis

Table 1 shows the system usage for two interfaces. The data indicate that participants extensively used both the control panel and explanation tabs to complete the tasks. The participants usually required more actions on the first task to familiarize themselves with the system. There is no significant difference on the action of change control panel and the click on the explanation tab between the interfaces in three tasks; although, in the SCATTER interface, the users tended to click the explanation functions less. The click frequency presented a significant difference between the two interfaces. This finding is not surprising because the RANK interface lacks the visualization information that pushes the participant to click more on the user profile page to inspect the necessary information. It is interesting to see that not every user clicked on the scatter plot. This data hints that some participants treated the scatter plot visualization as an explanation function rather than an interactive exploration interface. At the same time, we found no significant difference in the time spent on the tasks. The data hints that each action taken in the SCATTER interface delivered more interesting information with which to engage.

		RANK		SCATTER		P
Task	Action	M (SE)	User Count	M (SE)	User Count	
T1	Control Panel	3.88 (2.40)	24	4.12 (2.02)	25	
	Explanation Tab	34.28 (29.50)	25	7.96 (7.48)	19	
	Click - Rank	26.28 (29.50)	25	4.92 (6.75)	15	*
	Click - Scatter	-	-	3.04 (5.45)	13	*
	Time Spent	345.44 (209.86)	25	389.12 (235.29)	25	
T2	Control Panel	2.88 (1.64)	24	2.88 (1.12)	25	
	Explanation Tab	19.96 (17.47)	25	9.16 (6.28)	25	
	Click - Rank	16.96 (17.47)	25	2.68 (4.69)	15	*
	Click - Scatter	-	-	3.48 (5.41)	13	*
	Time Spent	216.6 (144.95)	25	190.84 (115.33)	25	
T3	Control Panel	2.56 (1.04)	24	2.84 (1.10)	25	
	Explanation Tab	20.08 (20.29)	25	6.4 (7.22)	19	
	Click - Rank	19.08 (20.29)	25	3.48 (7.80)	9	*
	Click - Scatter	-	-	2.92 (2.95)	15	*
	Time Spent	345.95 (156.39)	25	369.2 (169.77)	25	

Table 1. User action summary of study 1: the table shows the user interaction statistics while performing each of the three tasks using two interfaces. (Statistical significance level: (*) $p < 0.05$.)

User Feedback Analysis

To compare subjective feedback, we analyzed the responses of the post-stage questions using paired sample t-tests. Figure 3 shows the result of this analysis. We compared the twelve aspects of subjective feedback from the participants; among them, the SCATTER interface received a significantly higher rating for six aspects: Trust (Q4), Supportiveness (Q5), Interest (Q6), Satisfaction (Q8), Intention to Reuse (Q9), and Enjoyable (Q11). In two questions, facilitation (Q7) and the Reversed Benefit Question (Q12), the SCATTER interface scored higher, but not significantly so. It is interesting to see

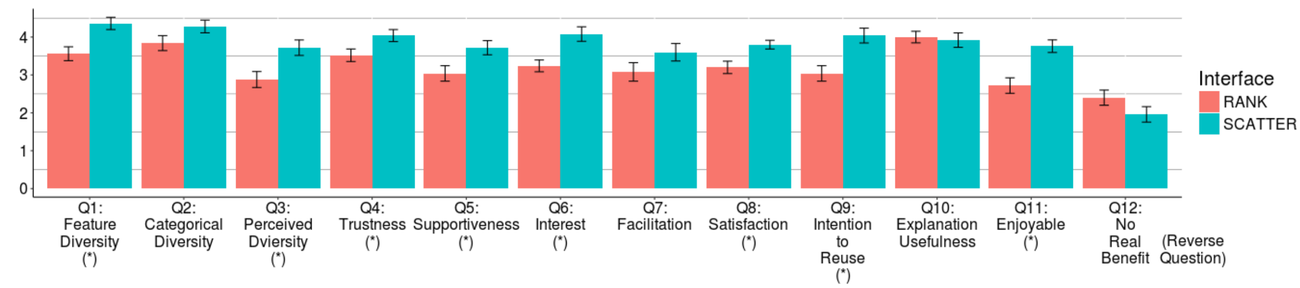

Figure 3. User feedback analysis results for Study 1 shows that the SCATTER interface received a significantly higher rating for six aspects. (A cut-off value was set at 3.5 on the 5 point scale. Statistical significance level: (*) $p < 0.05$.)

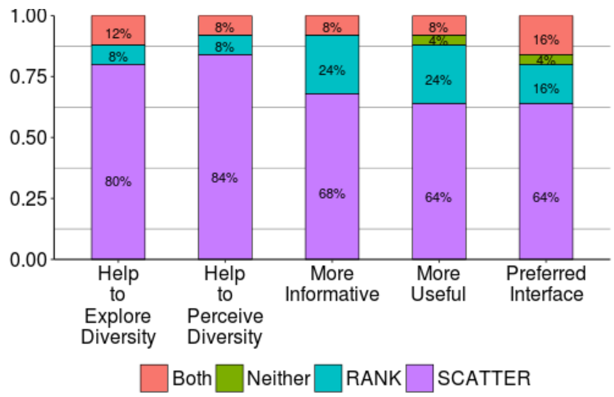

Figure 4. User preference analysis in study 1 (the preferences were collected after the users experienced both interfaces). The result shows the SCATTER interface was preferred by the users in all aspects.

that the RANK interface scored a bit higher (though not significantly so) on explanation usefulness, which hints that the lack of visualization made explanations more important in the RANK interface. In the final preference test, the SCATTER interface received much stronger support than the RANK interface in the user preference feedback (Figure 4). Most importantly, a majority of users (84%) considered the SCATTER interface to be a better system for recommending attendees and a better help in diversity-oriented tasks, as well as a better system for recommending.

Diversity Analysis

Table 2 shows the diversity analysis for each task and interface. The result shows the users' responses to the tasks with a different pattern of exploration, which caused a variance of diversity and coverage measurements. All three tasks are shown a least one significance between two interfaces but in the different aspects of features. For the SCATTER interface, task 1 (relevance-oriented) shows significance statically on less difference between academic/social & social/interest features, but more coverage on the title category. Tasks 2 & 3 (diversity-oriented) show higher selection diversity in the interest/distance and social/distance features, respectively, as well as higher selection coverage in the title & country category features. The data supports the finding that the SCATTER interface helped the participants to accurately filter the attendees in the relevance-oriented task, as well as extend the selection diversity in the diversity-oriented tasks.

		RANK	SCATTER	P
Task	Dimensions	M (SE)	M (SE)	
T 1	Academic + Social	0.14 (0.06)	0.11 (0.03)	*
	Academic + Interest	0.16 (0.09)	0.13 (0.08)	
	Academic + Distance	0.13 (0.07)	0.12 (0.04)	
	Social + Interest	0.27 (0.13)	0.21 (0.09)	*
	Social + Distance	0.27 (0.12)	0.24 (0.08)	
	Interest + Distance	0.26 (0.13)	0.25 (0.13)	
	Title	0.17 (0.12)	0.22 (0.07)	*
	Position	0.26 (0.12)	0.23 (0.10)	
	Country	0.49 (0.25)	0.46 (0.15)	
T 2	Academic + Social	0.14 (0.06)	0.13 (0.04)	
	Academic + Interest	0.14 (0.09)	0.14 (0.08)	
	Academic + Distance	0.12 (0.07)	0.14 (0.04)	
	Social + Interest	0.25 (0.13)	0.24 (0.11)	
	Social + Distance	0.26 (0.13)	0.28 (0.10)	
	Interest + Distance	0.23 (0.14)	0.27 (0.13)	-
	Title	0.17 (0.14)	0.31 (0.12)	*
	Position	0.29 (0.17)	0.25 (0.14)	
	Country	0.46 (0.31)	0.68 (0.26)	*
T 3	Academic + Social	0.14 (0.06)	0.13 (0.03)	
	Academic + Interest	0.12 (0.07)	0.12 (0.08)	
	Academic + Distance	0.14 (0.07)	0.16 (0.04)	
	Social + Interest	0.22 (0.12)	0.22 (0.14)	
	Social + Distance	0.24 (0.11)	0.31 (0.10)	*
	Interest + Distance	0.21 (0.14)	0.23 (0.11)	
	Title	0.17 (0.16)	0.32 (0.10)	*
	Position	0.19 (0.14)	0.15 (0.08)	
	Country	0.44 (0.26)	0.66 (0.26)	*

Table 2. Diversity analysis for study 1: the table shows selection diversity for three tasks in the feature and category dimensions. The result shows that the SCATTER interface can help users to explore a more diverse set of recommendation in diversity-oriented tasks (T2 & T3). (Statistical significance level: (*) $p < 0.05$; (-) $p < 0.1$.)

Discussion

In study 1, we evaluated a dual visual interface for recommending attendees at a research conference. A research conference context introduces several dimensions of attendee relevance, such as social, academic, interest, and distance similarities. Due to these factors, a traditional ensemble ranked list makes it difficult to express the diversity of recommended items (attendees). By spreading rankings over two dimensions, the suggested interface helps users to explore recommendations and recognize their diversity in several aspects. To assess the visual approach, we conducted a user study in a real conference environment to compare our interface (SCATTER) with a traditional ranked list (RANK) in three practical tasks. Our experimental result shows a tangible incremental impact on the metrics of system usage, efficiency, and diversity.

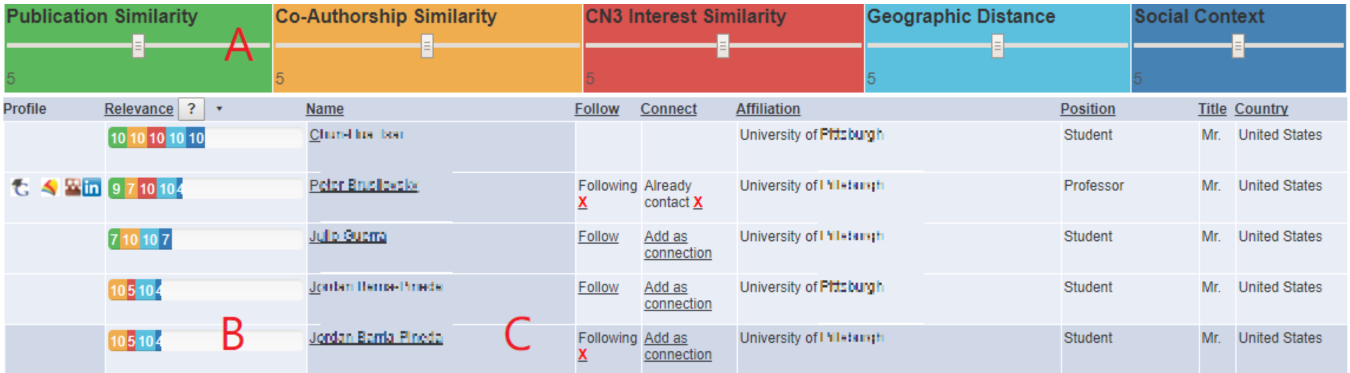

Figure 5. The design of the Relevance Tuner: (a) Relevance Sliders; (B) Stackable Score Bar; (C) User Profiles. The interface enables the user to adjust the feature weighting on-the-fly for retrieving a customized recommendation list. The user can examine the relevant aspects of the recommended item through the multicolored score bar. (The Name and Affiliation entities have been pixelated for privacy protection.)

We found that the **Scatter Viz** interface can improve user inspection on recommendations with multi-relevance, which leads to a higher selection diversity in the given tasks. However, we also noticed that some of the experiment participants still stick to the familiar ranked list, even when an enhanced visualization was provided. This finding helps us to realize a user preference on adopting the interface with lower learning efforts. Besides, the scatter visualization requires additional space to present, which may not be feasible in many real-world rank-based recommender systems. These findings lead to our second attempt at extending the ranked list with multi-aspect awareness, controllability, and diversity-aware designs.

SECOND ATTEMPT: RELEVANCE TUNER

The ranked list is widely applied to recommender systems for presenting recommendations to users. Even in visual recommender systems, a basic ranked list is still essential for user interactions [44, 29, 40]. A qualified recommender system usually ranks the recommended items from high to low relevance, which may reduce the users' cognitive decision loading [2]. However, in a context with multiple relevance aspects, a statically ensembled ranked list makes it more difficult for the user to recognize the impact of different aspects and to adjust the recommendations to different needs. In addition, the persuasive design of the ranked list causes users to pay more attention to items on top of the list [8], which may cause the low selection diversity and the fitter bubble effect [25].

According to the first study, we knew the users were able to correlate the multiple relevances in a two-dimensional scatter plot visualization, but that this ability came with a high learning curve. In our second attempt, we would like to see if the user can still correlate the multi-relevance within a ranked list extended with a controllable tuner and stackable color bars. We proposed the **Relevance Tuner** - a visual interface with user-driven control function and meaningful visual encoding. This design aims to help the users to inspect and filter recommendations of multi-relevance from a ranked list, which may facilitate the exploration of diversity-oriented tasks. The interface is extensible to various types of recommender systems without the limitation of dual interfaces. This design reduces the user's difficulties in getting familiar with the interface.

Visual Design

The design of the Relevance Tuner is shown in Figure 5. **Section A** contains five controllable sliders with the different colors representing the features of the Personalized Relevance Model. The scale of the slider ranges from 0 to 10. The user can change the weighting on the fly to re-rank the ranked list below. It provides controllability for the user to adjust the ranking to different recommendation needs and preferences. The interface also adds one new feature: Social Context. This feature computes the Google search result based on the scholar's name and affiliation information; that is, the text similarity of the homepage and other related search results. **Section B** shows the stackable relevance score bar of each recommended item in the ranked list. The color corresponds to the features in section A. It would adaptively adjust the bar score (length) from 0 to 20, based on the weighting percentage of the sliders. A stackable color bar interface is known for its ability to enhance controllability and transparency in a multi-aspect ranking [9]. In our system, the stackable color bars help the user to see how different relevant aspects of a recommended item are coordinated while adding transparency to the multi-aspect recommendation process. **Section C** shows the recommended scholar's meta-data, including name, social connection, affiliation, position, title, and country. The user can sort the ranked list by clicking the head of each column, or can inspect the explanation tabs (same as Section C in Figure 1) by clicking the name entities.

STUDY 2: SCATTER VIZ V.S. RELEVANCE TUNER

Data and Participants

Study 2 was conducted through the Conference Navigator 3 (CN3) system. The data was extended from Study 1 to a new conference: the 25th Conference on User Modeling, Adaptation, and Personalization (UMAP 2017). A total of 20 participants (7 female) were recruited for the user study. All of the participants were attendees at the UMAP 2017 conference. They were from 15 different countries; their ages ranged from 20 to 40 (M=31.19, SE=4.97). All of them had at least one publication from UMAP 2017. The background knowledge of recommender systems score was (M=3.85, SE=0.79) on a five-point scale.

		TUNER		SCATTER		P
Task	Action	M (SE)	User Count	M (SE)	User Count	
T1	Control Panel	38.4 (37.71)	20	2.85 (2.23)	18	
	Explanation Tab	9.35 (8.28)	20	22.95 (23.71)	20	*
	Click - Rank	5.05 (2.45)	20	9.8 (8.43)	17	*
	Click - Scatter	-	-	4.1 (6.03)	11	
	Time Spent	357 (289.04)	20	537 (596.98)	20	
T2	Control Panel	15.2 (13.63)	19	2 (1.71)	17	*
	Explanation Tab	6.5 (8.74)	20	8.45 (6.79)	20	
	Click - Rank	4.3 (1.21)	20	5.2 (3.76)	18	
	Click - Scatter	-	-	1.8 (2.94)	8	*
	Time Spent	201 (235.43)	20	294 (470.78)	20	
T3	Control Panel	12.2 (11.67)	17	2.25 (1.80)	19	*
	Explanation Tab	9.25 (8.75)	20	12.9 (14.38)	20	
	Click - Rank	5.15 (2.51)	20	10.2 (16.93)	17	
	Click - Scatter	-	-	2.05 (3.13)	10	*
	Time Spent	153 (92.28)	20	285 (470.78)	20	

Table 3. User action summary of study 2: the table shows the statistics of user interaction while solving each of the three tasks using each interface. (Statistical significance level: (*) $p < 0.05$.)

Experiment Design and Procedure

In Study 2, we compared the interface of the ranked list plus the relevance tuner (TUNER) with a baseline of the scatter plot plus ranked list (SCATTER). The experiment design and procedure repeat the setting of Study 1. We manipulated the new proposed interface and adapted data from different conferences: the SCATTER interface used papers and attendees from UMAP 2017, while the TUNER interface used the same data from UMAP 2016, to minimize the learning effect between the two manipulations.

Action Analysis

Table 3 shows the system usage for the two interfaces of Study 2. In TUNER interface, the control panel usage is defined as each time the user moves the sliders. The data supports the users interacting more frequently (it shows significance in all three tasks) with the control panel in TUNER than in SCATTER. Conversely, the users clicked more on explanation tabs in SCATTER than in TUNER. The data implies that the information listed on the table was sufficient for the users to inspect and make decisions in three proposed tasks. In task 1, the SCATTER has a significantly higher clicking frequency and longer time spent (not significant) than the TUNER interface. The same pattern repeats in task 2 & 3, which shows that the users took more time to get familiar with the SCATTER interface. The users were gaining familiar with the TUNER interface more rapidly than with the SCATTER interface.

User Feedback Analysis

Figure 6 shows the analysis of the post-stage survey. The high rating in both interfaces shows the positive user acceptance in Study 2 (no significance on all the factors). However, the user tends to favor the TUNER interface when considering the factors of Supportiveness (Q5), Interest (Q6), Facilitation (Q7), Satisfaction (Q8), Intend to Reuse (Q9), and Usefulness (Q10). The SCATTER interface performs better on the measures of Trustiness (Q4) and Enjoyable (Q11). This result supports the users in favor of rank-based list more than the visual-based interface, but the visualization shows an increased level of usability on gaining trust and enjoyment in using the interface. Surprisingly, the feedback also indicates that the TUNER interface would be better for the user to fulfill the task on the feature diversity (Q1) and category diversity (Q2), but that the SCATTER interface is outperformed on the ability to Perceive Diversity (Q3).

This result shows that a user tends to use the ranked list with better controllability and transparency to conduct diversity-oriented tasks. The scatter visualization would play the role of helping the users to perceive diversity in multiple areas of relevance. The final preference result in Figure 7 also confirms this conclusion. About half of the users select the TUNER interface as the one with an advantage at helping to explore diversity, providing more informative information, being more useful, and fitting their preference - but the users also agree that the SCATTER interface could better help to perceive diversity after they finished the three tasks on two interfaces.

Diversity Analysis

Table 4 shows the diversity analysis of Study 2. In task 1 (relevance-oriented), there is no significant difference between the two interfaces on the diversity measurement, showing both of the interfaces can support the user to fulfill a relevance-oriented task. However, in the diversity-oriented tasks 2&3, we found the TUNER group could achieve higher entropy than the SCATTER group. This finding hints that even with a ranked-list interface the user can achieve a good level of selection diversity if controllability and transparency for each considered dimension of relevance is available. At the same time, the SCATTER interface performed slightly (but not significantly) better than the TUNER interface in the *category* diversity metrics. This finding helps to highlight the value of color-coding data in the SCATTER interface, a feature not supported by TUNER. The diversity analysis of Social Context is omitted due to the page limitation.

Discussion

In Study 2, we presented a new rank-based interface for recommender attendees at a research conference. A total of five dimensions of relevance were proposed from the Personalized Relevance Model. We conducted a user study in a real conference environment to compare the two interfaces of an enhanced ranked list (TUNER) and the visualization interface (SCATTER). Our experimental results suggested the different suitable scenarios for the two interfaces. We found that, even in diversity tasks with multi-relevance settings, the users were still able to fulfill the diversity task with a rank-based interface, but it required the support of interface controllability

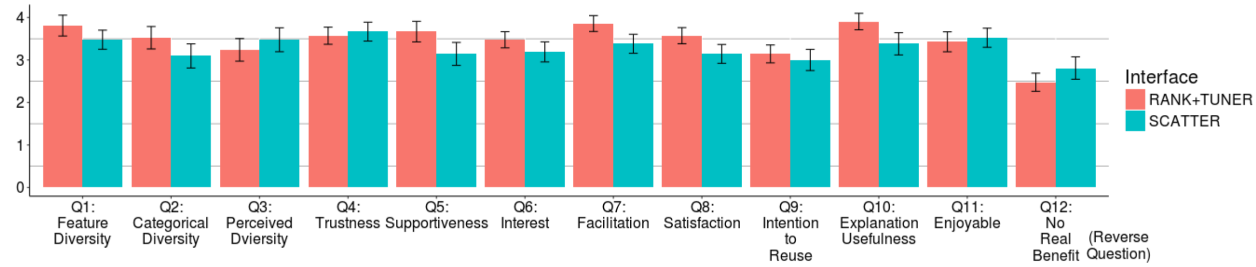

Figure 6. User feedback analysis results for study 2: we did not find significant difference in all aspects, which indicates that the usability of two interfaces was comparable. (A cut-off value was set at 3.5 on the 5 point scale. Statistical significance level: (*) $p < 0.05$.)

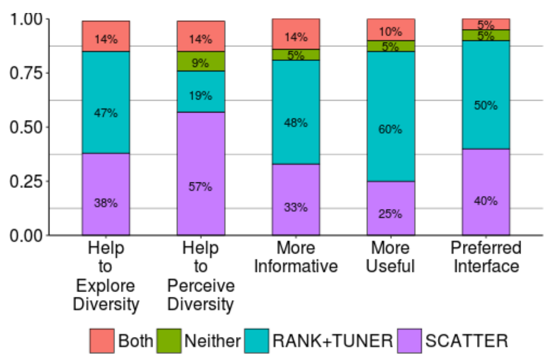

Figure 7. User preferences for study 2 collected after the users experienced both interfaces. The result shows that the TUNER interface was preferred by users in all aspects except perceived diversity.

Task	Dimensions	TUNER M (SE)	SCATTER M (SE)	P
T 1	Academic + Social	0.12 (0.05)	0.12 (0.05)	
	Academic + Interest	0.17 (0.10)	0.16 (0.07)	
	Academic + Distance	0.13 (0.06)	0.12 (0.04)	
	Social + Interest	0.23 (0.11)	0.23 (0.10)	
	Social + Distance	0.23 (0.10)	0.22 (0.09)	
	Interest + Distance	0.24 (0.16)	0.29 (0.12)	
	Title	0.12 (0.04)	0.16 (0.13)	
	Position	0.29 (0.15)	0.27 (0.12)	
	Country	0.35 (0.20)	0.57 (0.29)	*
T 2	Academic + Social	0.15 (0.02)	0.12 (0.04)	*
	Academic + Interest	0.17 (0.07)	0.13 (0.07)	*
	Academic + Distance	0.16 (0.06)	0.12 (0.05)	*
	Social + Interest	0.25 (0.07)	0.19 (0.08)	*
	Social + Distance	0.28 (0.07)	0.22 (0.10)	*
	Interest + Distance	0.31 (0.15)	0.24 (0.13)	*
	Title	0.15 (0.02)	0.17 (0.12)	
	Position	0.28 (0.12)	0.32 (0.15)	
	Country	0.41 (0.24)	0.58 (0.22)	*
T 3	Academic + Social	0.16 (0.03)	0.13 (0.05)	
	Academic + Interest	0.14 (0.07)	0.14 (0.06)	
	Academic + Distance	0.17 (0.05)	0.14 (0.05)	*
	Social + Interest	0.32 (0.09)	0.26 (0.16)	
	Social + Distance	0.32 (0.08)	0.26 (0.09)	*
	Interest + Distance	0.28 (0.12)	0.26 (0.13)	
	Title	0.15 (0.02)	0.14 (0.05)	
	Position	0.16 (0.07)	0.22 (0.16)	
	Country	0.48 (0.19)	0.54 (0.32)	

Table 4. Diversity analysis for study 2: the table shows selection diversity for three tasks for each feature combination and category dimensions. The result indicates that the TUNER interface enabled users to explore a more diverse set of recommendation in diversity-oriented tasks T2 & T3. (Statistical significance level: (*) $p < 0.05$.)

and transparency through visual encoding. Besides, while we found that the user would better perceive the diversity in the SCATTER interface, the user would prefer to adopt the TUNER interface to fulfill the diversity tasks.

Furthermore, when the user was interacting with the TUNER interface, the user spends more time on inspecting the information on each row, instead of checking the explanation functions. This result shows that the higher level of diversity exploration was not triggered by the diversity-enhanced visualization or explanation (fewer clicks on the explanation tabs), but was instead contributed by the intention of the user reaction to the simulated diversity-oriented tasks. In the SCATTER interface, the user relies more on the explanation function and multi-relevance visualization to explore diversity-oriented tasks. Although the result showed lower entropy measurement when the user adapted to the SCATTER interface, the SCATTER interface can better help the user to perceive the diversity among multiple-relevance dimensions, based on the user feedback analysis through post-study questionnaires.

USER-CENTRIC EVALUATION

To better understand the mediation effects across the two interfaces, we conducted a structural equation model (SEM) analysis [20] to inspect the effects of the proposed interfaces on the user experiences. We used the logged data and questionnaire feedback from the two studies. There are two conditions and five summarized factors in the model (as shown in Figure 8). In objective system aspects (OSA), there are two manipulations based on the proposed interfaces. In subjective system aspects (SSA) and user experience (EXP), we proposed four factors based on the classification by [30] and our post-experiment questions. In interaction (INT), we listed the entropy of the participant's selection diversity (average of Task 2&3 that shown in Table 2 and 4). The model fit the statistics of $\chi^2(96) = 234.68$, $p < 0.01$, $RMSEA = 0.18$, $90\%CI : [0.152, 0.211]$, $CFI = 0.941$, $TLI = 0.922$.

The model shows that the two manipulations have different positive effects on the system. The TUNER condition has a positive effect on the Perceived Usefulness and Perceived Diversity factors. It supports providing controllability and transparency on the ranked list, which helps the users to perceive the variety of recommendation results and increase the overall usefulness rating. The SCATTER condition only affects the Trust factor, which supports the argument for providing multi-relevance visualization to help the user to gain trust in the interface, but not mediates to the ratings of satisfaction or

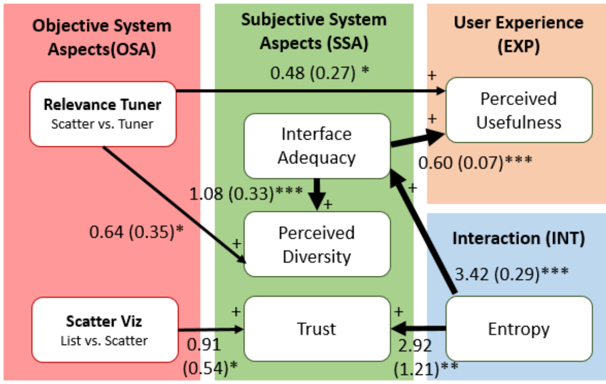

Figure 8. The structural equation model of the experiments. The number (thickness) on the arrows represents the β coefficients and standard error of the effect. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

usefulness. We also found the effect between Interface Adequacy and Perceived Diversity, which supports a better design interface that helps the user to perceive the recommendation diversity. Furthermore, a positive impact on Interface Adequacy to Perceived Usefulness implied that the usefulness is contributed by the rating of the interface design. The Entropy value mediates the additional effect of the two interfaces. We found that the Entropy factor has a positive impact on Interface Adequacy and Trust; that is, the users who can achieve higher selection diversity through the given interfaces gain more trust and interface adequacy ratings in the system. This finding explains why the rank-based interface, with a lower learning curve, is preferred by the participants on the proposed tasks. In other words, it also supports the multi-relevance visualization that is appreciated by the users only when they are capable of taking full advantage of the interface.

CONCLUSION

In this paper, we presented experiments on three different interfaces: RANK, TUNER, and SCATTER. We first showed that providing a scatter plot (SCATTER) can help the user to better fulfill the diversity-related tasks, as compared to a simple ranked list (RANK). However, despite the benefits of the new two-dimensional presentation, the users still extensively use the ranked list component of the interface. Based on the results of the first study, we attempted to integrate the ability to coordinate multiple aspects of relevance within the ranked list rather than offering it in a separate component as in SCATTER. To compensate for the biasing nature of the ranked list, we also provided controllable fusion of relevance aspects. The resulting TUNER interface offered both controllability and visual encoding of multiple relevance aspects. We showed that the users could adopt a rank-based list to fulfill diversity-oriented tasks with higher selection diversity. The usability analysis reveals that both SCATTER and TUNER were ranked by the conference users with high subjective ratings. However, TUNER requires less learning effort. We also discussed the mediation effects of the proposed interfaces on the user experience. The analysis helps to describe the benefits of the two proposed interfaces in social recommender systems.

We aim to understand how the user adopts the recommender interfaces to diversity-oriented tasks. The given tasks indicated a concrete goal of exploring the conference attendees with multiple types of relevance. We found that both of the proposed interfaces were capable of helping the user to fulfill the assigned tasks. The experimental result supported the finding that the participants were able to correlate multiple aspects of relevance using two dimensions of visualization in SCATTER and the controllable ranked list with multi-aspect visualization in TUNER. We found that an extension of a traditional ranked list was better than a separate visualization component in the sense of on getting familiar with the new interface, which led to a higher rating on the user preference. In other words, a separate diversity-enhanced visualization can also achieve the goal, but it came at the cost of a steeper learning curve for the user. However, once the users were familiar with the interface, it brought an advantage of helping the users to perceive diversity and gain trust in the recommendations.

The findings shed light on designing diversity-enhanced and diversity-aware interfaces in a social recommender system. In a diversity-enhanced system, the target users would have a concrete goal of exploring the recommendation results. A similar scenario was reported in this paper; that is, the conference attendees explore the social recommendation within a particular conference. The user has specific goals on the exploration from multiple aspects of relevance. However, the users may not have specific goals when interacting with the recommender system. For example, when accessing news article on social media, the user may not have a strong desire to read the news article from a different perspective. In this case, we need a diversity-aware interface to help the user to perceive the prospect diversity, which may be a solution to help to solve the filter-bubble effect [36]. In this paper, we find that the effects of usability and perceived diversity are additive. The best user experience may happen when the users can take advantages from both of the proposed interfaces. It remains an open question as to how to design an interface to reflect the needs of users better.

LIMITATIONS AND FUTURE WORK

Our study has some limitations. First, the within-subject user study was conducted using consecutive years of the same conference series. Some well-known and senior domain experts may appear in the recommendation list for both conferences. This repetition may cause bias in our user studies. Second, the data sparsity and cold-start problem may hurt the recommendation performance; for example, the Interest feature is less useful for users who never bookmarked any talks with CN3. We tried our best to send out emails both before and during the conference to improve interest-based recommendations. Third, the scale of reported user studies is relatively small. It may decrease the statistical power of the findings. Fourth, the experiment was conducted at mid-size conferences, so we were not able to explore scaling issues which might occur at conferences with much larger number of attendees or in a different recommendation context with a large set of items to explore. We hope to address some of these limitations in our future work.

REFERENCES

1. Jisun An, Daniele Quercia, and Jon Crowcroft. 2013. Why individuals seek diverse opinions (or why they don't). In *Proceedings of the 5th Annual ACM Web Science Conference*. ACM, 15–18.
2. Dirk Bollen, Bart P Knijnenburg, Martijn C Willemsen, and Mark Graus. 2010. Understanding choice overload in recommender systems. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 63–70.
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. Engin Bozdag and Jeroen van den Hoven. 2015. Breaking the filter bubble: democracy and design. *Ethics and Information Technology* 17, 4 (2015), 249–265.
5. 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.
6. 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.
7. Henriette Cramer, Vanessa Evers, Satyan Ramlal, Maarten Van Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob Wielinga. 2008. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction* 18, 5 (2008), 455.
8. Paolo Cremonesi, Franca Garzotto, and Roberto Turrin. 2012. Investigating the persuasion potential of recommender systems from a quality perspective: An empirical study. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 2 (2012), 11.
9. Cecilia di Sciascio, Vedran Sabol, and Eduardo E Veas. 2016. Rank As You Go: User-Driven Exploration of Search Results. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, 118–129.
10. Michael D Ekstrand, Daniel Kluver, F Maxwell Harper, and Joseph A Konstan. 2015. Letting users choose recommender algorithms: An experimental study. In *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 11–18.
11. Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1175–1184.
12. Dorota Glowacka, Tuukka Ruotsalo, Ksenia Konuyshkova, Samuel Kaski, Giulio Jacucci, and others. 2013. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 117–128.
13. Brynjar Gretarsson, John O'Donovan, Svetlin Bostandjiev, Christopher Hall, and Tobias Höllerer. 2010. Smallworlds: visualizing social recommendations. In *Computer Graphics Forum*, Vol. 29. Wiley Online Library, 833–842.
14. Jonathan L Herlocker, Joseph A Konstan, and John Riedl. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM, 241–250.
15. Rong Hu and Pearl Pu. 2011. Helping Users Perceive Recommendation Diversity.. In *DiveRS@ RecSys*. 43–50.
16. Marius Kaminskas and Derek Bridge. 2016. Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 7, 1 (2016), 2.
17. Hannah Kim, Jaegul Choo, Haesun Park, and Alex Endert. 2016. Interaxis: Steering scatterplot axes via observation-level interaction. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 131–140.
18. Khalil Klouche, Tuukka Ruotsalo, Diogo Cabral, Salvatore Andolina, Andrea Bellucci, and Giulio Jacucci. 2015. Designing for exploratory search on touch devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 4189–4198.
19. Khalil Klouche, Tuukka Ruotsalo, Luana Micallef, Salvatore Andolina, and Giulio Jacucci. 2017. Visual Re-Ranking for Multi-Aspect Information Retrieval. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*. ACM, 57–66.
20. 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.
21. Q Vera Liao and Wai-Tat Fu. 2013. Beyond the filter bubble: interactive effects of perceived threat and topic involvement on selective exposure to information. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2359–2368.
22. Jennifer Moody and David H Glass. 2016. A Novel Classification Framework for Evaluating Individual and Aggregate Diversity in Top-N Recommendations. *ACM Transactions on Intelligent Systems and Technology (TIST)* 7, 3 (2016), 42.
23. Sean A Munson and Paul Resnick. 2010. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1457–1466.

24. Mark EJ Newman. 2001. Clustering and preferential attachment in growing networks. *Physical Review E* 64, 2 (2001), 025102.
25. Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. 2014. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*. ACM, 677–686.
26. John O'Donovan, Brynjar Gretarsson, Svetlin Bostandjiev, Tobias Hollerer, and Barry Smyth. 2009. A visual interface for social information filtering. In *Computational Science and Engineering, 2009. CSE'09. International Conference on*, Vol. 4. IEEE, 74–81.
27. 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.
28. Valeria Orso, Tuukka Ruotsalo, Jukka Leino, Luciano Gamberini, and Giulio Jacucci. 2017. Overlaying social information: The effects on usersâĂŹ search and information-selection behavior. *Information Processing & Management* 53, 6 (2017), 1269–1286.
29. Denis Parra and Peter Brusilovsky. 2015. User-controllable personalization: A case study with SetFusion. *International Journal of Human-Computer Studies* 78 (2015), 43–67.
30. Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*. ACM, 157–164.
31. Tuukka Ruotsalo, Jaakko Peltonen, Manuel Eugster, Dorota Głowacka, Ksenia Konyushkova, Kumaripaba Athukorala, Ilkka Kosunen, Aki Reijonen, Petri Myllymäki, Giulio Jacucci, and others. 2013. Directing exploratory search with interactive intent modeling. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*. ACM, 1759–1764.
32. J Ben Schafer, Joseph A Konstan, and John Riedl. 2002. Meta-recommendation systems: user-controlled integration of diverse recommendations. In *Proceedings of the eleventh international conference on Information and knowledge management*. ACM, 43–51.
33. 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.
34. Robert Tarjan. 1972. Depth-first search and linear graph algorithms. *SIAM journal on computing* 1, 2 (1972), 146–160.
35. Choon Hui Teo, Houssam Nassif, Daniel Hill, Sriram Srinivasan, Mitchell Goodman, Vijai Mohan, and SVN Vishwanathan. 2016. Adaptive, Personalized Diversity for Visual Discovery. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 35–38.
36. Nina Tintarev. 2017. Presenting Diversity Aware Recommendations: Making Challenging News Acceptable. (2017).
37. 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.
38. Chun-Hua Tsai. 2017. An Interactive and Interpretable Interface for Diversity in Recommender Systems. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces Companion (IUI '17 Companion)*. ACM, New York, NY, USA, 225–228.
39. Chun-Hua Tsai and Peter Brusilovsky. 2016. A personalized people recommender system using global search approach. *IConference 2016 Proceedings* (2016).
40. Chun-Hua Tsai and Peter Brusilovsky. 2017a. Enhancing Recommendation Diversity Through a Dual Recommendation Interface. In *Workshop on Interfaces and Human Decision Making for Recommender Systems*.
41. Chun-Hua Tsai and Peter Brusilovsky. 2017b. Leveraging Interfaces to Improve Recommendation Diversity. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 65–70.
42. Chun-Hua Tsai and Peter Brusilovsky. 2017c. Providing Control and Transparency in a Social Recommender System for Academic Conferences. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 313–317.
43. 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.
44. 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.
45. Weiquan Wang and Izak Benbasat. 2007. Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. *Journal of Management Information Systems* 23, 4 (2007), 217–246.
46. David Wong, Siamak Faridani, Ephrat Bitton, Björn Hartmann, and Ken Goldberg. 2011. The diversity donut: enabling participant control over the diversity of recommended responses. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1471–1476.