

Expanding Controllability of Hybrid Recommender Systems: From Positive to Negative Relevance

Behnam Rahdari, Chun-Hua Tsai, Peter Brusilovsky

School of Computing and Information, University of Pittsburgh

Pittsburgh, PA 15260, USA

{ber58,cht77,peterb}@pitt.edu

Abstract

A hybrid recommender system fuses multiple data sources, usually with static and nonadjustable weightings, to deliver recommendations. One limitation of this approach is the problem to match user preference in all situations. In this paper, we present two user-controllable hybrid recommender interfaces, which offer a set of sliders to dynamically tune the impact of different sources of relevance on the final ranking. Two user studies were performed to design and evaluate the proposed interfaces.

Introduction and Related Work

In a modern digital world, users leave more and more traces of their behavior such as browsing trails, bookmarks ratings, or created social links. It enables modern recommender systems to use multiple sources of information about user interests and preferences to deliver better recommendations. This is most frequently done using parallel hybrid recommendation approaches (Burke 2007), which fuse together item relevance generated by different sources of information. Typically, the fusion is done by assigning static weights to different sources, where the optimal weights are picked or learned from ground truth data (i.e., known ratings). The problem with this approach is that users might seek recommendations for different reasons and in different contexts. Depending on the case, individual sources in a hybrid recommender might become more or less valuable. As a result, while the “optimal” static fusion could provide the best ranking across the cases, it might be suboptimal within each specific case.

The problem of optimal source fusion has been originally explored in the domain of information retrieval where it was demonstrated that the user might be in a better position to decide which weight should be assigned to each relevance source in each case (Ahn et al. 2008). The idea of user-controlled personalization has been further explored within the recommender systems domain by (Schafer, Konstan, and Riedl 2002; O’Donovan et al. 2008). More recently, (Bostandjiev, O’Donovan, and Höllerer 2012) introduced sliders as an approach to engage the user into tuning various parameters of recommendation approach. Following that, the use of sliders as a way to support user-controlled fusion has

been explored in the domains recommender systems (Parra and Brusilovsky 2015) and information retrieval (di Sciascio, Sabol, and Veas 2016) bringing additional evidence in favor of the user-controlled hybrid recommendation.

An extensive exploration of sliders as a tool for a user-controlled hybrid recommendation was performed by (Tsai and Brusilovsky 2018) in a research conference context where it was applied to suggest the most relevant attendees to meet. By comparing slider-controlled fusion of relevant sources to other approaches such as user-controlled scatter plot visualization this study discovered an important weakness of this approach: it was hard to use when the recommendation context required *reversing* the weight of a specific recommendation source. For example, while a recommendation source based on co-authorship links ranks attendees by its social similarity with the target user, the recommendation case might require to find attendees who are interested in similar topics while being most likely unknown to the target user (i.e., having the weakest social similarity).

The study reported in this paper attempts to re-examine the value of user-controlled hybrid recommendation in a different context. It also attempts to resolve the reported problem of slider-based control by exploring a more complex slider-based interface *Paper Tuner+* with an option to *reverse* relevance for each source. We start by introducing the design of the *Paper Tuner* and reviewing a pre-study, which was performed to select the best option for the reversible relevance slider. Following that, we present the *Paper Tuner+* design with reversible sliders and report the results of the second user study that assessed the value of both version of the *Paper Tuner* against a non-controllable baseline.

The Controllable User Interface: Paper Tuner

The Paper Tuner is an interface for user controllable recommendation of research papers. The interface combines several features that have been found beneficial by the past work including slider control of source importance (Parra and Brusilovsky 2015; Tsai and Brusilovsky 2018) and stackable bars for visualizing combined relevance (di Sciascio, Sabol, and Veas 2016; Tsai and Brusilovsky 2018). The Paper Tuner consists of three main parts (Figure 1). **Section A** contains five sliders to control the importance of recommendation sources used to generate the ranked list of the

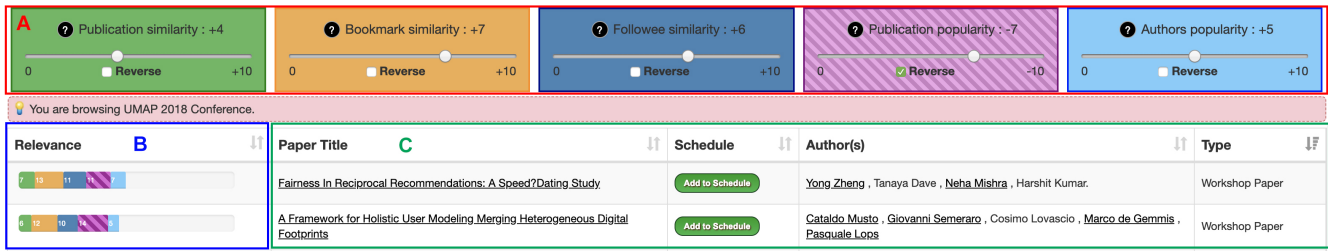


Figure 1: The interface of Paper Tuner with sliders, which allow adjusting the importance of different relevance aspects

results. Users can adjust the weight of each source from 0 to 10 by sliding to the right (increase) and left (decrease). **Section B** displays a stacked relevance bar next to each result. The full length of the bar displays the combined relevance of a recommended item to the target user. Each colored segment displays how much a specific source contributed to the total relevance given the current position of the source slider. The segments of the stacked bars update each time the user changes the sliders, i.e., the length of the “green” section will increase when the green slider is moved right. The ranked list of results also provides details for each recommended item (Figure 1: **Section C**). Users can click on the link on *Paper Title* and *Author(s)* columns to get more information such as the abstract of the publication, other people who are attending the presentation session, etc. More information about the relevance function is available at (Rahdari and Brusilovsky 2019)

The Design of Reversible Relevance Sliders

The base version of the *Paper Tuner* presented above enables users to have more control over the fusion of different relevance sources. The ability to choose and weight sources supports a range of recommendation cases. However, earlier research (Tsai and Brusilovsky 2018) indicates that this simple fusion might not be efficient for more complex cases when positive relevance (similarity) generated by one source has to be combined with negative relevance (dissimilarity) from the other source. For instance, to find popular publications which are outside of current research interests or find relevant papers authored by junior authors with few accumulated bookmarks.

To address this weakness, we attempted to extend the relevance slider interface with an option to reverse each relevance direction. We developed two versions of reversible sliders: one option converted reversed similarity into positive “dissimilarity” and used regular stacked bars (left design on Figures 2 and 3), while another option treated reversed similarity as a negative number and used balanced design for the stacked bars with positive evidence on the right and negative on the left (right design on Figures 2 and 3). We conducted a study to choose between the two designs. In the study, 66.7% of participants favored the first design. Based on the results of the pre-study we developed an enhanced version of Paper Tuner (*Paper Tuner+*), which enables users to reverse the direction of relevance for each of the contributing sources, i.e., use information from this

source to rank items from smaller to larger relevance. In this interface, users can use the reverse function to inverse the relevance value associated with each criterion. By clicking on the reverse checkbox for each criteria, the relevance values generated by this source will change from positive to negative. To stress it, the relevance value generated by this source is displayed as dissimilarity (a meshed pattern) in stacked color bars. Dissimilarity is defined as the distance between the current value of relevance and the maximum available value for that criteria. For example, in Figure 1 the relevance of Publication Popularity source is reversed and purple bars representing this source in section B are shown in the meshed pattern. It reminds users that longer purple bars correspond to less popular papers.

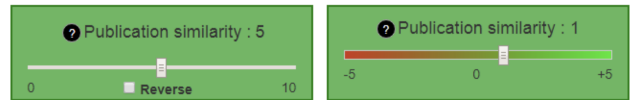


Figure 2: Reversible slider designs



Figure 3: Stackable relevance bar designs

Evaluation

We conducted a user study to evaluate two versions of the controllable user interface: *Paper Tuner* and *Paper Tuner+* against a non-controllable baseline. Our goal was to gain an understanding of users’ information seeking workflows in the tasks of finding conference presentations. This section summarizes the data and the study procedure. The next sections review the results obtained from log analysis and users’ feedback.

Data and Participants

To make the study similar to a real conference recommendation context, we used data from the *Conference Navigator* system. A total of ten doctoral students from a large research university were recruited for this user study. To make sure the participants are knowledgeable in conducting research work and attending academic conferences, we recruited second to fourth-year students (M=2.20, SD=1.14) who had at least one published research paper (M=7.20, SD=5.92). All participants had no prior knowledge of the *Conference Navigator* system.

Task	Action	Baseline		Paper Tuner		P	Paper Tuner+		P
		M (SE)	User Count	M (SE)	User Count		M (SE)	User Count	
T1	Click Title	3.7 (1.82)	9	2.3 (1.14)	8	*	2.10 (1.59)	7	*
	Click Author	1.7 (1.56)	7	0.8 (0.74)	6		0.60 (0.84)	4	
	Time (mins)	3.56 (1.17)	10	2.41 (0.70)	10	*	2.26 (0.80)	10	*
	Tune Sliders	-	-	16.30 (3.31)	10		14.40 (1.95)	10	
	Reverse Function	-	-	-	-		3.80 (1.22)	10	
T2	Click Title	2.70 (1.88)	8	1.80 (1.46)	7		1.60 (1.26)	8	*
	Click Author	1.1 (1.1)	6	0.7 (0.78)	5		0.6 (0.84)	4	
	Time (mins)	3.75 (1.28)	10	2.45 (0.94)	10	*	2.12 (0.95)	10	*
	Tune Sliders	-	-	14.80 (3.18)	10		13.50 (1.50)	10	
	Reverse Function	-	-	-	-		1.30 (1.33)	6	

Table 1: User action summary: the table shows the statistics of user interaction while solving the two tasks.

Experiment Design and Procedure

The study was designed to compare the experimental interfaces (*Paper Tuner* and *Paper Tuner+*) with a baseline (*Baseline*) ranked list interface. The baseline included sections B and C (Figure 1), but had section A disabled. All participants were asked to use each interface for two information seeking tasks and to fill out a post-stage questionnaire at the end of their work with each interface. At the end of the study, the participants were asked to pick their preferred interface. To minimize and control the learning effect, we used data from different conferences at each stage and the order of interfaces and conference data was randomized.

After a 30 minutes system tutorial, we asked the participants to conduct two realistic information seeking tasks for a *seed* scholar who has published three conference papers in the conference series, which we used for the study. Pretending to be the seed scholar, the participants were expected to complete the following tasks using each of the three compared interfaces:

- **Task 1** : To expand your understanding of the field, your advisor has asked you to find four conference papers, which are favored by the conference attendees but not relevant to your current research interests.
- **Task 2** : To expand your reading list, your advisor has asked you to find four conference papers, which are relevant to your current research interest and written by popular authors in this conference series.

The participants were asked to pick suitable conference presentations based on their best judgment in each task. When designing the tasks, we focused on multiple aspects of relevance, which are typical for real-world information seeking in a conference context. We consider task 1 is dissimilarity-oriented (fusing positive relevance from one source with negative relevance from another) and task 2 is similarity-oriented (fusing positive relevance from two sources). In addition to explicit user feedback, we also collected data from user logs including total time spent, number of clicks and the use of the reverse functionality while completing each task.

Results

User Log Analysis

Table 1 shows the system usage for the three interfaces. The data indicate that the participants extensively used the sliders to complete the tasks. In the group of *Paper Tuner* and *Paper Tuner+*, all participants adopted the sliders to complete the tasks. In the *baseline* group where the sliders were not available, the participants inspected paper and author information more frequently to find suitable presentations. We found that the number of clicks on paper title (*Click Title*) was significantly lower in the two experimental interfaces (compared to baseline), which also contributed to a significantly shorter searching time for the users to complete the tasks.

For *Paper Tuner+*, we found that all participants enabled the *Reverse Function* in the dissimilarity-oriented task (task 1) while only six users tried it in task 2. When the reverse function is provided, the user required less slider tuning as well as fewer clicks on the title and author links for completing both of the proposed tasks. The pattern of tuner re-weighting is task-dependent but similar in the two tuner interfaces. As expected, the participants extensively used the *Publication Similarity* (M=6.20, SD=1.47) and *Publication Popularity* (M=5.20, SD=0.91) in task 1, but switched to *Publication Similarity* (M=5.90, SD=1.31) and *Author Popularity* (M=4.70, SD=0.85) for task 2.

User Explicit Feedback Analysis

As is shown in Figure 4, the interface with reverse function (*Paper Tuner+*) received significantly higher ratings in all aspects ($p \leq 0.05$) except the “Ease of Use” (Q6). It points that although the *Paper Tuner+* interface was the hardest to use (among all three interfaces) users preferred this interface in all other aspects. This pattern was also observed for the *Paper Tuner* interface, which was always in the second place, ahead of the baseline for all criteria with the exception of the “Ease of Use”. This shows that using the sliders in general lead to higher satisfaction than using the non-controllable *baseline* interface, but for the price of ease-of-use.

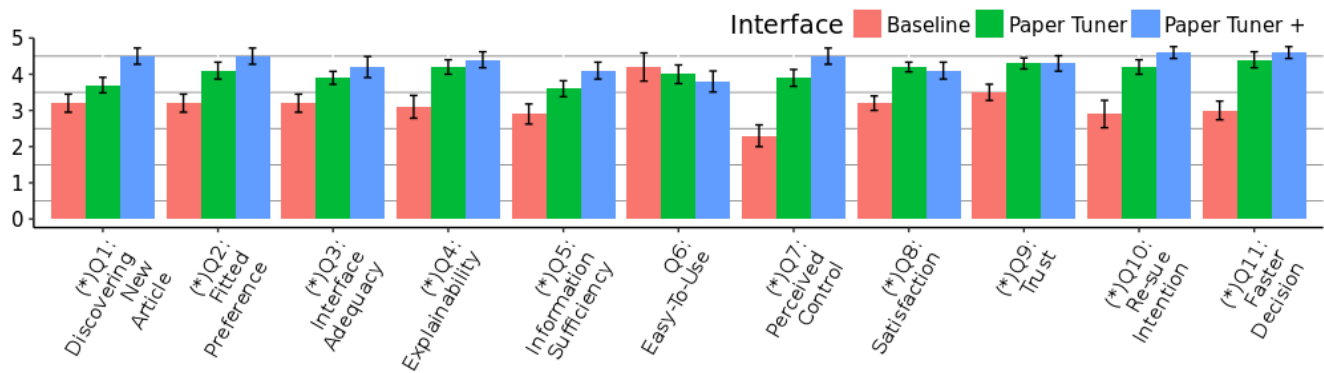


Figure 4: User feedback analysis: the result shows that the *Paper Tuner+* interface received a significantly higher rating

Summary and Discussion

In this paper, we presented the design and the study of two user-controlled interfaces for research paper recommendation at academic conferences. One of these designs (*Paper Tuner*) re-examined in a new context a popular slider-based approach to control and display hybrid recommendations. Another design (*Paper Tuner+*) extended *Paper Tuner* with an ability to reverse relevance for each contributing source. We evaluated these designs against a non-controllable baseline in a controlled user study, which simulated a realistic recommendation context with complex tasks.

The analysis of data obtained from the explicit feedback from the participants and log analysis provided valuable information about the values and trade-offs of the new interfaces. As we expected, the users of both experimental interfaces extensively used sliders to solve the recommendation tasks. Most actively, they worked with relevance dimensions required by the task. Moreover, participants actively used reverse functionality to answer the first task where the ability to rank items by dissimilarity to user’s publication could be of help. The extensive use of the control functionality hints that the decrease in problem-solving time observed in the controllable interfaces might be associated with the provided controllability.

User explicit feedback provides additional data in favor of the controllable interfaces, but also demonstrated that there is a cost associated with controllability. The increase of controllability from baseline to *Paper Tuner* to *Paper Tuner+* lead to the increased user satisfaction in all aspects except for ease of use, which followed a reverse trend. While the most advanced *Paper Tuner+* was picked as the top system in respect to user confidence in results, the baseline was the leader in the low learning cost.

While the results of the lab study provided good insights about a controllable hybrid recommendation, it had some limitations. Most importantly, the size of the focus group study was limited and the use of Ph.D. students who pretended to be a real conference attendees is just an approximation of real users. In our future work, we expect to obtain some additional evidence by assessing the system in a real-life scenario i.e., during a conference with real conference attendees as users.

References

- Ahn, J.-w.; Brusilovsky, P.; He, D.; Grady, J.; and Li, Q. 2008. Personalized web exploration with task models. In *17th international conference on World Wide Web, WWW '08*, 1–10. ACM.
- Bostandjiev, S.; O’Donovan, J.; and Höllerer, T. 2012. Tasteweights: A visual interactive hybrid recommender system. In *6th ACM Conference on Recommender System*, 35–42.
- Burke, R. 2007. Hybrid web recommender systems. In Brusilovsky, P.; Kobsa, A.; and Neidl, W., eds., *The Adaptive Web: Methods and Strategies of Web Personalization*, volume 4321 of *Lecture Notes in Computer Science*. Berlin Heidelberg New York: Springer-Verlag. 377–408.
- di Sciascio, C.; Sabol, V.; and Veas, E. E. 2016. Rank as you go: User-driven exploration of search results. In *21st International Conference on Intelligent User Interfaces (IUI'16)*, 118–129. ACM.
- O’Donovan, J.; Smyth, B.; Gretarsson, B.; Bostandjiev, S.; and Höllerer, T. 2008. PeerChooser: Visual interactive recommendation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08*, 1085–1088. New York, NY, USA: ACM.
- Parra, D., and Brusilovsky, P. 2015. User-controllable personalization: A case study with setfusion. *International Journal of Human-Computer Studies* 78:43–67.
- Rahdari, B., and Brusilovsky, P. 2019. User-controlled hybrid recommendation for academic papers. In *24th International Conference on Intelligent User Interfaces (IUI '19 Companion), March 17–20, 2019, IUI '19*.
- Schafer, J. B.; Konstan, J. A.; and Riedl, J. 2002. Meta-recommendation systems: User-controlled integration of diverse recommendations. In *Proceedings of the Eleventh International Conference on Information and Knowledge Management, CIKM '02*, 43–51.
- Tsai, C.-H., and Brusilovsky, P. 2018. Beyond the ranked list: User-driven exploration and diversification of social recommendation. In *23rd International Conference on Intelligent User Interfaces, IUI '18*, 239–250. New York, NY, USA: ACM.