

Acceptance Issues of Personality-based Recommender Systems

Rong Hu

Human Computer Interaction Group
Swiss Federal Institute of Technology (EPFL)
CH-1015, Lausanne, Switzerland

rong.hu@epfl.ch

Pearl Pu

Human Computer Interaction Group
Swiss Federal Institute of Technology (EPFL)
CH-1015, Lausanne, Switzerland

pearl.pu@epfl.ch

ABSTRACT

To understand users' acceptance of the emerging trend of personality-based recommenders (PBR), we evaluated an existing PBR using the technology acceptance model (TAM). We also compare it with a baseline rating-based recommender in a within-subject user study. Our results show that while the personality-based recommender is perceived to be only slightly more accurate than the rating-based one, it is much easier to use. The side-by-side comparison also reveals that users significantly favor the personality-based recommender and have a significantly higher intention to use such a system again. Therefore, we believe that if users accepted rating-based recommenders, they are most likely to accept personality-based recommenders and personality-based recommenders have a high likelihood to be widely adopted despite the fact that rating-based recommenders are now the industry norm. We further point out some preliminary guidelines on how to design personality-based recommender systems.

Categories and Subject Descriptors

H.5.2 [Information interfaces and presentation]: User Interfaces – *evaluation/methodology, interaction style*; H.1.2 [Models and Principles]: User/Machine Systems – *Software Psychology*

General Terms

Design, Experimentation, Human Factors

Keywords

Technology Acceptance Model, Rating-based Recommender, Personality-based Recommender, Personality Quiz, User Study

1. INTRODUCTION

As powerful tools to help users deal with information overload and to provide personalized services, recommender technologies have been incorporated into a number of applications, especially for e-commerce websites [1]. Understanding what impact users' acceptance to such new technologies is critical to enable service providers to achieve the maximum benefits. Jones and Pu [8]

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

RecSys'09, October 23–25, 2009, New York, New York, USA.
Copyright 2009 ACM 978-1-60558-435-5/09/10...\$10.00.

conducted an experiment to compare two popular music recommender websites, Pandora and Last.fm, to understand such issues. The study not only discovered users' preference between these two systems, but also revealed insights on the user adoption process. Analyzing both the objective and subjective variables, they showed that a simple interface design, little initial effort requirement, and good recommendation quality (e.g. accuracy, novelty and enjoyability) are the most influential factors that contribute to users' adoption of music recommenders.

Recently, inspired by findings in psychology, personality-based recommender systems are increasingly attracting the attention of researchers and industry practitioners [5, 9, 10]. Research has shown that personality is an enduring and primary factor that determines human behaviors [2] and that there are significant connections between personality and people's tastes and interests [5, 9]. It is reasonable to believe that personality-based recommenders could provide more personalized information/services, since they understand the customers better from the psychological perspective. As this is still an emerging trend, little research and user validation work have been done on this topic. However, the existing studies have shown that personality-based recommenders could be more effective and applicable to a broader range of applications.

In our previous work, we demonstrated that personality quiz-based recommender systems have the potential to be alternative to commonly used rating-based systems [11]. However, we questioned to what extent users would like to use the new personality-based recommender technology, and which design and algorithm aspects would influence users' acceptance.

The contribution of this paper is to utilize a well-studied intention model, the technology acceptance model (TAM) [4], to explore potential user acceptance issues of personality-based recommender systems, especially those based on personality quizzes. An empirical user study was conducted, comparing a personality-based recommender with a rating-based recommender in terms of the acceptance constructs in the TAM. The results validate the fact that perceived usefulness and perceived ease of use are two significant factors affecting users' acceptance to the personality-based recommender systems. Furthermore, the study pointed out some preliminary research directions on design issues. Despite there being other aspects that affect users' acceptance, such as interface display [3], we focused on the initial preference elicitation processes and recommendation approaches which are the two key innovations of personality-based recommender systems.

The remainder of this paper is organized as follows. Section 2 presents the technology acceptance model used in our study. Section 3 describes our user study in detail, including a brief introduction of the two compared systems, evaluation criteria, materials and experiment design. Section 4 presents the results analysis and discussion, followed by the section on conclusions.

2. TECHNOLOGY ACCEPTANCE MODEL

The Technology Acceptance Model (TAM) was introduced by Davis in the field of Information Systems [4]. Taking TRA (Theory of Reasoned Action) as the theoretical basis, TAM aims to provide an explanation to the determinants of computer technologies acceptance. In TAM, two particular beliefs, perceived usefulness and perceived ease of use, were postulated to determine an individual's intention to use a system. Perceived usefulness was also seen as being directly influenced by perceived ease of use (see Figure 1). TAM is considered to be a robust model for explaining user potential adoption towards technologies and is applied in various contexts, such as word processors, spreadsheets, e-mails, voice mails, and online trust studies [6]. In our present work, we utilize TAM to investigate the potential acceptance issues of personality-based recommender systems.

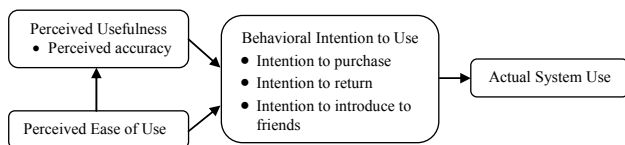


Figure 1. The Technology Acceptance Model (TAM)

3. USER STUDY

This user study was conducted to compare one personality-based system, Whattorent (<http://www.whattorent.com>), with the well-known rating-based recommender system, MovieLens (<http://www.movielens.org>), based on TAM.

3.1 Compared Systems

MovieLens is a research site run by the GroupLens research group at the University of Minnesota. It learns users' preferences from their ratings and uses collaborative filtering technology to recommend movies. At the time of our experiment, when a new user logs on, he/she was required to rate at least 15 movies which he/she had already seen.

WhattoRent provides the movie recommendation service according to users' personality and mood. It was built based on the LaBarrie Theory which states that "movie viewers emotionally interact with a film in the same manner that they interact with other human beings".¹ Accordingly, the developers designed a personality quiz with 20 questions, such as, "What would you say if a friend said, 'Let's go see a great new independent foreign movie with English subtitles that has some powerful yet difficult to handle scenes?'"

The designers hope that by asking seemingly unimportant life experiences they will obtain reliable responses. Thereby, WhattoRent tries to predict users' movie preferences and provide recommendations based on the deciphered personalities.

¹ <http://www.whattorent.com/theory.php>

3.2 Evaluation Criteria

We first establish the following specific evaluation criteria in terms of the three constructs of TAM (see Figure 1).

Perceived Accuracy. Even though the usefulness of a recommender is determined by various aspects (e.g. enjoyability or novelty), accuracy is still a primary evaluation criterion [7]. Thus, we adopted the perceived accuracy, i.e., how well a recommendation corresponds to users' preferences, as the perceived usefulness. More concretely, it is measured by both the participants' ratings of the recommended movies and the post-study questionnaire measuring their satisfaction.

Ease of Use. The perceived ease of use is referred to the users' perceived cognitive effort, that is, the amount of effort users feel they consumed. Since our study focuses on the novel preference elicitation process in personality-based recommenders, we specifically measured the perceived effort spent in the sign up phase. In addition, we recorded the actual completion time (i.e., the objective effort) to compare with the perceived effort.

Behavioral Intentions to Use. We evaluated this from three aspects: intention to purchase, intention to return and intention to introduce this system to friends. The first two measurements have been used in previous studies [3]. We added "intention to introduce to friends" to measure whether users would extend the acceptance range to their social community.

3.3 Materials and Participants

In order to avoid potential influences from the differences of the two interfaces [3], we rearranged the recommendation interfaces of the two systems by utilizing an open-source filtering HTTP proxy (PAW)² to render the two interfaces as similar as possible and show only one recommendation at a time. Employing PAW, we also appended a trailer link and rating buttons next to each recommended movie to help participants easily evaluate recommendations. A total of 30 participants (11 females) were recruited in our study. Every user was given a cash voucher valued at 20 Swiss Francs (a little more than \$18 on August 12, 2009). The details on the experiment setup can be found in [11].

3.4 Experiment Design

Our user study was conducted with a within-subjects design for two reasons. We want to avoid the possible negative influences from the individual differences, and we want to make a direct comparison of users' side-by-side evaluation of two elicitation methods (personality quiz vs. rating). Therefore, each of our participants evaluated both systems one after the other. To minimize the carryover effects (both practice and fatigue effects) inherent to the within-subjects design, we developed two experimental conditions and randomly but evenly assigned the 30 participants to those conditions. Therefore, 15 users in one condition evaluated MovieLens first and then WhattoRent. The other condition had the reverse sequence. When evaluating each system, all participants were required to register a new account in order to avoid any impact from the usage history and were asked to rate 6 recommended movies as a measurement of the perceived recommendation accuracy.

² <http://paw-project.sourceforge.net>

4. RESULTS AND ANALYSIS

4.1 Perceived Accuracy

The perceived accuracy was assessed by a positive question and a negative question, which are significantly related in the 2-tail Pearson Correlation (correlation coefficient = -0.806, $p < 0.001$). The averages of users' responses and the corresponding questions are shown in Figure 2. To facilitate comparison, we reversed the responses from the negative question, meaning that the higher the rating level, the more the participants are satisfied with the recommendations. Although the personality quiz-based system obtained higher scores, the difference between the two systems is not significant by the paired t-test ($t = 1$, $p = 0.326$). On the other hand, more movies recommended by the personality-based system were evaluated as interesting (54.9%), compared to those recommended by the rating-based system (42.7%). Similarly, its distribution difference is not significant by the chi-test (chi square = 7.274, $p = 0.122$).

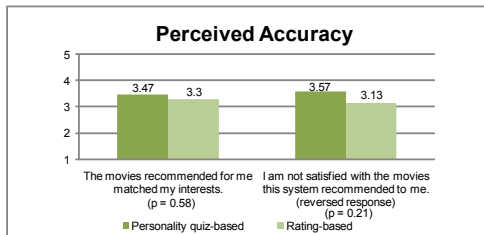


Figure 2. Mean difference of perceived accuracy.

4.2 Ease of Use

To compare users' perceived cognitive effort, we also asked two interrelated questions (correlation coefficient = -0.55, $p < 0.001$, in 2-tail Pearson Correlation). The results are presented in Figure 3. We reversed the responses from the negative question as well. The higher it scores, the easier it is to use. The results show that the overall cognitive effort is perceived to be significantly lower in the personality quiz-based system than in the rating-based system ($p < 0.001$). The difference achieves 52.5%. For the subjective effort, participants spent significantly more time with the rating-based system (6.8 minutes on average for personality-based system vs. 18.7 minutes for the rating-based system, $t = -9.49$, $p < 0.001$).

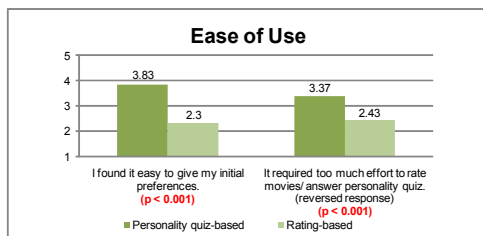


Figure 3. Mean difference of perceived cognitive effort.

4.3 Behavioral Intentions to Use

Figure 4 shows the results of users' acceptance intentions towards the two systems. Participants only expressed a slightly higher intention to rent or buy DVDs of the recommended movies proposed by the personality-based system rather than the rating-based system ($t = 0.108$, $p = 0.915$). On the other hand, they stated a significantly stronger intention to reuse the personality-based system in the future ($t = 2.163$, $p < 0.05$). Likewise, more

participants would introduce the personality-based system to their friends, even though the difference is moderately significant ($t = 2.02$, $p = 0.052 < 0.1$).

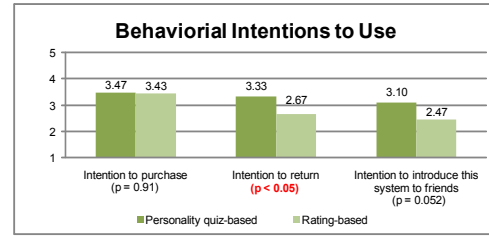


Figure 4. Mean difference of participants' behavioral intentions

Finally, the results on system preference show that 53% of users (16 out of 30) preferred the personality-based system over 13% of users who preferred the rating-based system ($p < 0.001$ with the chi-test).

In conclusion, if users accepted rating-based recommenders, they are most likely to accept personality-based recommenders. The side-by-side comparison also reveals that users significantly favor the personality-based recommender and have a significantly higher intention to use such a system again. Therefore, we believe that personality-based recommenders have a high likelihood to be adopted despite the fact that rating-based recommenders are now the industry norm.

4.4 Design Guidelines

To understand factors that influence users' acceptance of recommender systems at a more detail level and to identify some initial guidelines for the design of such systems, we further investigated the correlations among the measured constructs. Figure 5 and 6 show the correlation results for perceived accuracy, cognitive effort, and actual completion time and acceptance intentions for the respective system. First of all, perceived accuracy is highly correlated with users' acceptance intentions in both systems, especially the intention to purchase (0.82 in the personality-based system and 0.582 in the rating-based system). Therefore we propose:

Guideline 1: Improve perceived recommendation accuracy to ultimately convince users to purchase.

Perceived cognitive effort is another critical factor. It is negatively correlated with perceived accuracy and all behavior intentions ($p < 0.05$) for the personality-based recommender. That is, less perceived cognitive effort relates to higher perceived accuracy and stronger acceptance intentions. While we know from TAM that effort is directly related to acceptance, the first correlation link between accuracy and effort is surprising. This implies that a less demanding signing up process or preference elicitation process implies a higher perceived accuracy. On the other hand, the actual completion time has no significant correlation with cognitive effort ($p = 0.783$) in the personality-based recommender. This suggests that a long completion time does not necessarily imply that users would perceive more cognitive effort, and vice versa. Therefore, we reckon that there may be some other factors affecting users' perceived cognitive effort. For example, it has been demonstrated that perceived enjoyment is positively related to perceived effort [6]. Therefore, a more entertaining signing up process may cause

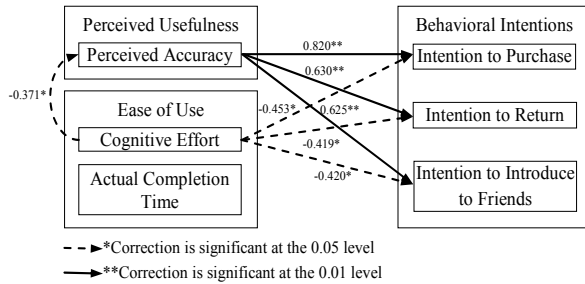


Figure 5. Correlation results for the personality-based system (Pearson Correlation).

users to perceive a more accurate recommender and because of that users may be more ready to accept such recommenders. Therefore, we propose:

Guideline 2: Design entertaining personality quizzes.

Figure 6 shows the correlations results in the rating-based system. Users' acceptance intentions are significantly correlated with only perceived accuracy ($p < 0.01$). Furthermore, there is no significant correlation between perceived accuracy and cognitive effort ($p = 0.52$). In other words, even if users perceive they spend less effort in the sign up phase, it cannot predict that users would express higher perceived recommendation accuracy. Cognitive effort is highly positively correlated with actual completion time for this system ($p < 0.01$). While these correlation results did not help us establish more novel guidelines, we focused on users' comments.

Participants who favored the rating-based system said they had more controls over the recommendation process in the rating-based system since they could directly change their profiles by updating their ratings to movies. Moreover, the participants had more confidence in the recommendations since they knew how the system works. Contrarily, they thought there was a certain magic in the way the personality-based systems came up with recommendations. Therefore, we propose:

Guideline 3: Increase transparency and user control.

5. CONCLUSION

In this paper, we conducted a within-subject user study to investigate potential user acceptance issues of the personality-based recommender systems based on the Technology Acceptance Model. Consistent with prior studies, our experiment results validate that perceived accuracy plays a predominant role in determining user acceptance of a recommender system. Perceived ease of use has a significant correlation with users' acceptance intentions and perceived accuracy in the personality-based recommender system tested. In our future work, we will investigate how to increase users' perceived usefulness and ease of use in personality-based recommender systems, and accordingly hone effective design guidelines.

6. ACKNOWLEDGMENTS

We thank the EPFL and Chinese Government for sponsoring the reported research work. We are grateful to the participants of our user studies for their patience and time.

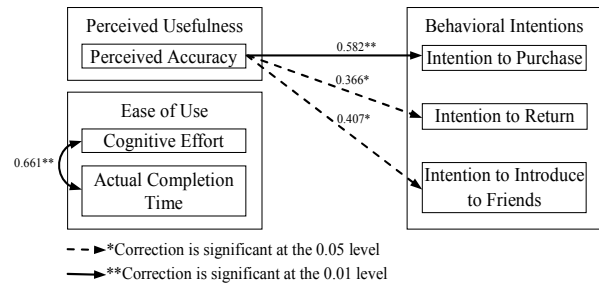


Figure 6. Correlation results for the rating-based system (Pearson Correlation).

7. REFERENCES

- [1] Adomavicius, G. and Tuzhilin, A. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Trans. Knowledge and Data Eng.*, 17, 6(2005), 734-749, 2005.
- [2] Carl, J. 1923. *Psychological Types*. New York, Harcourt Brace.
- [3] Pu, P. and Chen L. Trust Building with Explanation Interfaces. In *Proceeding of IUI'06*, pages 93-100, Sydney, Australia, 2006.
- [4] Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982-1003, 1989.
- [5] Dunn, G., Wiersema, J., Ham J. and Aroyo, L. Evaluating Interface Variants on Personality Acquisition for Recommender Systems. To appear in *proceedings of UMAP'09*, Trento, Italy, 2009.
- [6] Hassanein, K. and Head, M. Building Online Trust through Socially Rich Web Interfaces. In *Proceedings of PST'04*, Fredericton, Canada, October 13-15, 15-22, 2004.
- [7] Herlocker, J., Konstan, J.A., Terveen, L.G. and Reidl, J. Evaluating Collaborative Filtering Recommender Systems. *ACM Trans. on Information Systems*, 22(1), 5-53, 2004.
- [8] Jone, N. and Pu, P. User Technology Adoption Issues in Recommender Systems. In *Proc. of NAEC'07*, p. 379-394, 2007.
- [9] Lin, C. and McLeod, D. Exploiting and Learning Human Temperaments for Customized Information Recommendations. *IMSA*, 218-233, 2002.
- [10] Nunes, M. A. S. N., Recommender System Based on Personality Traits. PhD thesis. Université Montpellier 2-LIRMM. 2008.
- [11] Hu R. and Pu P. A comparative user study on rating vs. personality quiz based preference elicitation methods. In *Proceedings of IUI'09*, pages 367-372, Sanibel Island, Florida, USA, February 08 - 11, 2009.