Recommender Systems: Investigating the Impact of Recommendations on User Choices and Behaviors

Robin Naughton, Xia Lin

The iSchool at Drexel, College of Information Science and Technology
3141 Chestnut Street, Philadelphia, PA 19104 USA
{rnaughton,xlin}@ischool.drexel.edu

ABSTRACT

Recommender systems have been used in many information systems, helping users handle information overload by providing users with a way to receive specific recommendations that fulfill their information seeking needs. Research in this area has been focused on the recommender system algorithms and improving the core technology so that recommendations are robust. However, little research is focused on the user-centered perspective of recommendations provided by recommender systems and the impact of recommendations on user's information behaviors. In this paper, we describe the results of an exploratory survey study on a book recommender system, LibraryThing, and the impact of recommendations on user choices, particularly what users do as a result of getting a recommendation. Based on survey respondents, our results indicate that users prefer member recommendations rather than the algorithm-based automatic recommendations and about two third of users that responded are influenced by the recommendations in their various information activities.

Categories and Subject Descriptors

H5.3 [Information Interfaces and Presentations]: Group and Organization Interfaces *collaborative computing, organizational design, web-based interaction*

General Terms

Computer applications, Design, Evaluation

Keywords

Recommender systems, user-centered design, survey study, user information behaviors

1. INTRODUCTION

Recommender systems offer a solution to the problem of information overload by providing a way for users to receive specific information that fulfill their information needs. These systems help people make choices that will impact their daily lives and according to Resnick and Varian [10], "Recommender Systems assist and augment this natural social process." As more information is produced, the need and growth of recommender systems continue to increase. One can find recommender systems in many domains ranging from movies (MovieLens.org) to books (LibraryThing.com) to e-commerce (Amazon.com). Research into this area is also growing to meet the demand, focusing on the core recommender technology and evaluation of recommender algorithms. However, there's a need for user-centered research into recommender systems that looks beyond the algorithms to people's use of the recommendations and the impact of those recommendations on people's choices. With this in mind, the study objective is to understand the impact of recommendations on user choices and behavior through the use of recommender systems, and this paper presents the results from an exploratory survey of users of a book recommender system, LibraryThing, focusing on whether users follow the recommendations they receive and how those recommendations impact their choices, particularly what users do as a result of getting a recommendation.

2. LITERATURE REVIEW

2.1 Recommender Systems

Resnick and Varian [10] chose to focus on the term "recommender system" rather than "collaborative filtering" because "recommender system" may or may not include collaboration and it may suggest interesting items to users in addition to what should be filtered out. By using the term "recommender system," it becomes clear that the system is not just about the algorithm, but rather the overall goal. It also becomes an umbrella term for different types of recommender systems that uses various algorithms to achieve their goals. Recommender systems can have algorithms that are constraint-based (question and answer conversational method) [3], content-based (CB) (item description comparison method), collaborative filtering (CF) (user ratings and taste similarity method), and hybrid (a combination of different algorithms) [7, 15]. The collaborative filtering technique has gained in popularity over the years [5] and the social networking aspects help to strengthen the filtering techniques. The hybrid technique combines collaborative filtering with content-based techniques to capitalize on the strength of each method.

2.2 Evaluation of Recommender Systems

Research on recommender systems algorithms is very active and seeks to enhance current recommender systems. However, as recommender systems improve, it is important that there is user-centered research on the evaluation of recommender systems. According to Herlocker, et al [5], "To date, there has been no published attempt to synthesize what is known about the evaluation of recommender systems, nor to systematically understand the implications of evaluating recommender systems for different tasks and different contexts." Herlocker, et al [5] focused extensively on the problems of evaluating recommender systems, presenting methods of analysis and experiments that provides a framework for evaluation. Identifying three major challenges, they point out that algorithms perform differently on different datasets, evaluation goals can differ, and deciding on measurement in comparative evaluation can be a challenge [5]. Hernandez del Olmo and Gaudioso [6] proposed an alternative evaluation framework for recommender systems that focuses on the goal of the recommender system. They indicate that there's a shift in the field to a broader and general definition of recommender systems that focuses on guiding users to "useful/interesting objects" [6]. This redefining of the recommender system goals also frames the redefining of the recommender system framework, implying that evaluation can be based on goal achievement of guiding the user and providing useful/interesting items [6]. By dividing recommenders into these subsystems, the authors suggest that each recommender system will have one of the two subsystems more active than the other and the closer they are in terms of activity, the closer they are to achieving the global objective of the recommender system.

The work of Herlocker, et. al [5] and Hernandez del Olmo and Gaudioso [6] offer evaluation frameworks that function across different domains and algorithms. However, they are still steps away from focusing on evaluating recommender systems from the user perspective. A few steps closer is research focused on improving the user experience. Celma and Herrera [2] "Item- and User-centric evaluation" methods to identify novel recommendations based on CF and CB systems, and found that users perceive recommendations through CF are of higher quality "even though CF recommends less novel items than CB" [2]. O'Donovan and Smyth's [8-9] research on trust in recommender systems defines two trust levels, context-specific and system/impersonal trust to help to create and preserve accuracy and robustness within recommender systems. Ziegler and Golbeck's [16] research into trust and interest similarity focused on the link between trust and a person' interest, concluding that the more trust users have between each other, the more their ratings are similar. Tintarev [13] and Tintarev and Masthoff [14] argue for effective explanations that can increase user trust, help users make good decisions and improve user experience.

Although much of the research is based on improving the algorithms, the literature shows movement towards a focus on the user. Tintarev and Masthoff [14] use of two focus groups to determine how participants would like to be recommended or dissuaded from watching a movie indicate a change in the field towards direct contact with users. Accuracy metrics of algorithms is not enough to determine the true impact on user choices.

3. LIBRARYTHING

Book recommender systems (LibraryThing, GoodReads, BookMooch, Amazon, All Consuming, Shelfari, etc.) allow users to catalogue books, and receive and share recommendations within a social community. Since its launch in 2005, LibraryThing has grown to over 920,000 users with the largest group representing librarians, 45.5 million books have been catalogued, and where some book recommender systems offer a single algorithm, LibraryThing has multiple recommender algorithms [1]. According to the founder, Tim Spalding, "We've got five algorithms so far, and a few more I haven't brought live, or which lie underneath the current ones. … LibraryThing's data is particularly suited to it, the books you own being a much better representation of taste than the books you buy on a given retailer" [11]. It is a robust book recommender system with a strong social network that offers a fertile area for user-centered research.

LibraryThing users can add book titles to their accounts and receive book recommendations directly from LibraryThing algorithms (automatic recommendations) or other users of the website (member recommendations). Member recommendations are submitted through a manual process that allows LibraryThing users to submit recommendations for any book by going to the book's recommendation page. The majority of recommendations are automatic and for each book, LibraryThing offers six types of recommendations: 1) LibraryThing Combined Recommendations, 2) Special Sauce Recommendations, 3) Books with similar tags, 4) People with this book also have... (more common), 5) People with this book also have... (more obscure), and 6) Books with similar library subjects and classification. Most of the titles of the recommendation types are self-explanatory in that a user can easily get the general idea of the type of recommendations being offered. For example, the "LibraryThing Combined Recommendations" represents a combination of other types of automatic recommendations. However, the "Special Sauce Recommendations" seems to be the one title that is not self-explanatory and offers no immediate understanding of what users should expect. Spalding says, "Our Special Sauce Recommendation engine is the only one we don't talk about how it works," [11].

4. RESEARCH DESIGN

This study used an online survey ("LibraryThing Recommendation Impact Survey") to explore the impact of LibraryThing recommendations on user choices. No personal or identifying information was collected. There were 10 questions using both open and closed question types. Two of the ten questions focused on capturing demographic data (gender and age range) so that responses could be grouped within a larger context. The other eight questions focused specifically on LibraryThing recommendations and user preferences, influences and actions. Before administering the survey, permission was obtained from Tim Spalding, and an IRB approval from the University.

4.1 Implementation

On October 27th, 2009, the recruitment letter with a link to the survey was posted to "Book Talk," a LibraryThing group recommended by Tim Spalding as a place for major discussions. Spalding pointed out that postings can be tagged for spamming if posted to multiple groups and the goal was to reach the LibraryThing users rather than have the posting removed. However, after a few weeks within the "Book Talk" group, the posting was added to the "Librarians who LibraryThing" group because they were one of the largest groups of LibraryThing users, which helped with getting survey respondents. The posting was repeatedly checked to make sure that it was still on the first page of the active group discussion and if it wasn't, it was adjusted to remain prominent to improve visibility and opportunity for user response. The survey was posted on LibraryThing for five months, from October, 27th, 2009 to March 27th, 2010.

4.2 Participants

Participants were 18 years and older who have previously or were currently using LibraryThing that volunteered to take the survey by clicking the link to the survey from the LibraryThing group. The expectation was that the survey may receive about 100 self-selected respondents and within the five months, there were 62 survey respondents.

5. RESULTS

The data gathered from the survey used descriptive statistics to generate percentages and iterative pattern coding of qualitative data to identify major themes [4].

5.1 Demographic

Two demographic questions (gender and age range) helped to frame the population responding to the survey. For gender, there were 50 females (81%) and 12 males (19%) who responded to the survey. All age range groups had at least 3 participants. The 25-34 years old range accounted for 42% (26) of participants and the 45-54 years old range accounted for 26% (16) of participants, representing the two largest groups responding to the survey. Overall, there were no age ranges that had zero participants, but the 55-64 age range was the only group with no male participants.

5.2 Member vs. Automatic Recommendations

In their own words, participants described their preferences regarding automatic and member recommendations, and from the data five participant preference categories were developed: automatic, member, both, neither, and no preference. Of the 62 participants that responded to the survey, the majority 48% (30) preferred member recommendations while only 24% (15) preferred automatic recommendations. The other 28% (17) of the participants preferred neither, both or had no preference (Table 1).

Table 1: User Recommendation Preferences

User Preference	*# of Participants*
Member	30 (48%)
Automatic	15 (24%)
Neither	9 (15%)
Both	4 (6.5%)
No Preference	4 (6.5%)

In addition, there was an even split of participants (50%) between those who have submitted member recommendations and those who have not. Participants were also asked to identify their preference for a specific type of LibraryThing automatic recommendation, the top two preferences were "LibraryThing Combined Recommendations" and "People with this book also have…. (more common)" (Table 2).

Table 2: Users' Most Valuable Automatic Recommendations

# of Participants	*Automatic Recommendation Type*
15	LibraryThing Combined Recommendations
14	People with this book also have… (more common)
12	Other
9	Books with similar tags
5	Special Sauce Recommendations
4	Books with similar library subjects and classification
3	People with this book also have... (more obscure)

5.2.1 Discussion

The data suggested that twice as many participants preferred member recommendations over automatic recommendations. Based on reasons provided by participants, a distinction could be made between preferring member or automatic recommendations. Participants that preferred member recommendations seemed to be interested in the social connection between the recommendation and the recommender where they were able to assess the recommender and recommendation as it relates to their own tastes. As one participant described, "Even though automatic recommendations may more 'accurately measure' my tastes and interests based upon the books I have in my library, I feel recommendations from real human beings have the advantage of the recommender's intuitive understanding of what I would find interesting based upon their own impressions of books they know I've read." Alternatively, participants that preferred automatic recommendations seem to be interested in the logical connection of the recommendation and user libraries where the algorithm looks at all items. As one participant stated, "I prefer automatic recommendations because they are based on all users with a particular book, not just on one member who thinks a book is like another." In both cases, the preference for member or automatic recommendations is influenced by the user's trust in particular aspects of the system, which has an impact on the level of trust that the user has of the system and their fellow users. Research into trust models such as a user's trust in another user based on that other user's profile or a user's trust in the system based on the items can begin to offer another dimension for developing recommendations [8-9].

The top preferences for automatic recommendations (Table 2) suggest that LibraryThing users want recommendation types that are additionally filtered (combined recommendations) and socially connected (people also have). The other preferences suggest that there may be overlap with the combined recommendations, lack of knowledge ("What is special sauce? I missed that!"), or an alternative approach to getting recommendations ("People whose library is similar to mine," "Top 1,000 on my recommendations page," "The stars, recommendations in forums").

Since automatic and member recommendations present different ways of getting recommendations within the system, as expected, Table 1 shows that some participants preferred both (6.5%) or had no preference (6.5%). However, the neither category suggested that participants (15%) actively did not prefer automatic or member recommendations, but instead, preferred to get their recommendations from other sources such as message boards ("message boards on the site--it's much more useful for me to read another member's opinion about a book or to see a dialogue about a book on the message boards than to just see a list") or chat ("The recommendations that I DO pay attention to, however, are the ones made personally from people I regularly chat with on LT, and whose tastes I know I share"). The neither category presents an opportunity to understand why some participants are not using the traditional automatic and member recommendations, and how recommender systems can be improved to service this population that seeks alternative methods of getting recommendations that combine multiple sources. These results also suggest looking at the overall goal of the recommender system to identify how best to guide users and filter content appropriately to satisfy user wants and needs [6].

5.3 Recommendation Impact

Users were asked if they checked their recommendations, what they did with the information, and how it influenced their choices. Table 3 shows that only 8 (13%) participants never checked their recommendations while 46 (74%) participants checked their recommendations daily, weekly or periodically. Most of the 8 (13%) participants that chose "Other" checked their recommendations on a different schedule than what was presented in the survey question.

Table 3: Frequency of User Checking Recommendations

User Checks	*# of Participants*
Periodically	22 (35%)
Weekly	15 (24%)
Daily	9 (15%)
Other	8 (13%)
Never	8 (13%)

After checking their recommendations, 61% (38) of participants read and followed-up on recommendations (Table 4).

Table 4: Participant Follow-up on Recommendations

Follow up	*# of Participants*
Read and follow-up on recommendations	38 (61%)
Only read recommendations	6 (10%)
Never read or follow-up on recommendations	9 (15%)
Other	9 (15%)

Participants were asked to select specific actions that they took as a result of recommendations and could select multiple responses to indicate the types of influence the recommendations had on their choices. As a result, there were 167 responses, which exceed the number of participants (62), with an average of 2.7 responses per participant. Table 5 shows the selection options and the number of responses per selection.

Table 5: Recommendation Influence

Recommendation Influence	*# of Responses*
Added books to my library.	36
Purchased the recommended book or added to a list for purchase.	35
Browsed user libraries that have the recommended book	31
Reminded you of something else.	29
Submitted a recommendation.	19
Other	17

5.3.1 Discussion

It was important to know whether users were actively engaging the recommender system or taking a passive approach by just reading whatever appears on the homepage. The data show that a majority of the participants checked whether they had new LibraryThing recommendations (Table 3) and followed up on those recommendations by adding books to their libraries, purchasing recommended books or putting recommended books on a list to purchase, and browsed other user libraries with recommended book (Table 4). Table 5 shows 17 "Other" responses, suggesting a need for additional options for users to describe the influences of LibraryThing recommendations, such as no influence, added to wishlist within or outside of LibraryThing, borrowed from local library, and discovery research leading to additional information. Most participants, 46 (74%), found LibraryThing recommendations useful and stated that the recommendations helped them to find books they would not have found otherwise. One participant pointed out the international nature of LibraryThing, "Useful as an introduction to unknown authors and series - particularly American titles - often difficult to source in the UK." Nine (15%) participants found the recommendations "somewhat" useful, and 7 (11%) participants did not find recommendations useful. One participant stated, "I suppose I feel the recommendations function is less useful because it doesn't account for shifting literary interests," highlight an issue for user satisfaction and perceived usefulness.

Perceived usefulness is another area of research that can help to shed light on recommender systems from the user's perspective. Swearingen and Sinha's [12] research comparing online and offline recommendations, focused on perceived usefulness and found that what mattered most was whether users got useful recommendations, the reason for using the recommender system. Overall, LibraryThing participants checked, followed, acted upon and found useful the recommendations they received from LibraryThing and on multiple questions, indicated the impact of recommendations on their choices.

6. LIMITATIONS & FUTURE

One limitation of this study is the self-selected nature of the online survey, which limits the respondents to frequent users of LibraryThing who chose to respond to the survey. This can create a self-selected group of users that do not represent the full range of LibraryThing users. As a consequence, the results are not easily generalized to the larger population and an exploratory survey only scratches the surface of the user perspective. However, this research provides a valuable starting point for future research into user experience with recommender systems, particularly focusing on user preference, user actions and perceived usefulness of recommendations. Based on the themes identified, future research would include creating a more robust method of soliciting data directly from users and in-depth analysis of the "other" categories identified as these categories seem to indicate that users are using the system in unexpected ways, which in turn can help to improve recommender systems.

7. CONCLUSION

The main research goal of this study was to explore the impact of recommendations through recommender systems on user choices and behaviors, particularly what users did as a result of getting a recommendation. Much of the literature on evaluation has focused on the algorithms [5-6], but research into trust [8-9, 16], explanations [13-14], design and usefulness [12] are getting closer to the user of the system. Understanding impact directly from users is an important aspect of developing recommender system research on evaluation and this study has contributed to this effort.

For LibraryThing, the results from this exploratory study indicate possible areas of improvement such as limiting automatic recommendation types because participants preferred only 2-3 out of 6 automatic recommendation types, improving submission of member recommendations because twice as many participants preferred member recommendations over automatic recommendations, and providing alternative recommendations from other areas of LibraryThing because participants indicated a growing need to get recommendations from alternative sources such as tags, message boards, and other areas of LibraryThing.

The research has shown that twice as many participants preferred member recommendations over automatic recommendations, and participants checked, followed-up, acted upon and found recommendations useful. The findings indicate that there's more to uncover within the evaluation of recommender system and that users are an important aspect of understanding whether recommender systems are indeed useful and impactful in people's daily lives.

8. ACKNOWLEDGMENTS

We thank Tim Spalding, LibraryThing Founder, and IMLS fellowship funding for making this research study possible.

9. REFERENCES

[1] *LibraryThing*. Available from: http://www.librarything.com.

[2] Celma, O. and P. Herrera, *A new approach to evaluating novel recommendations*, in *Proceedings of the 2008 ACM conference on Recommender systems*. 2008, ACM: Lausanne, Switzerland. p. 179-186.

[3] Felfernig, A. and R. Burke, *Constraint-based recommender systems: technologies and research issues*, in *Proceedings of the 10th international conference on Electronic commerce*. 2008, ACM: Innsbruck, Austria.

[4] Glaser, B.G. and A.L. Strauss, *The constant comparative method of qualitative analysis*, in *The discovery of grounded theory: Strategies for qualitative research*. 1967, Aldine de Gruyter: Hawthorne, NY. p. 101-115.

[5] Herlocker, J.L., et al., *Evaluating collaborative filtering recommender systems.* ACM Trans. Inf. Syst., 2004. **22**(1): p. 5-53.

[6] Hernández del Olmo, F. and E. Gaudioso, *Evaluation of recommender systems: A new approach.* Expert Systems with Applications, 2008. **35**(3): p. 790-804.

[7] Huang, Z., et al., *A graph-based recommender system for digital library*, in *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*. 2002, ACM: Portland, Oregon, USA. p. 65-73.

[8] O'Donovan, J. and B. Smyth, *Trust in recommender systems*, in *Proceedings of the 10th international conference on Intelligent user interfaces*. 2005, ACM: San Diego, California, USA. p. 167-174.

[9] O'Donovan, J. and B. Smyth, *Is trust robust?: an analysis of trust-based recommendation*, in *Proceedings of the 11th international conference on Intelligent user interfaces*. 2006, ACM: Sydney, Australia. p. 101-108.

[10] Resnick, P. and H.R. Varian, *Recommender systems.* Commun. ACM, 1997. **40**: p. 56-58.

[11] Starr, J., *LibraryThing.com: The Holy Grail of Book Recommendation Engines*, in *Searcher*. 2007. p. 25-32.

[12] Swearingen, K. and R. Sinha. *Beyond Algorithms: An HCI Perspective on Recommender Systems*. in *Proceedings in the SIGIR 2001 Workshop on Recommender Systems*. 2001.

[13] Tintarev, N., *Explanations of recommendations*, in *Proceedings of the 2007 ACM conference on Recommender systems*. 2007, ACM: Minneapolis, MN, USA.

[14] Tintarev, N. and J. Masthoff, *Effective explanations of recommendations: user-centered design*, in *Proceedings of the 2007 ACM conference on Recommender systems*. 2007, ACM: Minneapolis, MN, USA. p. 153-156.

[15] Torres, R., et al., *Enhancing digital libraries with TechLens+*, in *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*. 2004, ACM: Tuscon, AZ, USA. p. 228-236.

[16] Ziegler, C.-N. and J. Golbeck, *Investigating interactions of trust and interest similarity.* Decision Support System, 2007. **43**(2): p. 460-475.