Hey Alexa, What Should I Read? Comparing the Use of Social and Algorithmic Recommendations for Different Reading Genres

Huiwen Zhang, George Buchanan(✉), and Dana McKay

iSchool, University of Melbourne, Parkville, VIC 3010, Australia
huiwen.zhang@student.unimelb.edu.au,
george.buchanan@unimelb.edu.au, danamckay@gmail.com

Abstract. Users often seek reading recommendations for what to read, across a variety of topics of interest and genres. While there has been extensive research on the development of recommender algorithms, our understanding of social factors relating to reading recommendation in the digital era is poor. We have no holistic view of how readers interact with diverse resources, social and digital, to obtain reading recommendations. Users can consult computer-generated summaries and human-created reviews. How much or how often the typical user relies on one or other source, or what variations there are by genre of intended reading, are both open questions.

To narrow these research gaps, we conducted a diary study to capture a comprehensive picture of readers’ use of algorithm- and social-sourced information to inform their future reading choices. Based on a qualitative analysis of these diaries, we produced a survey to investigate in-depth readers’ recommendation preferences across fictional reading, factual reading, academic resources, and news and articles. We show that users rely on different sources of recommendation information in different ways across different genres, and that modern social media plays an increasing role alongside established mass media, especially for fiction.

Keywords: Digital reading · Reading recommendations · Social information seeking · Discovery

1 Introduction

The explosion of digital reading material, such as fanfiction, blogs, self-published ebooks, and guidebooks [2] has resulted in a situation where, for readers, ubiquitous internet access isn’t so much a valley of riches as an avalanche. This information overload [1] has made narrowing one’s options as key a feature of the reading experience as finding something at all; one way to address both types of information need is using recommendations.

Recommendations are frequently used to try to identify information, or a book, that matches a user’s preferences or needs. Different services, focused on varied content such as ebooks [1], music [2], movie [3], or daily tasks such as job-seeking [4], shopping [5]

© Springer Nature Switzerland AG 2021
K. Toeppe et al. (Eds.): iConference 2021, LNCS 12645, pp. 346–363, 2021.
https://doi.org/10.1007/978-3-030-71292-1_27
etc., all have their own recommendation strategies to improve the user experience. This study addresses both online and offline sources of reading recommendations.

Reading is not simply a solitary activity. Rather, through reading together [6], sharing reading experiences, and gathering suggestions as to what to read, reading is a profoundly social activity [7, 8]. Readers may also get useful information outside the reading material to enrich and assist their reading experience, either through traditional reading guides, finding blogs and social media posts, or digesting newspaper reviews [7, 9]. In this varied context, readers can obtain recommendations through multiple channels that may inspire or motivate them to read.

In the digital domain, some mainstream platforms such as Amazon.com provide automated, algorithmic recommendations, and collect large scale reader data. Human-generated material such as readers’ reviews are also common among the digital reading platforms (such as Amazon-Kindle), and review-focussed websites (such as Goodreads). However, this all sits alongside a reader’s own personal networks and social media connections.

This rich information environment is not yet well understood. Research on reading recommendations near-exclusively addresses development of algorithms to provide more accurate recommendations (such as [10]). Limited research, qualitative or quantitative, has been taken to gain an empirical understanding of how readers seek and obtain recommendations from today’s diverse channels. Most of our understanding of users’ recommendation sources and preferences dates back to an era where print was the dominant medium, or predate the widespread adoption of social media [9, 11, 12]. How well those earlier patterns translate into the digital era, or even what the digital era’s preferences are, is not clear. Therefore, we consider the following questions: Do readers rely more on summative and algorithmically generated information, or human-generated prose? Are these preferences consistent for different types of material? While some limited information around specific types of reading is available, overall, there is a major gap in our understanding of what material users rely on when looking online for reading recommendations.

2 Literature Review

The main purpose of reading recommendations is to help a reader in discovering useful or interesting reading material effectively and efficiently. From the reader’s perspective, based on the different channels that they obtain the recommendation, we classify the reading recommendations into two categories: Algorithm-based and Social-sources.

Algorithm-based recommendations are automatically generated: for example, by analysing data from the individual user and the reading material, and matching it with the activity of many other users. In contrast, readers obtain social recommendations through direct or indirect social interactions with others.
2.1 Algorithm-Based Recommendations

Recommender algorithms have long been a popular research topic in information retrieval. Digital reading platforms use these techniques to provide readers with recommendations across diverse genres of reading material, including general books [1, 13], news [14, 15], and academic papers [16, 17].

Profiles and data about both readers and reading material are considered by typical algorithms [18]. There are three kinds of metadata that algorithms frequently use: the reader’s profile, the information about or contained within the reading material, and reader interaction behaviour [19–21].

The reader’s profile data includes their gender, age, location etc. [18]. These data are usually used in demographic-based filtering algorithms [22] to provide recommendations across a specific group (e.g. a book recommendation to children).

The information about reading material can be keywords (e.g. title), genre or the main reading material itself [13, 21]. Reader interaction behaviours capture elements such as how far into a previous book a reader has read to date, their own rating of a book, purchase records, or average reading time per day etc. [23]. Two major approaches exist: content-based filtering, which emphasizes book content, and collaborative filtering, which emphasizes user profiles and book-rating data.

Content-based filtering uses the main text of books to recommend a new item that is similar to the ones that the user has purchased or read before [13, 24]. This algorithm is widely used in digital news [25] and other text-focussed reading. Its strength is providing relevant recommendations with similar characteristics. Content recommendation can start from even a small amount of data about the user, making it more suitable in the cold start situation than the collaborative filtering algorithms, especially when the user is new to the system [26]. However, this algorithm cannot well perform when there is a lack of content data, and recommendations can be of poor quality as many factors are not captured by book content alone [23].

Compared to content-based filtering, collaborative filtering is better at providing high-quality recommendations [23]. It predicts the user’s preference based on their similarity with user other users, and predicts their preferences based on others with a record of similar choices. Thus, it has been compared to a computer-driven “word-of-mouth” recommendation process [27, 28]. A typical example of collaborative filtering is from Amazon.com that “(other) people who buy this item also buy this item” [29]. Although collaborative filtering can generate new suggestions for readers, this approach has some major drawbacks. One such drawback is that it cannot perform well in a cold start situation because it relies strongly on having a large amount of behavioural data about users [20, 23].

To avoid the shortcomings and adopt the good points of different algorithms, many recommendation systems are a hybrid of two or more basic algorithms together to improve the accuracy and quality of recommendations [20], such as [22, 30]. Recent algorithms also consider contextual features such as the user’s recent location [18, 31], to provide more personalized recommendations which may better meet the user’s preference.

Overall, recommender algorithms collect data from the reader, the reading material or other context features. They then use this data to predict and generate personalized
Reading recommendations. This technique has been widely adopted, and two different approaches meet different user needs. Different domains have adopted techniques that rely more on one or other of the two approaches, so variations in the role of algorithmic recommendations can be expected in the context of different reading genres.

2.2 Socially Sourced Recommendations

Reading is a social activity rather than a solitary work [7]. During the reading process, readers directly communicate with other people or indirectly refer to some assistant materials to obtain reading recommendations.

“Reading advisory” is a service typically provided by libraries or bookshops to provide suggested reading options for readers [32, 33]. Besides reading suggestions from a professional librarian, readers also seek recommendations from other people in their social circle. Previous studies have observed that children will seek advice from their parents, teachers or friends [11, 34]. Adults also rely on their social networks [9].

In the meanwhile, traditional mass media such as the review of a book from a TV show and radio can also provide reading ideas for the readers [9], and readers usually treat these as a “trustworthy choice” [35]. Additionally, Ooi [9] observed that some movies which are adapted from a novel can inspire the reader to read the original book. This is an implicit but effective recommendation for the reader [36]. Recommendations from sources unknown to the reader, though, are often disregarded as irrelevant or not authoritative, such as the reader in [37] who commented “oh, there’s a curator[of a section]. But I don’t know who she is, so don’t really care. But if it [the curator] was an author I loved, I might...”.

In the digital reading era, the Internet and diverse social media platforms expand the routes through which the reader can obtain reading recommendations, reducing the limitations of time and location [8]. Many digital reading platforms, such as Kindle, integrate social functions such as sharing, add comments and reviews [38]. These functions make it easier for readers to communicate and exchange ideas with each other. There are also many book review platforms such as Goodreads which provide a specific place for readers to discuss books and seeking for reading recommendations [39]. Social networking sites also benefit readers’ reading journeys, e.g. readers can communicate with each other through instant messaging platforms [7].

All these examples reflect that nowadays, readers have a richer variety of sources for social reading recommendations.

2.3 Research Gaps and Research Aim

Reading recommendations are an established research topic in information retrieval. Most of the previous work examines particular recommender algorithms used in a specific reading context, such as [40] for news recommendations. There is a strong evidence base to understand how effective different algorithms are compared to each other. The vast majority of these comparative studies of algorithms have used either automatic analysis against gold-standard outcomes or, more often, subjective user feedback in controlled laboratory tests [41, 42]. How well the recommendations of these algorithms are accepted outside the laboratory, compared to other sources of recommendations is
less clear. The prevalence of algorithmic recommendations on websites such as Amazon suggests they must be effective to some degree at least, but to what degree do they complement, replace or pay second fiddle to established social sources of recommendations, such as a user’s friends or colleagues? There is a significant lack of qualitative and quantitative data to summarize the whole landscape of the reading recommendation in today’s society.

As for social reading recommendations, previous research has revealed some key behaviours that appear when the reader, directly or indirectly, seeks recommendations. However, most of these studies have happened in the context of physical libraries, bookstores, or printed reading material \([6, 24]\). The sources used by a typical reader when seeking reading recommendations in the digital era are still not clear.

The internet has substantially changed the reading recommendation landscape. Where in the past, book reviews were predominantly available from mass media, \([24]\), in the digital era, readers can easily get access to these reviews through websites or book review platforms such as Goodreads \([23]\). To our knowledge, there is only one previous paper, from 2001, that compares user assessments of recommendations from friends versus online systems \([43]\). However, this paper was 20 years ago, and the adoption of digital reading as a whole, the sophistication and prevalence of recommendation systems, and the emergence of large-scale social media platforms have all signified and driven major changes in the reading environment. Thus, one of our goals in this study is to investigate whether readers prefer algorithm-based or social-sourced recommendation in today’s reading environment.

Previous research has indicated that genre has many effects on the reading process, including recommendations. In the case of book reading, several studies have reported that readers digitally or physically establish a social relationship with each other \([11]\) and get access to the diverse social media (such as TV shows) to obtain reading recommendations \([9]\). Researchers of news recommendation techniques have highlighted the importance of contextual features like time (up-to-date News) and location (where the news happened) as key features, beyond the features such as the topic or keywords that the normal book recommendation would consider. To understand how much user characteristics, as from book reading, or algorithmic characteristics, in news recommendations, either interact or generalize, in this study we look at more than one genre of recommendations. We investigate both fictional and factual books \([1]\), academic resources \([16]\), and news \([40]\) as four typical reading genres which have targeted in the past. We hope to demonstrate whether readers show different preferences for different sources of reading recommendation when considering different genres.

Above all, this empirical study first aims to provide an overview of the relative importance of different sources of reading recommendations, social and algorithmic, play in real users’ choice of different kinds of digital reading material. We focus on digital reading specifically, and do not consider the choice of printed reading material. In consideration of social factors, we include recommendations from other people, whether those recommendations are communicated online or in the real world.
3 Method

Our research process was divided into two phases: first, a diary study and second a survey. Given the lack of previous research, we wished to first gain an overall picture of the variety of sources from which readers obtain recommendations. We started with a diary study to capture real readers’ selection of recommendations over three weeks. This allowed us to gather information without the intrusions into privacy that are associated with systematic logging. Having built a picture of the range of recommendations in use, we then undertook a large-scale survey to both validate the types that we had identified in the diary study, and also to obtain more robust quantitative information about their relative role at a larger scale [44]. Both parts of our study were approved by our institutional human ethics board.

3.1 Diary Study

Our first goal was to capture as complete a range of types of reading recommendations as we could. This could have been done using interviews, but these require participants to recall their memory of recommendations in the past. Interviews are good at eliciting information where the context is unclear, as the dialogue between interviewer and interviewee allows for disambiguation. A diary study, on the other hand, is a longitudinal method which can follow participants with different reading tastes and get themes to record their reading experiences much more closely to when they occur [45].

For the diary study, we randomly recruited our participants through posts shared via several social network applications (including WeChat, QQ, etc.), and through snowball sampling starting with participants from previous studies. The posts included the research background, an introduction on recording a diary, and the rate of payment for participation. Fourteen participants from China (eight males and six females, aged from 25 to 45) joined the study. All participants are frequent digital readers, and they represented a range of different reading tastes and habits, and their occupation of the participants are varied, including students, engineers, and designers. Thus, we could collect a rich and diverse set of data. The reading material included digital books (including both fictional and factual books), digital news, blog articles, and academic papers. We use the anonymized identifiers P1 to P14 to identify the participants in our results.

The diary study lasted from 5 to 29th May 2020. Each individual participant spent about two and a half weeks logging their diaries. We asked the participants to record the platform they read and how they obtained recommendations (from whom and through which channel). We collected a total of 176 diaries with an average of nearly 13 diary entries from each participant. After the diary stage, each participant gave an exit interview (about one hour) to investigate more details about their reading journey and their opinion about the different types of reading recommendations.

The diary study was conducted through WeChat. WeChat is the most popular instant messaging platforms in China, which all our participants used. WeChat supports image, video, audio and even document and web page sending. This allowed participants to flexibly record their diary in various formats, such as taking a screenshot of the reading application interface and sending that to the researcher, or directly sharing a post or a chapter of an ebook etc. Additionally, the video or audio calling functions also make it
easier for the participants and researchers to communicate with each other, especially for interviews. Researcher also can effectively monitor the whole logging process and responded to the participants requirements or questions.

After coordinating the diaries and interview transcripts, we utilized the open coding method [46] and thoroughly inspected and coded the reading material type, the reading platform, and the behaviours which are related to reading recommendation.

Two key findings emerged from the diary study. Firstly, we observed that the way that readers interact with the fictional books, factual books, academic papers, and digital news and articles are different, and so are the routes through which they obtain reading recommendations for the four genres. Secondly, we collated three typical algorithm-based reading recommendations and seven social-sourced recommendations (Table 1) which together covered the types of reading recommendation information reported in the diary study.

Based on this framework, we set up a survey to investigate the reading recommendation of among fictional books, factual books, academic papers, and digital news and articles in depth. This second step would help us determine the proportion of what kinds of recommendations that readers use [44].

### 3.2 Survey

The survey contained two parts: in the first part, we investigated the background of the participants, including their gender, age, occupation and the reading material types that they usually read (a multiple-choice question which contains the four kinds of reading material mentioned above). Based on the respondent’s selection of the types of reading material they read, in the second part of the questionnaire, the answered questions specific to the selected reading material types. The second set of questions contained details of specific reading platforms and what kinds of recommendation they draw on for the genre, but this is not reported in detail.

The options in each question were generated from the framework created from the diary study results. We listed all the platforms and activities from the diaries related to reading recommendations so that the participants can easily recognize and select the types of recommendation that they regularly use. According to the diary study results, reader’s recommendation preferences across different kinds of reading material are significantly different. The “platforms” and “recommendations” listed under each genre of reading material were also different, reproducing the unique sources used for different types of reading material, according to the diary study. For example, TV series and movies were specifically only used as a source of recommendation in the case of fictional reading. It was never seen in the context of factual books, academic papers and digital news and articles. Even the same kind of recommendation may also have a different form of expression in each genre or platform, so we provided a short example for each option to help the participant understand the meaning of it. We would introduce the details of the diverse recommendation sources in Sect. 4.1.

Additionally, we provide an “other” option in each “reading platforms” and “reading recommendation” questions so that the participants can supply any other source not already listed in the question’s provided answers. The survey produced no new reading platforms and no further types of recommendation. This suggests that our diary study
Hey Alexa, What Should I Read? 353

results did accurately capture the reading recommendation landscape among different genres of reading material, and the differences we noted between genre are also consistent with larger-scale behaviour among the Chinese community.

We collected 312 (132 male and 180 female, aged from 18 to 55) responses to the survey in three days. All the data was collected from China, and the survey participants’ background is consistent with those from the diary study.

The data were analysed using 2-way chi-squared test, to identify any variations in how often the different types of recommendations were used in deciding what to read for different genres. The details of this analysis are given in the results.

4 Findings

In our results we first provide an overview of the types of sources readers use for reading recommendations in the context of different genres, as reported in the diary study. In the second part of the results we then address the quantitative findings from our survey. Throughout the analysis, we treat any item as a source of recommendation information if our participants reported it as such. Some features could be argued as not being purpose-built as sources of recommendation data, but we exclusively endorse the reported experiences of our participants.

4.1 Reading Recommendation from Multiple Channels

After analysing the data from the dairies, we identified ten common types of source of reading recommendations. We grouped these recommendation sources into two major categories: algorithm-based and social-sourced. Through this section, we will present example screen snapshots to illustrate the types of information our participants reported. These illustrative examples are from English-language interfaces to make them accessible to the non-Chinese reader.

Algorithm-Based Recommendations. There were three typical types of algorithm-based data used as sources of recommendation by our users. First is the order in which reading material is presented by default, particularly articles on a news site, search results for a query, or browsing a booklist. In this case, the system automatically sorts the reading material in a specific order (e.g. Fig. 1). This could be, for example, the documents that best match the search keywords, or in the order of popularity in the context of news, this was often the recency of the article. Items appearing at the top of the list would be treated as “high quality” or a better match to the user’s interest. P1 recorded that “after I search a keyword [of a factual book], there would have plenty of similar books in the searching result. I would inspect the books in the first few pages and would not go down too much. Because these are more likely to well match my reading requirement and usually have higher quality than those appears below. If I didn’t find what I want within this range, I would consider whether the keywords I search is correct and re-search.”

Secondly, the platforms would also automatically recommend related reading material to the reader (e.g. Fig. 2). These recommendations are usually generated from the reader’s reading history or reading taste [18]. Typically, they appear on the home page of
the platform, the book or article description page or the bottom of a blog: e.g. “[after reading this] you may also like to read”, “similar books/news/articles”. It could also simply be the most popular books is a specific group, e.g. the bestselling Masters of Business Administration textbooks. This example relates to the third kind of algorithm-based recommendation sources: the rating or scoring of reading material.

While the explicit recommendations consider multiple features together and provide comprehensive advice, the rating and scoring of the reading material is much simpler. It uses a specific number to measure the quality of the reading material. Usually, readers can give a grade to the books, and the system would calculate these grades to see which book have a higher score. This function is common in both digital reading platforms (such as Kindle) and book review platforms (such as GoodReads) [7, 39] (e.g. Fig. 3). Readers would also refer to other features, as mentioned above, such as the best seller. For the academic paper, a commonly used score could be number of citations. However, there were no clear examples of this sort of information being perceived as recommendation material in the context of blog articles or digital news. Thus, this third type is only used in three of our four genres.

Social Recommendations. We found seven types of data that users perceived as recommendation that have a social source. The first three recommendations are directly from other users or indirectly from the content generated by others. To protect the privacy of third parties, we provide no screen examples in this section.

Online comments on a book or an article can inspire a new reading journey of the reader. P6 told us that “… one of the comments on this book [a science fiction] showed that this author has another popular book which is similar to this one. I like this book
very much, so I am also eager to read the book mentioned in these comments.” This has also been seen in the news and article reading: when P13 was reading an article about the history of the camera, one comment on that article mentioned a specific type
Polaroid; this inspired P13 to search for more information about this Polaroid and start a new reading journey. In academic reading, the comments play a different role. Academic databases or digital libraries don’t provide a comment function as with standard digital book platforms such as Kindle. However, within academic papers, the authors would cite the previous work and add present comments on it. P1 wrote in the diary “I found plenty of relevant papers through the literature review section of this paper. The authors of this paper provide a comprehensive summary of the previous work and their comments are also quite helpful for me to have a deep understanding of the knowledge in this area.”

Additionally, readers may also seek reading recommendation from other readers’ sharing, or a discussion group. The discussion groups could be a normal reading club. It could also be a group focussed on some other shared interest, e.g. a teamwork group or a group with family members or friends. Sometimes readers obtained recommendations through face-to-face communication, e.g. P11 said: “today my boss recommended me to read this [C++] book in our meeting.” It could also happen through multiple online social network platforms such instant messaging where people may share a link to a book or article. P13 said, “My daughter found news about COVID-19 in our city, and she shared this with me.” And P12 noted, “we have a WeChat discussion group in our department, and we usually share some reading material [pdf document or link to a website] which is related with our work in this group and encourage the employees to read and learn.”

The next three social recommendations are from multiple mass media. Like a previous study [9] we found that readers get reading inspiration from movies or TV series that are adapted from a novel. It was a prevalent type of recommendation that 9 out of 12 participants who read fictional books reported. Additionally, some user-generated videos may also provide reading recommendations for the reader. These UGC videos usually can be divided into two categories. One is book reviews which can be found in both fictional and factual reading. The others are common in fictional reading where publishers of novels adapt the stories into a few short videos and publish these on video-sharing social networks such as TikTok to attract readers. P2 and P9 both mentioned these UGC videos from TikTok several times in their diary.

Furthermore, blog posts are also a critical way for readers to expand their reading channel and obtain more useful reading advice. The forms of these blog articles also vary from different kinds of reading material. For fictional reading, these blogs usually contain a few chapters of the novel for sample reading, or it could be fan fiction stories which are generated by the readers and published in a social network platform such as a fan forum. For factual books and academic papers, these blog articles are usually a piece of knowledge that summarises works from third parties and published in a knowledge-sharing forum. In the case of news and other digital articles, these would be a link within the article that points the user to another related article.

Finally, a task in daily life could also be a motivation for the readers to read. It could be a small reading goal such as “to finish reading five books this month” (P8), or seeking relevant handbooks or academic papers to assist in work or study, or to solve a personal task. These real-life tasks help the reader set a reading goal and seek reading material to meet it.
Above all, readers relied different kinds of reading recommendations when they read different genres of reading material. Even the same type of recommendation may also have diverse forms among various genres of reading material.

Both Algorithmic and Social recommendations are a critical component of readers’ reading journeys. To identify which recommendation play a more important role, we conducted a survey and invited more readers to share their experiences with us.

### 4.2 Survey

The summative data from the survey is seen in Table 1 below.

| The source of recommendation | Fictional reading | Factual reading | Academic reading | News and articles |
|------------------------------|------------------|-----------------|------------------|-------------------|
| Algorithm-based              |                  |                 |                  |                   |
| Initial order of the reading resources | 35.06% (61) | 26.81% (37) | 43.86% (50) | 48.67% (110) |
| Automatically recommended by the platform | 25.86% (45) | 21.74% (30) | 21.05% (24) | 59.29% (134) |
| Aggregated rating or score | 54.6% (95) | 41.3% (57) | 54.39% (62) | – |
| Social-sourced               |                  |                 |                  |                   |
| Comments and reviews         | 45.98% (80) | 43.48% (60) | 44.74% (51) | 39.82% (90) |
| Personal Sharing and communication | 28.16% (49) | 39.13% (54) | 37.72% (43) | 34.07% (77) |
| Discussion groups            | 14.37% (25) | 26.09% (36) | 28.95% (33) | 15.04% (34) |
| TV series or movie           | 36.78% (64) | – | – | – |
| User-Generated Video         | 18.97% (33) | 10.14% (14) | – | – |
| Blogs and articles           | 38.51% (67) | 32.61% (45) | 28.07% (32) | 34.51% (78) |
| Task-oriented motivation     | 14.37% (25) | 49.28% (68) | 53.51% (61) | 31.42% (71) |
| Total                        | 312 participants | 174 | 138 | 114 | 226 |

As explained in the Method section above, the survey questions followed the patterns of recommendation found during the diary study. The survey data, therefore, allows us a
more reliable quantitative analysis of the importance of different elements than the diary study alone.

We first tested for an overall pattern in the relative frequency of algorithm-based versus social-sourced recommendations. In total, algorithmic sources were used 705 times (37.3% of cases), and social sources 1187 times (62.7%). Testing the totals for the two major recommendation types versus the four genres yielded a significant result ($p = 0.017$, $\chi^2 = 10.2$, df = 3). This demonstrates that the relative importance of the two different types does vary between genres. Social sources are particularly frequent in the context of factual reading, and algorithmic sources with news and article reading.

When we examine differences within the two types by genre, further marked patterns of difference appear. In the case of algorithmic recommendation information, the absence of scoring or ranking data for news articles means we cannot do a single test to check for variances. Instead we did two tests. First, we tested all three types of recommendation against the three genres for which the data was available, (fiction, factual, and academic); the result was was not significant ($p = 0.59$, $\chi^2 = 2.80$, df = 4). Therefore, these three factors are used similarly for all three genres. Second, we tested the two types of recommendation that were available for all four genres, which proved significant ($p = 0.0038$, $\chi^2 = 13.4$, df = 3). In the news and article genre users markedly relied on the default ordering of the platform. This combination of results indicates that for fictional, factual and academic reading the sources of algorithmic recommendation are broadly similar, but that news reading is influenced in markedly different ways than other genres. The absence of score information seems to lead to a greater emphasis on recommendations by the platform and the presentation order of articles.

Turning to social sources of recommendation, we again have complexities that arise from the absence of certain types of recommendation for different genres. In this case, we have five sources of recommendation that are used by all four genres. Testing those together to check if genre influences which information is relied on, a chi-squared test proves significant ($p < 0.0001$, $\chi^2 = 43.65$, df = 12). The underlying differences are many. Task-related influences are markedly, and perhaps unsurprisingly, low for fiction reading. In contrast, discussion group recommendations have a high influence for both factual and academic reading; and blog content was particularly low for academic reading. There are broad similarities in the frequency with which particular sources are used between genres. One example is the consistent reporting of the use of online comments. However, blogs and discussion group materials are more strongly associated with particular genres. A local test for the influence of UGC video between fiction and factual reading is just significant ($p = 0.046$, $\chi^2 = 3.97$, df = 2).

The role of films and TV programmes is a unique feature associated with recommendations for fictional reading, and so we cannot reliably test for it alone. However, it is the third ranked source for fiction, and this suggests it is worth further exploration.

5 Discussion

In our investigation, we sought to understand the degree to which users rely on algorithmic versus social information as sources of inspiration for what to read. We now reflect on the findings from both diary study and survey to build a clear picture of how Chinese readers adopt information from both sources when deciding what to read.
There is relatively little previous research on this question for any user group. The key previous paper is that of Sinha and Swearingen [43]. Like our own study, this combined quantitative and qualitative approaches, but in their case, there was no longitudinal component, and participants simply rated recommendations provided to them in isolation from each other in a closed study. That study also considered only recommendations from friends chosen by the participants and recommendations from three chosen systems. Our study differs on a number of points: ours is longitudinal, includes a variety of social sources, not just immediate friends, and in the diary study is naturalistic, capturing actual reader activity in real life, rather than in a controlled experiment. We also captured a wide range of qualitative data in the diary study: something not gathered in the numerical ratings given in the Sinha and Swearingen study. While the previous study considered books and movies, we focus only on text, but maintain a wider scope of reading material. Furthermore, the previous study treated each algorithmic recommendation system in isolation, not as part of a wider technical environment. They found that friends were ranked better than any of the three systems they tested did individually. However, there was not a holistic comparison of friends versus the algorithmic recommendations as a whole. A further previous paper addresses sources of academic reading discovery [47], and investigated readers use of a wide range of sources including some algorithmic recommendations (such as Amazon) and some social recommendations (such as reading lists and personal recommendations) alongside other sources. Like Swearingen, though, this study is not longitudinal, and it further focuses specifically on a single genre of reading: academic books. It further does not focus on recommendations, but addresses discovery as a whole [47].

Our study therefore makes a number of novel contributions. We demonstrate that, overall, social recommendations are used as sources more frequently than algorithmic ones, this difference is moderate, with 37% of our recommendation information coming from algorithmic sources, and 63% from social sources. This average, though, is not the same when different genres of reading are considered.

Sinha and Swearingen [43] did not find significant differences between book and movie recommendations, but we find that algorithms are trusted less often when factual information is being sought, and more often in the case of news articles. Therefore, users do rely on different sources when choosing reading material of different types. Furthermore, some sources of recommendation are particularly associated with particular genres, e.g. film and TV seem to strongly influence fiction reading, but were not reported outside of fictional reading. There are reliable differences between genres: while online comments are a consistent influence on future reading choices, blogs, for example, are little used when seeking ideas for factual reading material.

The reading environment has changed a lot in the twenty-years since Sinha and Swearingen [43]. One key example is that in 2001, social media was at best in its infancy. As their data suggested, we can now confirm that in today’s digital reading era, friends are not the only channel that readers when seeking reading recommendations. Our widened view of the social context, beyond immediate friends, demonstrates that readers also rely on socially sourced information from more distant connections (e.g. friends of friends) and even social posts by strangers or acquaintances.
5.1 Limitations and Future Work

In this research, both the diary study and survey’s data are collected from Chinese adults. Reader’s reading habits and social interactions may vary between countries or in different age groups. It would be worthwhile conducting a comparative study in other contexts to see whether our findings can hold under diverse cultural backgrounds. Also, investigating how children and young teenagers seek reading recommendations may reveal current differences between age groups. It is known that children often rely more on parental suggestions for reading [11], and their access to and use of social media may be much lower for a variety of societal reasons [48]. This suggests that the picture for this age group is likely to be very different.

We also observed across the diary study and survey that the sources of recommendation among different types of fiction reading varies. Fictional reading is, it appears from our data, relatively more complex than the other three genres. One diversity is in the commonly read forms, from professionally published books, to user-generated novels, and fanfiction etc. The platforms that the readers drew on and interacted with included digital reading platforms, social networking platforms, and forums. Different platforms will appear in the future, and specific sub-genres of fiction can be studied so that we could provide more accurate suggestions to promote these works.

Additionally, a similar study could be conducted on the sources of recommendations for other media, such as music or film. A similar approach to our own, commencing with a diary study, may reveal different sources particular to those media, and addressing the particular associations users have with these different forms.

6 Conclusion

Drawing on a diary study and survey, we have established that readers turn to different sources of information about or recommendations of potential reading material when seeking different genres. When seeking factual material, social factors are more influential, and in the context of reading news articles, automatic algorithm-created data is used more. There are also sources that are strongly associated with only one genre, e.g. film and television provide inspiration for fiction reading in particular, but not for factual, academic or news reading. As a first contemporary study, there is ample room for future work. First, our data came from China, and other countries may demonstrate different patterns; similarly we revealed a high level of variation with recommendation for the reading of fiction, and further studies could examine this complex area in detail.

References

1. Tewari, A.S., Kumar, A., Barman, A.G.: Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining. In: 2014 IEEE International Advance Computing Conference (IACC), pp. 500–503. IEEE (2014)
2. Rosa, R.L., Rodriguez, D.Z., Bressan, G.: Music recommendation system based on user’s sentiments extracted from social networks. IEEE Trans. Consum. Electron. 61, 359–367 (2015)
3. Subramaniyaswamy, V., Logesh, R., Chandrashekhhar, M., Challa, A., Vijayakumar, V.: A personalised movie recommendation system based on collaborative filtering. Int. J. High Perform. Comput. Netw. 10, 54–63 (2017)

4. Kenthapadi, K., Le, B., Venkataraman, G.: Personalized job recommendation system at linkedin: practical challenges and lessons learned. In: Proceedings of the Eleventh ACM Conference on Recommender Systems, pp. 346–347. Association for Computing Machinery (2017)

5. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. Expert Syst. Appl. 39, 10059–10072 (2012)

6. Pearson, J., Owen, T., Thimbleby, H., Buchanan, G.R.: Co-reading: investigating collaborative group reading. In: The 12th ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 325–334 (2012)

7. Cordón-García, J.-A., Alonso-Arévalo, J., Gómez-Díaz, R., Linder, D.: Social Reading: Platforms, Applications Clouds and Tags. Elsevier, Amsterdam (2013)

8. Kutzner, K., Petzold, K., Knackstedt, R.: Characterising social reading platforms—a taxonomy-based approach to structure the field. In: 14th International Conference on Wirtschaftsinformatik (2019)

9. Ooi, K.: How adult fiction readers select fiction books in public libraries: a study of information-seeking in context. School of Information Management. Victoria University of Wellington (2008)

10. Huang, Z., Li, T., Xiao, S.: Research on library recommendation reading service system based on adaptive algorithm. Wirel. Pers. Commun. 102, 1963–1977 (2018)

11. Cunningham, S.J.: How children find books for leisure reading: implications for the digital library. In: 11th Annual International ACM/IEEE Joint Conference on Digital Libraries, pp. 431–432. Association for Computing Machinery (2011)

12. Cunningham, S.J., Vanderschantz, N., Timpany, C., Hinze, A., Buchanan, G.: Social information behaviour in bookshops: implications for digital libraries. In: Aalberg, T., Papatheodorou, C., Dobrev, M., Tsakonas, G., Farrugia, C.J. (eds.) TPDL 2013. LNCS, vol. 8092, pp. 84–95. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40501-3_9

13. Mathew, P., Kuriakose, B., Hegde, V.: Book recommendation system through content based and collaborative filtering method. In: 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), pp. 47–52. IEEE (2016)

14. Liu, J., Dolan, P., Pedersen, E.R.: Personalized news recommendation based on click behavior. In: Proceedings of the 15th International Conference on Intelligent User Interfaces, pp. 31–40. Association for Computing Machinery (2010)

15. Li, L., Wang, D., Li, T., Knox, D., Padmanabhan, B.: SCENE: a scalable two-stage personalized news recommendation system. In: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 125–134. Association for Computing Machinery (2011)

16. Bulut, B., Kaya, B., Alhajj, R., Kaya, M.: A paper recommendation system based on user’s research interests. In: 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 911–915. IEEE (2018)

17. Zulkarnain, T.D.P.: Proposed model of academic reading material recommendation system. In: The 3rd Asia Pacific Conference on Research in Industrial and Systems Engineering 2020, pp. 105–109. Association for Computing Machinery (2020)

18. Seyeddenezhad, S., Cozart, K.N., Bowllan, J.A., Smith, A.O.: A review on recommendation systems: context-aware to social-based. arXiv preprint arXiv:1811.11866 (2018)

19. Amri, C., Bambia, M., Faiz, R.: Behavior-based approach for user interests prediction. In: 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), pp. 541–548. IEEE (2017)
20. Chandak, M., Girase, S., Mukhopadhyay, D.: Introducing hybrid technique for optimization of book recommender system. Procedia Comput. Sci. 45, 23–31 (2015)

21. Hu, J., Shao, H., Tang, J., Zhang, S.: Research on reading subject recommendation of library based on data analysis. J. Phys: Conf. Ser. 1607, 012022 (2020)

22. Kanetkar, S., Nayak, A., Swamy, S., Bhatia, G.: Web-based personalized hybrid book recommendation system. In: 2014 International Conference on Advances in Engineering & Technology Research (ICAETR-2014), pp. 1–5. IEEE (2014)

23. Sachan, A., Richariya, V.: A survey on recommender systems based on collaborative filtering technique. Int. J. Innovations Eng. Technol. (IJIET) 2, 8–14 (2013)

24. Mooney, R.J., Roy, L.: Content-based book recommending using learning for text categorization. In: Proceedings of the Fifth ACM Conference on Digital Libraries, pp. 195–204. Association for Computing Machinery (2000)

25. Kompan, M., Bieliková, M.: Content-based news recommendation. In: Buccafurri, F., Semeraro, G. (eds.) EC-Web 2010. LNBIP, vol. 61, pp. 61–72. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15208-5_6

26. Wang, N., Zhao, H., Zhu, X., Li, N.: The review of recommendation system. In: Xie, Y., Zhang, A., Liu, H., Feng, L. (eds.) GSES 2018. CCIS, vol. 980, pp. 332–342. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-7025-0_34

27. Rafsanjani, A.H.N., Salim, N., Aghdam, A.R., Fard, K.B.: Recommendation systems: a review. Int. J. Comput. Eng. Res. 3, 47–52 (2013)

28. Sharda and, U., Maes, P.: Social information filtering: algorithms for automating “word of mouth”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 210–217. ACM Press/Addison-Wesley Publishing Co. (1995)

29. Smith, B., Linden, G.: Two decades of recommender systems at Amazon.com. IEEE Internet Comput. 21, 12–18 (2017)

30. Gao, Q., Dong, A.: A conditional context-awareness ontology based personalized recommendation approach for e-reading. In: 2017 First IEEE International Conference on Robotic Computing (IRC), pp. 365–370. IEEE (2017)

31. Spiteri, L.G.: Readers’ Advisory Service in the Public Library. American Library Association, Chicago (2005)

32. Mohr, K.A.: Children’s choices for recreational reading: a three-part investigation of selection preferences, rationales, and processes. J. Literacy Res. 38, 81–104 (2006)

33. Rehberg Sedo, D.: Richard and Judy’s Book Club and ‘Canada Reads’: Readers, books and cultural programming in a digital era. Inf. Community Soc. 11, 188–206 (2008)

34. Hall, R.M.: The “Oprahfication” of literacy: reading “Oprah’s book club.” Coll. Engl. 65, 646–667 (2003)

35. Makri, S., Chen, Y.-C., McKay, D., Buchanan, G., Ocepek, M.: Discovering the unfindable: the tension between findability and discoverability in a bookshop designed for serendipity. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) INTERACT 2019. LNCS, vol. 11747, pp. 3–23. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29384-0_1

36. Barnett, T.: Social reading: the Kindle’s social highlighting function and emerging reading practices. Aust. Humanit. Rev. 56, 141–162 (2014)

37. Nakamura, L.: “Words with friends”: socially networked reading on Goodreads. PMLA 128, 238–243 (2013)
40. Lommatzsch, A.: Real-time news recommendation using context-aware ensembles. In: de Rijke, M., et al. (eds.) ECIR 2014. LNCS, vol. 8416, pp. 51–62. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06028-6_5

41. Pu, P., Chen, L., Hu, R.: A user-centric evaluation framework for recommender systems. In: Proceedings of the fifth ACM Conference on Recommender Systems, pp. 157–164. Association for Computing Machinery (2011)

42. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. 22, 5–53 (2004)

43. Sinha, R.R., Swearingen, K.: Comparing recommendations made by online systems and friends. DELOS 106 (2001)

44. Preece, J., Sharp, H., Rogers, Y.: Interaction Design: Beyond Human-Computer Interaction. Wiley, Hoboken (2015)

45. Lazar, J., Feng, J.H., Hochheiser, H.: Research Methods in Human-Computer Interaction. Morgan Kaufmann, Burlington (2017)

46. Flick, U.: An Introduction to Qualitative Research. Sage, Thousand Oaks (2018)

47. Rowlands, I., Nicholas, D.: Understanding information behaviour: how do students and faculty find books? J. Acad. Librariansh. 34, 3–15 (2008)

48. Mascheroni, G., Ólafsson, K.: The mobile Internet: access, use, opportunities and divides among European children. New Media Soc. 18, 1657–1679 (2016)