Research on Book Recommendation System for People with Visual Impairment Based on Fusion of Preference and User Attention

Zhi Yu¹,²,³(✉), Jiajun Bu¹,²,³, Sijie Li¹,²,³, Wei Wang¹,²,³, Lizhen Tang⁴, and Chuanwu Zhao⁵

1 Zhejiang Provincial Key Laboratory of Service Robot, College of Computer Science, Zhejiang University, Hangzhou, China
{yuzhirenzhe, bjj, 21621177, wangwei_eagle}@zju.edu.cn
2 Alibaba-Zhejiang University Joint Institute of Frontier Technologies, Hangzhou, China
3 Ningbo Research Institute, Zhejiang University, Ningbo, China
4 China Braille Press, Beijing, China
eicatang_1@163.com
5 Zhejiang Toman Intelligent Manufacturing Technology Co. LTD., Xinchang, China
zhaocw@zjtoman.com

Abstract. With the development of the Internet, the information explosion problem comes into being and it is challenging for users to search for the information they needed from e-books. Although the book recommendation system can help users find their focuses, it is not applicable for visually impaired users when using ordinary visual reading methods for knowledge acquisition. Therefore, a book recommendation system that suits their behavior habits is required. In order to provide accurate and effective book sets for users, we propose an algorithm based on fusing their preferences. For intelligently ranking the candidate book sets and help users find the right book quickly, we propose a context-aware algorithm based on users’ attention. Meanwhile, we introduce an improved calculation method for users’ attention to solving the problem of inaccurate prediction on users’ current attention when their action history is cluttered. We use the self-attention to preserve the users’ reading tendencies during the reading process, analyze users’ personal features and book content features, and improve the accuracy of the recommendation by merging the feature space. Finally, the improved algorithm proposed and comparative experiments were employed on the dataset collecting from the China Blind Digital Library, and the effectiveness of the improvement is proved in each experimental comparison results.

Keywords: People with visual impairment · Digital book · Recommendation system · User attention
1 Introduction

With the rapid development of the Internet, massive amounts of data can be applied to analysis and research. The problem of how to choose useful and personalized information brings more challenges to the information service system. The recommendation system can effectively solve such problems, so it has rich practical applications in the fields of entertainment, content, e-commerce, and services. As a part of the content recommendation field, book recommendation is also being developed and improved. Many reading platforms have developed their own recommendation systems to serve them.

However, a large part of people still does not benefit from it. According to the statistics of the China Disabled Persons’ Federation, as of the end of 2010, there are approximately 12.63 million people with visual disabilities in China, and this number continues to grow. For visually impaired users, because of their visual abilities, it is difficult to obtain information through normal visual channels, and can only rely on other sensory channels such as touch and hearing. Ordinary reading methods cannot meet their needs for knowledge acquisition. As developers and beneficiaries of the Internet, when enjoying the convenience and knowledge brought by massive data, we should also pay attention to these groups so that the visually impaired can also enjoy the information from the Internet.

Visually impaired people have different behaviors and methods when they browse the Internet: they are more willing to receive voice messages, often browse fixed websites, and often use the favorites function; at the same time, they have many challenges during browsing the webpages. For example, it is difficult to switch webpage windows, browse the entire webpage, and useless information greatly affects the browsing experience. When they want to acquire e-books in a huge amount of internet data, how to find what they are interested in quickly is an essential problem. Even because of visual defects, screening information is more difficult for them than sound people. Book recommendation systems in the market are recommended for the reading habits and interests of healthy people, which is not suitable for visually impaired people. Therefore, we need to set up a Braille book recommendation system for visually impaired people in accordance with their behavior habits.

We elaborate on the development of book recommendation systems and the problems encountered by visually impaired people when reading books, and propose a context-aware recommendation algorithm based on fusion preferences and user attention, which respectively solves the problems encountered in book recall and intelligent ranking. Insufficient recalls and inaccurate rankings, and the algorithm was reasonably integrated into the accessible book recommendation system, and corresponding improvements were made to the characteristics of the visually impaired population. Experimental results on real datasets of the Chinese Braille Club library show that this method has outstanding advantages in terms of recommendation accuracy and intelligent ranking of recommendation lists.
2 Related Work

We give a general introduction to the related technologies used, mainly introduce the development of recommendation systems, and concepts of user attention mechanisms.

In 1990, with the development of computer algorithm development and application deployment, recommendation systems began to appear. In 1992, Goldberg [1] and others first proposed collaborative filtering. Two years later, the first automated recommendation system GroupLens [2] was proposed by the University of Minnesota and applied collaborative filtering technology to the recommendation. In 1997, the Recommendation System was first proposed and began to become a research area. Subsequently, the recommendation system gradually became a popular research direction due to the strong role of consumption guidance in e-commerce. Amazon applied the recommendation system to product recommendation as early as 1998 and launched collaborative filtering based on products. Through continuous improvement and Exploration, and finally reached a GMV contribution rate of 20% to 30% [3], illustrating the significant role of the recommendation system. After more than 20 years of accumulation and precipitation, the recommendation system has gradually formed a complete and comprehensive research system. At the algorithmic level, it is generally believed that there are three basic recommendation methods, namely content-based recommendation [4], user behavior-based recommendation [5], and hybrid recommendation.

The Encoder-Decoder based model is a variant of Recurrent Neural Network(RNN) [6, 7], where the Encoder encodes the input X into a fixed-length hidden vector C, and the Decoder decodes the hidden vector C to the target output Y, where the process from X to C and from C to Y is built with RNN. However, such models lack discrimination for the input sequence ABC, but all become the same hidden vector C. Therefore, in 2015, Kyunghyun Cho et al. [8] introduce the attention mechanism into the model to solve this problem. Attention is a mechanism used to change the calculation method of the hidden vector C, thereby improving the effectiveness of the Encoder-Decoder model.

Self-attention [9] is very different from the traditional attention mechanism. It is performed separately in the calculation of the encoder and decoder. The encoder and decoder each calculate the self-attention related to their input and output. The obtained result is only the dependency between its input and output. Finally, the self-attention of the encoder is added to the self-attention of the decoder, and finally, the attention of the entire model is obtained. The advantage of Self-attention is that it can directly calculate in parallel without resorting to the cyclic characteristics of RNN. It directly uses the attention model to model the sequence, and the model effect is better. Therefore, we chose self-attention as the attention in this article. Baseline algorithm of the force mechanism.
3 Recommendation Method

The recommendation algorithm based on fusion preferences proposed in this paper combines user behavior data and content data. The main idea is to obtain user preference vectors for books by mining user behaviors, user characteristics, and book characteristics, thereby recommending books of interest to users.

In terms of content calculation, due to a large amount of text data (such as abstracts, keywords, categories, tags, etc.) in the book recommendation system, we will focus on the text function, which is used to calculate the user’s interest in books and books. Secondly, in order to retain the user’s personalized preference interest expressed by the content data, we quantify the user’s implicit feedback on the book based on the user’s historical behavior record and behavior type and generate the user’s behavior for the user’s personalized preference expression vector. In order to retain the neighborhood preferences expressed by user behavior, we propose a novel co-occurrence book pair selection method that models users to read all co-occurrence book pairs in a book and extract relevant text from them based on the characteristics. Theme Similarity and text similarity can help a lot to find similar content books, while users reading behaviors on the similar content book are in common to some extent. We will predict multiple candidate book sets to keep the diversity of the results and sparse user preference matrix is used to present the user feature, we combine the user’s preference with the content text’s preference to facilitate user recommendation.

Our proposed context-aware recommendation algorithm based on user attention is divided into three parts: In the first part, we use a recurrent neural network RNN to mine the sequence information read by the user, train the user to read the book’s probability model and then personalize the book for the user recommend. In the second part, we introduce the user self-attention model to further accurately describe the user’s reading behaviors. In the third part, we use the feature space fusion attention mechanism to further simulate the user’s real attention. By constructing different behavioral features in different feature spaces, we retain the part of global attention which is biased towards the feature space to accurately characterize user behavior attention. During the training part, we obtain the probability of each book in the candidate book set and recommend to each specified user with achieving intelligent book ranking and improving the ranking effect through model improvement.

4 Experiment

4.1 Dataset

The data set used in this paper is divided into two parts: reading records and book content information. The reading record collected data from the Chinese Braille Library for 830 days from April 2016 to August 2018. It contains 44375 reading records of 3282 books by 3368 visually impaired users, 7013 bookmark records and 1193 Articles are added to the collection record, which is extremely sparse. Since our records describe the user’s reading, bookmarking, and favorite records, they are a data set consisting of the user’s implicit feedback. Due to the inconvenience of visually
impaired users, user operations are very difficult for them, so there is no explicit scoring mechanism. Each of our reading records contains a user unique identifier, a book unique identifier, a reading start timestamp, and a reading progress bar at the end of the reading. We have four types of display for the form of book contents: e-books, oral images, audiobooks, and e-braille, with a total of 8986 books. For each book, we collected the book’s unique identifier, book display type, book title, author, keywords, book introduction, and book content type to extract the content features of the book.

Figure 1 shows the number of books read by the user (a), distribution of the number of times the book was read (b), and the distribution of books read by visually impaired users (c).

Figure 1 shows the number of books read by the user (a) and the distribution of the number of times each book was read (b). It can be seen that nearly half of the users (49.4%) have read less than 5 books, and only a few users have read them. More than 50 books; most books (60.7%) have not been read more than 50 times, and only 2% of popular books. Figure 1 (c) shows the distribution of reading for various categories of books by visually impaired users. It can be seen that users prefer to listen to audio directly, and have fewer times to listen to oral images. It can be seen that our user behavior is extremely sparse, and most users lack a sufficiently long behavior sequence and enough user ratings, but since our proposed algorithm can perform sequence input without limiting the fixed length, it can effectively adapt to us. At the same time, we use the updated user preference vector to augment user ratings, and can also handle the preference evaluation of sparse data. The behavior distribution of books is more saturated. More than two-thirds of the books have been read more than ten times, which can better calculate the similarity of books.

This article mainly studies book recommendations for this special group of visually impaired people. It is divided into two stages: recalling the Top-N recommendation set and intelligently ranking the recommendation set. During the recall phase, we calculated a recommendation list of length N for a given user’s reading history, so that the recommendation list contained as many books as the user would read. In the sorting stage, for a given user’s reading context information and book collection, we calculate the prediction of whether each book in the collection will be read, so that the possibility of a book with a large prediction probability is read. Therefore, our main evaluation indicators are Mean Average Coverage (MAC) and AUC.
4.2 Experiment on Fusion of Preference

In this section, a comparative experiment will be performed to prove that the recommendation algorithm based on user fusion preferences has a good effect on the recommendation recall of Top-N. In the selection of comparison experiments, since this algorithm calculates the recommended recall set based on the improved similarity and recommendation vectors, as a baseline comparison, we introduce content-based collaborative filtering, and to explore four similarity calculation methods separately. Under the recommendation results generated by the recommendation vector, we also conducted comparative experiments. The specific experimental scheme is as follows:

1. **BaseSimi similarity** calculation based on common co-occurrence pairs
2. **NormalSimi similarity** calculation based on common co-occurrence pairs
3. **StableSimi similarity** calculation based on stable co-occurrence pairs
4. **TextSimi similarity** calculation based on text similarity
5. **FusionSimi similarity** calculation based on fusion similarity.

To generate the top-N recommendation, method 1 ranks similar books according to similarity based on traditional ICF, while methods 2, 3, 4, and 5 rank similar books based on similarity calculated by themselves.

![Mean Average Coverage](image)

**Fig. 2.** MAC comparison between the proposed method and the comparative methods

As can be seen from Fig. 2, the coverage of FusionSimi is significantly higher than the other four solutions. This is because the similarity calculation of the fused features takes into account the user’s strong association, weak association and book content association, especially the book content association. The collection of books it covers is very different from the collection covered by NormalSimi in order to be able to mine as many book associations as user behavior can’t, thereby recalling books that most users might be interested in.
4.3 Experiment on Fusion Attention

In this section, comparative experiments will be performed to prove that context-aware recommendation algorithms based on feature space fusion attention have better performance in the ranking. In the choice of comparison experiments, because the algorithm uses contextual information to predict and improve the user’s personalized book preferences, we use Bayesian personalized ranking algorithm BPR as the comparison experiment. For the ranking algorithm under the force mechanism, we also conducted comparative experiments on the RNN without any attention, with the self-attention (ARNN), or with the fusion-attention (FARNN).

As can be seen from Table 1, the time complexity of BPR is the lowest. This is because BPR is calculated based on the Bayesian model. Compared to neural networks, it naturally saves time. Therefore, the model performance of BPR is worse than all RNN-based solutions. In the three neural networks of RNN, ARNN, and FARNN, it can be seen that FARNN greatly saves the training time of the recurrent neural network, and its AUC is also the highest among the three neural network models.

5 Conclusion

This paper proposes a context-aware recommendation algorithm based on fusion preferences and user attention. This algorithm solves many problems encountered in recommending books to visually impaired people and integrates into an accessible book recommendation system. To providing visually impaired users with accurate and effective books of interest, we propose a recommendation algorithm based on fusing preferences. To predict the preferences of the visually impaired when selecting books, we explore the user’s behavior-based interests and content-based interests and combined these two features to predict the books that users want to read more accurately. To intelligently categorize collections so that users can quickly find books of interest when browsing sequentially, we propose a context-aware recommendation algorithm based on user attention to intelligently categorize collections. The experimental comparison results show the effectiveness and improvements of our method.

Table 1. Time and AUC comparison between the proposed method and the comparative methods

| Method | Time (s) | AUC   |
|--------|---------|-------|
| BPR    | 153.3   | 0.8798|
| RNN    | 4340.5  | 0.9035|
| ARNN   | 4132.2  | 0.9042|
| FARNN  | 377.7   | 0.9060|
The current experiment is an offline experiment based on the Braille database dataset, so the recommendation level cannot be adjusted for real-time feedback from users. Next, we plan to optimize the recommendation method during training by applying an online recommendation algorithm based on user feedback.

Acknowledgments. This work is supported by Alibaba-Zhejiang University Joint Institute of Frontier Technologies, The National Key R&D Program of China (No. 2018YFC2002603, 2018YFB1403202), Zhejiang Provincial Natural Science Foundation of China (No. LZ13F020001), the National Natural Science Foundation of China (No. 61972349, 61173185, 61173186) and the National Key Technology R&D Program of China (No. 2012BAI34B01, 2014BAK15B02).

References

1. Goldberg, D., Nichols, D., Oki, B.M., et al.: Using collaborative filtering to weave an information tapestry. Commun. ACM 35(12), 61–70 (1992)
2. John, R., Mitesh, S., Neophytos, I., et al.: GroupLens: an open architecture for collaborative filtering of netnews. In: ACM Conference on Computer Supported Cooperative Work. ACM (1994)
3. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput. 7(1), 76–80 (2003)
4. Pazzani, M.J., Billsus, D.: Content-Based Recommendation Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web. LNCS, vol. 4321, pp. 325–341. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72079-9_10
5. Sarwar, B.M.: Item-based collaborative filtering recommendation algorithms. In: International Conference on World Wide Web. ACM (2001)
6. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to Sequence Learning with Neural Networks. (2014)
7. Cho, K., Van Merrienboer, B., Gulcehre, C., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. Computer Science (2014)
8. Cho, K., Courville, A., Bengio, Y.: Describing multimedia content using attention-based encoder-decoder networks. IEEE Trans. Multimed. 17(11), 1875–1886 (2015)
9. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. (2017)