The Research on User Interest Mining Based on Deep Learning

Yi Song
Dept of Computer of electronic and Information Engineering Harbin Huade Universit
Harbin, China
songyiyouxiang@163.com

Abstract. This paper, user behavior depth technology is studied, so that user information and
search engine can interact intelligently. The system can get information that is conducive to
mining users'potential search interests and preferences, so as to use this information to guide the
search process and results filtering, and achieve efficient and accurate retrieval. The research
process uses the mining of user logs of large data search engines, the establishment of user query
classification system, and the deep mining of the behavior of user groups with different
characteristics. User social network can be used to create a retrieval system in which users
cooperate with each other because of their similar interests. This paper studies the calculation
and method of user interest similarity to discover user groups with similar interest. User
relationship network and system mining user relationship network are compared to promote the
integration of the two and provide a basis for the application of recommendation system.

1. Introduction
The difficulty of information retrieval comes from the mismatch between user requirement space and
document information space. On the one hand, user query as a description of users'actual needs is too
simple and inaccurate; on the other hand, in the Internet era, the prosperous Internet environment
contains rich information, but also brings a lot of noise and redundant information. In the existing
research, people have done a lot of research on the analysis and modeling of document space. Whether
from the content analysis or link analysis of documents, they have obtained rich results. In contrast, the
analysis and research of user space by researchers is far from enough. The problems of vagueness,
ambiguity and diversity of user queries have become the bottleneck of modern information retrieval. In
this project, through the acquisition and learning of user interest and social network, a large number of
user space analysis is carried out. This research not only has important theoretical significance for
personalized information retrieval, but also has important problems for information filtering and
personalized information evaluation. Especially in personalized information retrieval, most of the
methods currently used are manual evaluation. This method is time-consuming, laborious, and involves
a lot of manual elements, and a lot of work. Through the research of evaluation technology, this project
studies the automatic evaluation method suitable for personalized retrieval system [1-2], which is not
only of great significance to personalized retrieval, but also of great reference value to other related
research.

User interest similarity calculation is the basis of user social network mining. On this basis, user
interest network is constructed and user relationship network is mined. Since it is difficult for a single
relational network to accurately characterize user interests, this paper will study the interaction of
complex relational networks, and finally apply these results to personalized information retrieval based on collaborative filtering.

2. **User Interest Similarity Computing**

User similarity calculation is to calculate the similarity of user's interest through the characteristic information representing user's interest, so as to cluster users, so that users with similar interests can learn from and share their experience, and achieve the purpose of collaborative filtering and resource recommendation [10].

To calculate the similarity of user interests, we can use content feature comparison method based on user information. Firstly, feature information is extracted from user description file to express user's interest, and weighted keyword vector method is used. Then, vector space similarity calculation method is used to find k most similar neighbors for current users to predict current user's interest. There are two ways to determine the neighbor users. One is to select the neighbor users whose correlation is greater than the threshold according to the pre-determined similarity threshold. The other is to select the first N users with the greatest correlation as the neighbor users according to the pre-determined number of neighbors N. This method encounters two difficult problems in practice, one is sparsity, that is to say, in the early stage of system use, because the system resources have not been evaluated enough, it is difficult to use these evaluations to find similar users; the other is scalability, that is to say, with the increase of system users and resources, the performance of this method will become lower and lower [9].

This paper intends to use classification and clustering methods to analyze user's interest similarity. The core of this method is to describe user's similarity by classification and clustering. The user's interest network is created and described by using automatic text classification, network clustering and other technologies within the framework of user's personalized classification system and classification identification. This method of describing user's interest in the form of category information conforms to the thinking habit of expressing user's interest. To a certain extent, it overcomes the shortcomings of computing complexity, data sparsity, semantic bias and poor expansibility which may be caused by describing user's interest similarity by terms and weights. At the same time, it can reveal user's interest similarity relationship flexibly and multi-dimensionally. Quickly and effectively establish links between similar users. Model Architecture “Fig.1”.

![Fig.1 Model Architecture](image)

3. **Construction of User Interest Network**

Complex network structure can describe a variety of complex systems, such as the Internet is described as a complex network linking routers and computers through various physical or wireless connections [3-5] knowledge, ideas and social relations spread in the social network, its nodes are human, edges
represent various social relations; and so on. Similarly, the user interest relationship we studied can also be described by relational network. Nodes are users and edges representing various interest relationships. User relationship network can be automatically mined, discovered and organized by the system. For example, users can access the information of the content of the web page, and think that users who access similar documents have the same interest, that is, there are relations in the interest relationship network. The discovery of this kind of relational network requires users to access historical data of document information or access logs of retrieval system. By calculating the similarity of user interests, for each user, a set of users with similar interests can be found in the user library. We need to fully mine the same interest groups on this kind of user interest information, study the generation algorithm of virtual user interest network based on graph structure or tree structure, and construct a user relationship network topology map that conforms to the normal network distribution to truly characterize the user interest relationship network.

At the same time, it can also establish a trusted user relationship network through the form of system assistance and user participation, so that users can find users with similar interests and establish network relationships in a network virtual community or group. User participation system or user data of similar functional system need to be constructed. User Interest Network “Fig. 2”.

4. Deep Mining of User Relational Networks
Based on the constructed user interest network, complex network analysis and network data mining technology are studied to obtain valuable structured information. Personalized information retrieval system contains a variety of complex interest relationship networks with multi-nodes [6-7]. Analysis and clustering of these relationship networks can help to promote the clustering of objects within the system and the grouping of people; to establish linkages between contents or functions with intrinsic relevance, to improve the combination operation of functions of the system and the relevance retrieval of content; to enable common tasks or objectives; and Interested users can establish a relationship, which helps to promote and improve collaboration and collaboration among users.

5. Conclusion
User interest is not constant for a long time. User short-term interest accumulated to a certain extent may evolve into long-term interest, which requires us to learn and update user interest model. If not updated, user interest drift may occur. In user model updating and learning, we use forgetting factor to update learning for short-term interest, and LRU elimination algorithm for long-term interest, eliminating the category with the least recent interest. Interest Classification Feature “Fig. 3”.

![User Interest Network](image-url)
Fig.3 Interest Classification Feature

Acknowledgments
This work is supported by the Natural Science Foundation of Heilongjiang Province of China (No. F2015046), the National Science Foundation of China (No. 61672185).

References
[1] Learning image and user features for recommendation in social networks. Geng Xue, Zhang Hanwang, Bian Jingwen, et al. IEEE International Conference on Computer Vision. 2015
[2] Connecting social media to e-commerce: cold-start product recommendation using microblogging information. ZHAO W, LI S, HE Y, et al. IEEE Transactions on Knowledge & Data Engineering. 2016
[3] Deep collaborative filtering via marginalized denoising auto-encoder. Li S, Kawale J, Fu Y. Proceedings of the 24th ACM International Conference on Information and Knowledge Management. 2015
[4] Meta-Prod2Vec: product embeddings using side-information for recommendation. Vasile F, Smirnova E, Conneau A. ACM Conference on Recommender Systems. 2016
[5] Improved recurrent neural networks for session-based recommendations. Tan Y K, Xu X, Liu Y. Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. 2016
[6] Tag-aware personalized recommendation using a deep semantic similarity model with negative sampling. Xu Z, Chen C, Lukasiewicz T, et al. Proceedings of the 25th ACM International Conference on Information and Knowledge Management. 2016
[7] Neural collaborative filtering. Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, Tat-Seng Chua. Proceedings of the 26th International Conference on World Wide Web. 2017
[8] ZhongMing Ma, Gautam Pant, Olivia, R. Liu Sheng. Interest-Based Personalized Search. ACM Transactions on Information Systems, New York. 2007. ACM: 1–5
[9] Xujuan Zhou, Sheng-Tang Wu, Yuefeng Li, Yue Xu, Raymond Y.K. Lau, Peter D. Bruza. Utilizing Search Intent in Topic Ontology-based User Profile for Web Mining. Proceedings of the 2006 IEEE International Conference on Web Intelligence. 2006. IEEE: 1–3
[10] Knijnenburg, B.P., Willemsen, M.C., Gantner, Z., Soncu, H., Newell, C.: Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction 22(4–5), 441–504 (2012). DOI 10.1007/s11257-011-9118-430.
[11] Konstan, J.A., Riedl, J.: Recommender systems: from algorithms to user experience. User Modeling and User-Adapted Interaction 22(1–2), 101–123 (2012)