

Personalized recommendation system of UGC (User Generated Content) video resources based on user interest graphs

Tong Wua*, Yuanqing Lib, Yiping Wangc

New Media College, Communication University of Zhejiang, Hangzhou, China

^a*516868946@qq.com, ^b1049480485@qq.com ^c3242782698@qq.com

ABSTRACT. *The research object of this paper is a personalized recommendation system of UGC (User Generated Content) video resource based on user interest graph. This article takes the personalized recommendation system of UGC video resources under the interest graph of the whole network as the research object, and generates and integrates user interest graph, development evolution and feedback of user interest graph, analysis and recognition of video content based on deep learning. The key technology theories and implementation plan of the video recommendation system based on user interest graphs are studied deeply. The goal is to form a perfect user interest graph and UGC video resource personalized recommendation framework, and providing solutions for personalized recommendation of network video resources.*

KEYWORDS: *Interest graph, Deep learning, UGC video resources, Recommendation system*

1. Introduction

The personalized recommendation system based on the interest graph of the whole network discovers the user's interest through the user's historical behavior data, interest algorithm automatically discovers user interest, and then actively recommends the information of interest to the user according to the user's information needs. So solving the contradiction between the large amount of information and the difficulty of information selection. Recommendation system by tracking user historical behavior, keep updated and iterate user interest, make recommended information always the same as the user's points of interest. Let users get their favorite information more conveniently. The ultimate goal is to achieve all-round personalized customized push for users. For example, with the development of social networks, big data and cloud computing technologies, network users can publish valuable comment information through more platforms (including shopping websites, blogs, forums, etc.). The opinions and emotions

expressed by these reviews will become an important reference basis for recommendation on other platforms. The user's various behavior information in the network environment will reflect its personalized value, and correct analysis and effective response to it will be of great significance to the development of the platform recommendation system. The recommendation theory based on the interest graph of the whole network brings new hope to the breakthrough of the recommendation system. It can be based on the subscriptions (such as on Weibo) or purchased goods (such as on Taobao) or ratings (such as on Douban) or running searches (such as on Baidu) or Comment on the problem (such as Zhihu) integrated to generate interest graphs, to make more effective personalized recommendations for users and make up for the shortcomings of traditional recommendation algorithms. A good recommendation system can not only bring users an excellent experience, but also bring significant economic benefits. Therefore, the research and development of a personalized recommendation system based on the interest graph of the whole network is very necessary.

2. Related work

The research on the recommendation system introduced in this article involves some fields like interest graphs and deep learning, etc. And it mainly includes two parts: interest graph construction and video content analysis. The following introduces the relevant work of these two aspects:

2.1. Construction of interest graph

The interest graph is a network graph formed by clues of common interests among people. It has two layers of connotation. First, the interest graph can acquire and collect the interests of users, and connect it to discover the common characteristics of users. Second, the interest graph can mark the user's characteristics and deverse interests. Once the concept of interest graph was proposed, it was widely concerned by academia, and doing a lot of research about the interest graph. For example, Lynne[1] proposed the construction method of interest and it was applied in the data collection of social networking sites and the selection of user interest. This way achieved user interest collection and interest integration. Berkovsky[2] proposed the user interest model integrated arbitration framework. In the case where some source files are missing, it still can achieve a good recommendation effect. Some scholars use the vector space model and the complex network analysis tool Gephi to construct the interest graph based on social e-commerce users, and it achieves ideal results ([3]);

2.2. Video content analysis

In the field of video content analysis, video content analysis technology with artificial intelligence technology as the core is the focus of scholars at home and abroad. Many scholars have researched video prediction, image retrieval, and video

content recognition technology based on deep learning algorithms, which achieves intelligent recognition analysis of video content and indicates the direction for video content analysis ([4]-[6]). Although domestic and foreign scholars have rich research results in this field at this stage, how to use interest graphs to achieve personalized recommendation of online video resources has always been a problem that puzzles the academic community because of the lack of key technologies in the recommendation field based on interest graphs at this stage and it needs further research and discussion ([7]). Therefore, it is necessary to study the personalized recommendation system of online video resources based on the interest graph.

3. System architecture design

3.1. Overall framework

The video resource intelligent recommendation system is mainly composed of these four main modules.

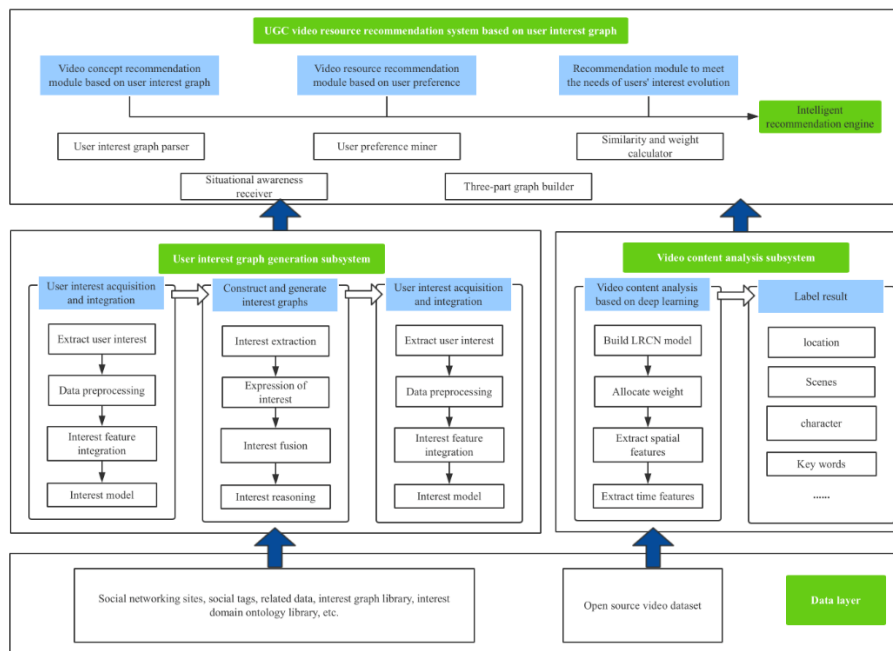


Figure 1 Overall framework

3.2. UGC video resource intelligent recommendation system based on user interest graphs

This subsystem provides user interest graph analyzer, similarity and weight calculator, three-part graph builder, user preference miner, situational awareness receiver, personalized recommendation engine and 3 recommendation modules. The user interest graph parser is used to parse the user interest graph according to the established user interest graph library; the similarity and weight calculator is used to calculate the user similarity based on interest, and the importance weight of user, interest and video resources; the three-part graph builder is used to construct the three-part graph among user, interest and video concept; the user preference miner is used to mine the user's preference for video attributes and provides basic data for video resource recommendation; The situational awareness receiver is used to sense the current situation of the user and filter the recommendation results; the three modules include a video concept recommendation module based on user interest graphs, a video resource recommendation module based on user preferences and a recommendation module that meets user interest evolution needs. Based on the source data, the intelligent recommendation engine adopts various processing technologies and recommendation module functions to recommend video resources based on its interest graph for users.

3.3. Research on the generation of interest graphs of users' entire network and the dynamic evolution of interest graphs

First of all, adopting the bottom-up interest graph construction method. Then, selecting one with higher confidence to join interest library by extracting entities from some open-link data. And finally constructing the top-level ontology mode to complete the construction of interest graph. In the process of constructing interest graphs, automatic or semi-automatic means is used to extract interest from the original database and the third-party database from the original data such as structured, semi-structured and unstructured data and store it in the database and mode layer of interest library. At last, I can use interest extraction, interest representation, interest fusion and interest reasoning to build an interest graph. The following introduces these technologies in detail. First, interest acquisition and presentation technology, that is, extracting the text from the user's unstructured data. After extracting the text, it uses natural language technology to identify the entity in the article. While entity recognition uses named entity recognition technology to identify the entity in the article. In the process of processing unstructured data, you can use the wrapper to learn the extraction rules of semi-structured data. When the user's data is stored in the database, use ETL tools to reorganize and clean the data to obtain the final data; Second, interest Fusion technology, that is, a fusion of knowledge extracted from multiple data sources. In the process of interest fusion, the ontology matching algorithm of pattern matching is used for knowledge fusion, and the dynamic merging method is established based on the risk minimization of Bayesian decision to obtain the ideal matching result; Third, interest reasoning technology, that is, using entity relationship learning methods to learn the relationship between instances and instances in the knowledge graph, and interpreting triplets based on the potential features of the instances.

The interest graph dynamic evolution technology is composed of three parts: interest dynamic evolution, interest adoption, and interest graph evolution. The interest graph evolution technology achieves the goal of user interest graph dynamic evolution through technologies based on time series prediction and graph structure link prediction. The interest graph adoption technology uses the construction of stereotyped models for research. The interest graph evolution technology is implemented through a feedback mechanism based on user interest adoption and video browsing behavior.

3.4. Research on online video content recognition based on deep learning

Based on deep learning technology, the LRCN model proposed by Jeffrey Donahue is used to analyze network video resources. The deep learning model simulates the attention characteristics of the human brain, and uses selective attention weights to assign different weights according to the correlation size of the video theme-related areas to extract distinguishable spatial and temporal features. Based on the time characteristics of the video, the BLSTM network is used to extract the forward and reverse time characteristics, relying on the time information of the video to identify the video content, realize the accurate recognition and analysis of the network video content, and obtain the video label result, to recommend to the user Video resources of interest characteristics and personal preferences.

3.5. Data collection and processing

The system obtains relevant data sets based on API interfaces provided by sites such as social networking sites, social tag sites, UGC video platforms, and associated data cloud LDB, and extracts data that can represent user interests (such as basic information, video resources, hobbies, tags, Attributes, situational awareness, etc.). Generate local interest graphs according to different data sources, and use integration technology to realize the integration of global interest graphs to form each user's personalized interest graph. Use the current open-source video data set for training and testing of video content analysis models based on deep learning.

4. Technical solutions

4.1. Basic ideas

The article is based on the idea of Requirement analysis---Research on key theory and technology---experimental verification to study the recommendation system. Firstly, the application background and requirements of the UGC video resource personalized recommendation system based on user interest graph is analyzed. then, key technologies are studied, such as the generation and integration as well as dynamic evolution and feedback of user interest graph, video content

analysis and recognition based on deep learning, and video recommendation systems based on user interest graph. finally, the framework and technical solutions of UGC video resources personalized recommendation system based on interest graph are built. Thus, existing data sets to simulation verification are used, and then we can analyze and demonstrate the results of the project research.

4.2. Technical route

The research route of this project is as follows:

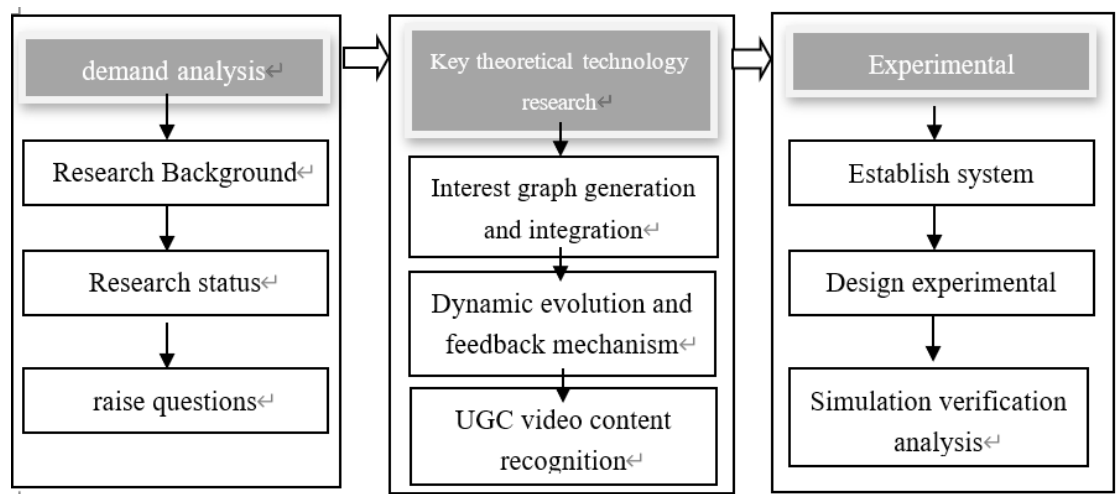


Figure 2 Research route

4.3. Research methods

4.3.1. Literature research method

We use methods of information retrieval through CNKI, Wanfang, VPCS, and Foreign journal literature database, collecting research results related to interest graph, video content analysis, personalized recommendation systems, etc., to provide a theoretical basis for the research of recommendation systems.

4.3.2. Mathematical modeling method

The space vector model is used to model user interest. By extracting the interests of different users which expressed with keywords, topic interests can be formed, which is classified to build the user interest model. Finally, the user interest graph of the whole network can be constructed.

Based on the LRCN model, the BLSTM network is used to replace the unidirectional LSTM network, and propose the LRCN model based on the BLSTM network.

4.3.3. Case study method

Taking the Pear Video website as an example, we collect the public information of its users with web crawler technology, construct user interest graph with Neo4j, and research interest inference, evolution, etc.

4.3.4. Experimental research methods

We use the PyTorch deep learning framework and existing open-source datasets (e.g. HMDB-51, UCF-101, THUMOS-2014, etc.), to conduct simulation analysis of the video content analysis model based on deep learning, verifying its feasibility and effectiveness.

5. Conclusion

With the development of plenty of video platforms and live broadcast platforms, the problem of personalized recommendation of UGC video resources needs to be solved urgently, and this paper proposes a solution. In this paper, by improving the whole network user interest graph and UGC video resource personalized recommendation framework, traditional recommendation limitations can be addressed, such as low precision and poor recommendation quality. In general, due to the lack of key technologies in the field of recommendation based on interest graph at present, how to use interest graph to achieve personalized recommendation of online video resources has always been a difficult problem for the academic community, which still needs further research and discussion.

Acknowledgement

Supported by “College Students Innovation and Entrepreneurship Training Program” (Grant No.201911647014)

References

- [1] LYNNE G. The interest graph architecture-social modeling and information fusion [C] //Proc of SPIE.[S.l.]: [s.n.], 2012:1-46 .
- [2] BERKOVSKY S, KUFLIK T, RICCI F. Mediation of user models for enhanced personalization in recommender systems [J]. User Modeling and User-adapted Interaction, 2008, 18(3): 245-286 .
- [3] Ding Xuwu, Wu Zhong, Xia Zhijie. Research on the construction of social

- e-commerce user interest graphs [J]. Information Theory and Practice, 2015, 38 (3): 90-94
- [4] Mo Lingfei, Jiang Hongliang, Li Xuanpeng. Based on Review of research on video prediction for deep learning . Journal of Intelligent Systems, 2018 (1): 85-96
- [5] Hu Erlei, Feng Rui. Image retrieval system based on deep learning [J]. Computer System Application, 2017, 26 (3): 8-19
- [6] He Qingqiang. Research on video content recognition technology based on deep learning [J]. University of Electronic Science and Technology of China, 2017
- [7] Zhang Yaming, Liu Haiou. Network community marketing recommendation mechanism of user interest graphs of cloud environment [J]. Intelligence Magazine, 2013 (3):