2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)

A User-based Collaborative Filtering Algorithm Considering Interest Attenuation for Short Video Recommendation

Tong Chen
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China
e-mail: 22121015@bjtu.edu.cn

Chunyan Li
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China

Minshu Ma*
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, China
*Corresponding author: mshma@bjtu.edu.cn

***Abstract*—To solve the problems of recommendation delay and repeated recommendation in short video recommendation, this paper propose a user-based collaborative filtering algorithm considering interest attenuation for short video recommendation. First, we extracted the user's positive feedback behavior from the user's historical interaction behavior records and quantified the user's interest through weighted calculation. Then, we build an interest time decay function by taking into account the features of user interest attenuation, which in turn improves the methods of similarity calculation and rating prediction in the collaborative filtering algorithm. The offline experimental results show that the algorithm we proposed outperforms the traditional algorithm in accuracy, proving the effectiveness of the improved method and providing implications for solving the problem of users' changing interest.**

Keywords- short video recommendation; collaborative filtering; interest attenuation

I. INTRODUCTION

With the development of the information society, short video satisfies the pan-entertainment needs of modern people with the characteristics of short and concise, rich in content, and rapid transmission speed. It quickly became the mainstream medium and attracted numbers of users [1]. Given the huge amount of short video corpus and the huge number of users, the recommendation system can efficiently help users get the short video content they are interested in. The recommendation algorithm has become one of the core competencies of short video platforms.

However, in practice, the personalized recommendation of short video has come to problems of repeated recommendation, recommendation delay, filtering bubbles and so on [2]. The quality of short video content is uneven and the timeliness is short. On the other hand, the user's personalized needs are significant and their interest changes rapidly. Therefore, the recommendation algorithm must take into account the user's changing interest [3].

To alleviate the problem of data sparseness, one approach is to extend the information sources such as video features for deformation similarity calculation [4, 5]. Another approach is to extend the information sources for feature learning in model-based recommendation to capture deep features [6, 7, 8]. As to processing users' interest changing, one approach is to mine user long- and short-term interest through sequential modeling [9,10]. A multi-aggregator temporal heterogeneous graph neural network based on sequential sessions was proposed by Han J. et al., which synthesizes the features of short videos to capture users' dynamic preferences [11]. There is also an approach that incorporate the time factor as a weight into the recommendation algorithm [12, 13]. Li J. et al. utilized the exponential moving average method to assign weights to social tags based on their frequency over time, in order to model users' preference changes [14]. Existing research primarily focuses on whether users has browsed or liked the short video, however, other types of user's interaction behaviors are less concerned. The recognition of user interest is not accurate enough, and the characteristic of short video user interest attenuation are not fully considered.

Based on the processes of user-based collaborative filtering, we design an algorithm for short video recommendation. The main contribution of this work is: we describe the law that the interests of users in short video platform decay over time based on Gossen's first law i.e. the theory of diminishing marginal utility. And for fitting this law, we adapt an exponential time interest decay function to recommendation algorithm.

II. METHODOLOGY

A. Process Overview

The process of the user-based collaborative filtering algorithm considering interest attenuation for short videos recommendation that we proposed is illustrated in Fig. 1. When

2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT) | 979-8-3503-8555-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/AINIT61980.2024.10581527

generating a recommendation list for a user, firstly, it the user's interest ratings are calculated based on the user's historical interaction records. Then, the function for interest time decay is then introduced to calculate user similarity and the user similarity matrix is obtained. And the weight of interest ratings is determined by the time interest decay function to predict ratings of candidate feeds. Finally, a list of recommended feeds for the user is generated based on the predicted ratings.

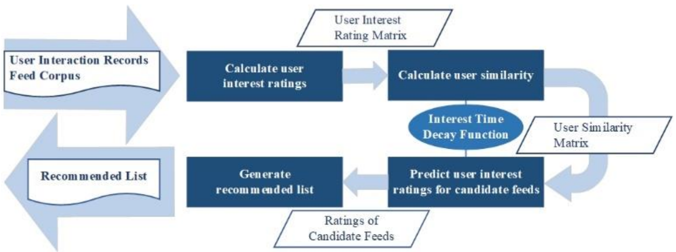

Figure 1. Process of the user-based collaborative filtering algorithm considering interest attenuation for short video recommendation

B. Calculation of Interest Rating Based on User Interactions

When users watch feeds on the short video platform, a large number of interaction records are generated, which can reflect the user's interest in the short video item [15]. We extract user interactions from the user behavior log and normalize them according to whether the feedback of user behavior is positive.

Specifically, for the user's comments, text sentiment analysis is used to analyze the comment content, and judge whether the feedback of this comment is positive according to its emotional tendency [16]. For the play behavior, based on the fuzzy theory, we use the play rate (play duration/short video duration) to measure users' interest.

According to the contribution of each interaction behavior to represent user's interest preferences, the corresponding weights are given in Table 1.

Table 1 Processing of users' historical Interaction records

User Interaction Behavior	Values	
	Field value (a_{uk})	***Weight (w_k)***
like	{0,1}	1
read comment	{0,1}	1
comment	{0,1}	2
favorite	{0,1}	2
forward	{0,1}	2
follow	{0,1}	2
click avatar	{0,1}	2
play rate	[0,1]	1

According to the user interaction behavior values and weights in Table 1, the user's interest score for short videos is calculated by (1).

$$r_{ui} = \frac{\sum_{k\in\{ACTIONS\}} w_k \times a_{uk}}{\sum_{k\in\{ACTIONS\}} w_k} \quad (1)$$

Where r_{ui} denotes the interest score of user u for feed i.

C. Algorithm Considering User Interest Attenuation

1) Time interest decay function

Users browsing short video is a kind of consumption behavior that satisfies their psychological needs. By data analysis, we discover that the changes of users' interests in short video platform follows Gossen's first law, exhibiting attenuation characteristics [17]. Users satisfy their interest needs in some kind of short videos by watching them. After watching such short videos for several days, the level of their interest in such short videos will decline. When the level of interest declines to 0, that is, this interest is fully satisfied. If they watch this kind of short videos again, they will only be disgusted by such short videos. That is, the users' interest decreases over time.

The exponential time decay model can well characterize the law of user interest decay [18]. To characterize the law of user interest decay, we use the exponential time decay model as the time interest decay function, which is shown in (2).

$$f(t) = \frac{2}{1+e^{-\frac{T-t_0}{\alpha(T-t)+bias}}} - 1 \quad (2)$$

where T is the current time, t_0 is the time of the earliest user interaction behavior records, t is the time of the interaction behavior, $bias$ is a tiny value added to prevent operational errors when the denominator is 0. α is a time decay factor, which controls the decay rate of $f(t)$ as $(T-t_0)/(T-t)$ decreases, and is set according to the range of $T-t_0$ and $T-t$.

2) Similarity calculation added user interest decay

Based on the concept of collaborative filtering, users are considered similar if they have shown interest in the same items. However, due to user interest attenuation, the interests of different users may be asynchronous.

For instance, suppose that both user u and v rated the same item 7 days ago, while another user w rated the same item just 1 day ago. Considering user interest attenuation only, from 7 days ago to 1 day ago, the user's interest in this item has been attenuated. Because users v exhibit more consistent changes with user u in their interests than users w and u, they are more similar. Therefore, to ensure consistency of interests across different users, it is necessary to adjust the user interest ratings to the same time when measuring user similarity.

Thus, we add the time interest decay function to the user similarity calculation as the weight of the interest ratings, as shown in (3).

$$sim(u,v) = \frac{1}{1+e^{-N(I_u \cap I_v)}} \times \frac{\sum_{i\in I_u\cap I_v}(r_{ui}\times f(t_{ui})-\overline{r_u})\cdot(r_{vi}\times f(t_{vi})-\overline{r_v})}{\sqrt{\sum_{i\in I_u\cap I_v}(r_{ui}\times f(t_{ui})-\overline{r_u})^2}\times\sqrt{\sum_{i\in I_u\cap I_v}(r_{vi}\times f(t_{vi})-\overline{r_v})^2}} \quad (3)$$

where t_{ui} and t_{vi} are the times when users u and v rated item i, respectively.

3) Rating prediction added user interest decay

To make a rating prediction, it is necessary to obtain the K nearest neighbor user set of the target user for the target item. Then, the algorithm calculates the predicted rating based on the ratings of each nearest neighbor user on the target item and their similarities with the target user. As users' interest decay over time, the effectiveness of their interest ratings also decreases. Therefore, the algorithm utilizes the interest time decay function to weigh the ratings. (4) is used to predict the rating of user u on item i.

$$p(u,i) = \overline{r_u} + \frac{\sum_{v \in S(u,i,K)} sim(u,v) \times (r_{vi} \times f(t_{vi}) - \overline{r_v})}{\sum_{v \in S(u,i,K)} sim(u,v)} \quad (4)$$

III. EXPERIMENTS

We conduct an offline experiment using a public dataset of 2021 WeChat Big Data Challenge (algo.weixin.qq.com/2021/). The dataset is sourced from WeChat Video platform and consists of two files: one containing the interaction records of 20,000 users over a 14-day period, and the other containing 106,444 feed information records.

The algorithm's prediction accuracy and recommendation effect were tested using the user interaction behavior data from the first 13 days as the training set and the data from the 14th day as the test set.

We use the user interaction behavior data from the first 13 days as the training set, and the data from the 14th day as the test set to test the proposed algorithm performance in prediction and recommendation.

A. Metrics

For the rating prediction task, we evaluate the accuracy of the prediction rating by mean absolute error (MAE) and root mean square error (RMSE). The smaller the value, the higher the prediction accuracy.

$$MAE = \frac{\sum_{(u,i) \in Test} |r_{ui} - p(u,i)|}{N(Test)} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{(u,i) \in Test} \left(r_{ui} - p(u,i)\right)^2}{N(Test)}} \quad (6)$$

For the recommendation task, we evaluate the recommended results by F1-score. The F1-score indicates the accuracy of the recommendation, with higher values indicating greater accuracy.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

B. Parametric Experiments

Parameter experiments are conducted to determine the time decay factor α of the exponential time interest decay function. Specifically, the time decay factor α is set according to the range of $T - t_0$ and $T - t$. According to the dataset, $T = 14$ and $t_0 = 1$, so $bias = 0$. Therefore, α ranges within [1/13, 13]. Different levels of time decay factor α are set in the function to predict the users' ratings on the 14th day based on their historical interaction behavior records from the previous 13 days. The prediction results of the user-based collaborative filtering algorithm considering interest attenuation (TR-UserCF) are evaluated for MAE and RMSE and shown in Fig. 2.

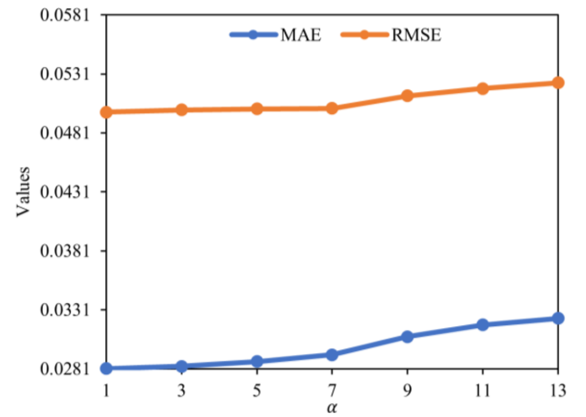

Figure 2. Prediction results in accuracy of TR-UserCF algorithm with different time attenuation factor α

Fig. 2 shows that as the time attenuation factor α increases, the MAE and RMSE slightly decrease and then increase, indicating a slight increase and decrease in prediction accuracy. The smallest values of MAE and RMSE occur when α is 1, making it the optimal value.

C. Comparative Experiments

1) Rating Prediction

To assess the effectiveness of the time interest decay function (ExponentialT) in characterizing user interest changes, a comparative experiment in predicting ratings is conducted. Keeping other conditions unchanged, we select three methods as comparisons to predict the ratings on the 14th day: the classical time decay method (InverseT) [19], the time forgetting curve fitting method (EbbinghausT) [20], and the user collaborative filtering method without considering time (UserCF). The results of the experiment are presented in Table 2.

Table 2 Prediction Accuracy of Different Methods

Prediction Accuracy of Different Algorithms	**Metrics**	
	MAE	***RMSE***
InverseT	0.05591	0.081231
EbbinghausT	0.062235	0.085811
ExponentialT	0.045997	0.072752
UserCF	0.046026	0.07278

Upon comparing the MAE and RMSE of the rating prediction results, it can be concluded that the method we proposed (ExponentialT) has a higher prediction accuracy than other methods. Thus, the algorithm based on the exponential time decay function can better characterize the user interest attenuation.

2) Recommendation Generating

To evaluate the recommendation performance of our proposed method, while keeping other conditions constant, we compared recommendation results based on user collaborative filtering (UserCF), rating-based user collaborative filtering (R-UserCF) and our proposed algorithm (TR-UserCF) on F1-scores, across different levels of TopN.

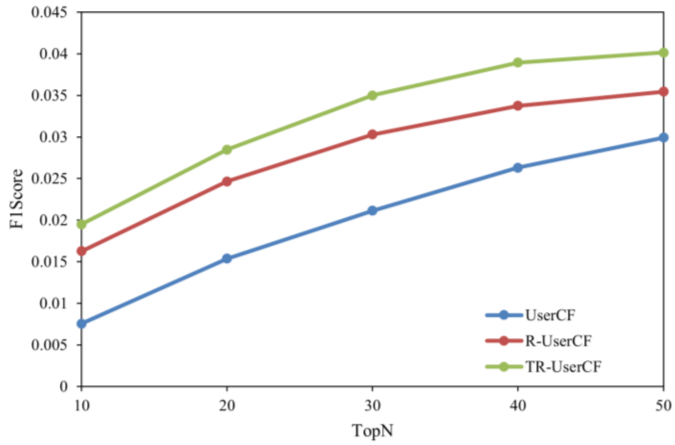

Figure 3. Recommendation results of UserCF, R-UserCF, and TR-UserCF in different levels of TopN

The experimental results are shown in Fig. 3. It can be seen from Fig. 3, the F1-scores of UserCF, R-UserCF, and TR-UserCF increase as TopN increases, but at different rates. Our proposed TR-UserCF has a significantly higher F1-score than UserCF and R-UserCF.

Overall, our proposed algorithm performs better in the above comparative experiments, and our improvements are effective.

IV. CONCLUSIONS

In short videos recommendation, the traditional user-based collaborative filtering utilizes little the user historical interaction behavior records and hardly concerns the characteristics of user interest attenuation. This leads to problems such as repeated recommendations and recommendation delay in practice. Our proposed algorithm based on user collaborative filtering extracts positive feedback from the user's historical interaction behavior records and measures their interest through weighted calculation. It considers the user interest attenuation by adding a time interest delay function to user similarity calculation and rating prediction.

However, the algorithm's high time complexity means it does not address the issue of poor scalability in collaborative filtering. We think the inherent memory-based characteristic of collaborative filtering make it exist this defect. We will study further to reduce algorithm’s computational time and memory complexity for short video recommendation.

REFERENCES

[1] CNNIC, 2024. The 52nd Statistical Report on the Development of Internet in China. https://www.cnnic.net.cn/n4/2023/0828/c88-10829.html

[2] Ji, Y. (2023) Algorithm Perception and Interaction of Short Video Recommendation System:An Interview Study Based on College Students. Research on Transmission Competence, vol. 7, pp. 154-156. https://qikan.cqvip.com/Qikan/Article/Detail?id=7109202541

[3] J. Davidson, et al., 2010. The YouTube video recommendation system. In RecSys '10. NY, USA. pp. 293-296. https://doi.org/10.1145/1864708.1864770

[4] S. Zhu, S. Wei, S. Wei, and D. Yu. (2021) Video recommendation algorithm based on danmaku sentiment analysis and topic mode. Journal of Computer Applications, vol. 41, pp. 2813-2819. https://link.cnki.net/urlid/51.1307.tp.20210714.1916.026

[5] N. Zhao, W. Pi and C. Xu. (2020) Video Recommendation Algorithm for Multidimensional Feature Analysis and Filtering, Computer Science, vol. 47, pp. 103-107. https://link.cnki.net/urlid/50.1075.TP.20200320.0711.002

[6] D. Cai, S. Qian, Q. Fang, J. Hu, W. Ding, and C. Xu. (2023) Heterogeneous Graph Contrastive Learning Network for Personalized Micro-Video Recommendation, IEEE Transactions on Multimedia, vol. 25, pp. 2761 - 2773. http://dx.doi.org/10.1109/TMM.2022.3151026

[7] S. Liu, H. Liu, Z. Chen, and X. Hu, 2019. User-video co-attention network for personalized micro-video recommendation. In WWW '19. San Francisco, CA, United states. pp. 3020-3026. http://doi.org/10.1145/3308558.3313513

[8] X. Du, H. Yin, L. Chen, Y. Wang, Y. Yang, and X. Zhou. (2020) Personalized Video Recommendation Using Rich Contents from Videos. IEEE Transactions on Knowledge and Data Engineering, vol. 32, pp. 492-505. http://dx.doi.org/10.1109/TKDE.2018.2885520

[9] P. Symeonidis, et al., 2020. Recommending the Video to Watch Next: An Offline and Online Evaluation at YOUTV.de, in RecSys '20. NY, USA, pp. 299–308. https://doi.org/10.1145/3383313.3412257

[10] C. Gao, Y. Li, and D. Jin. (2021) Video Recommender System with Graph Neural Networks. ZTE Technology Journal, vol. 27, pp. 27-32. https://link.cnki.net/urlid/34.1228.TN.20210120.1010.002

[11] J. Han, W. Li, Z. Cai, and Y. Li, 2022. Multi-Aggregator Time-Warping Heterogeneous Graph Neural Network for Personalized Micro-Video Recommendation. In CIKM '22. NY, USA, pp. 676 – 685. https://doi.org/10.1145/3511808.3557403

[12] J. Wu and M. Xu. (2021) Video Personalized Recommendation Based on User Profile and Video Interest Tags. Information Science, vol. 39, pp. 128-134. DOI: 10.13833/j.issn.1007-7634.2021.01.017

[13] Y. Huang, B. Cui, J. Jiang, K. Hong, W. Zhang, and Y. Xie, 2016. Real-time Video Recommendation Exploration. In SIGMOD '16. NY, USA, pp. 35–46. https://doi.org/10.1145/2882903.2903743

[14] J. Li, C. Li, J. Liu, J. Zhang, and M. Wang. (2019) Personalized Mobile Video Recommendation Based on User Preference Modeling by Deep Features and Social Tags. Applied Sciences, vol. 9, p. 3858. DOI:10.3390/app9183858

[15] T. Gao. (2021) Research on short video recommendation strategy based on big data analysis. Journal of Physics: Conference Series, vol. 1941, pp. 012071 (9pp). DOI:10.1088/1742-6596/1941/1/012071

[16] H. Xu, A. Wang, W. Che, Y. Du, W. Sun, and Y. Wang. (2022) Micro-Blog Text Sentiment Analysis Model Based on Multi Model Fusion. Journal of Shenyang University (Natural Science), vol. 34, pp. 112-118+133.https://link.cnki.net/doi/10.16103/j.cnki.21-1583/n.2022.02.005

[17] A. Jolink and J. van Daal. (1998) Gossen's laws. HISTORY OF POLITICAL ECONOMY, vol. 30, pp. 43-50. DOI:10.1215/00182702-30-1-43

[18] S. Liang, X. Peng, B. Qin, W. Lin, and Z. Hu. (2021) Research on a Collaborative Filtering Algorithm of Time Context Optimization for Book Recommendation. Library Tribune, vol. 41, pp. 113-121. https://link.cnki.net/urlid/44.1306.G2.20201013.1404.002

[19] L. Yang, Y. Hu and J. Shao. (2016) Preference prediction method based on time attenuation and preference fluctuation. Journal of Computer Applications, vol. 36, pp. 2011-2015. https://kns.cnki.net/kcms2/article/abstract?v=S5uBaE2M3OfXSve4_5nq8H9gLXq9r7iIjuwOlG3WUA5i-iuJd3kJULUd6FJNq-vshRMwp4eapr-rgdkpwb2LDwRBRatiii3YBW0rVvVpcj_FyID0LI3g_fsjJp8RSd3WzrqKTlWAfHU=&uniplatform=NZKPT&flag=copy

[20] H. Yu and Z. Li. (2010) A collaborative filtering recommendation algorithm based on forgetting curve. Journal of Nanjing University (Natural Science), vol. 46, pp. 520-527. https://qikan.cqvip.com/Qikan/Article/Detail?id=35470883