Short video recommendations based on analytic hierarchy process and collaborative filtering algorithm

Dongyuan Ma¹, Li Luo², Yuanchen Fang³*
¹Sichuan University Business School, Chengdu, P.R.China
²Sichuan University Business School, Chengdu, P.R.China
³Sichuan University Business School, Chengdu, P.R.China

Corresponding author and e-mail: Yuanchen Fang, y.fang@scu.edu.cn

Abstract. In video-sharing mobile apps, personalized recommendation systems can help users quickly finding the mobile videos of interest through analyzing videos characteristics and user behaviors, and thus have been widely applied in many applications. Collaborative filtering is one of the most popular technologies in the personalized recommendation systems. To improve the performance of personalized recommendation services, this paper proposes a new personalized collaborative filtering algorithm recommendation by combining the analytic hierarchy process, which captures various video information and user behaviors, and comprehensively evaluates users’ interest in videos. Experiments on NetEase Cloud Music app cloud village dataset shows that the proposed algorithm outperforms traditional method in recall rate, accuracy rate and coverage rate, offering more accurate personalized recommendations.

1. Introduction
Personalized recommendation is an important method for preventing user information overload [1]. Because collaborative filtering algorithms require little specific field knowledge and are easy to implement in engineering, they have become one of the most popular methods used in recommendation systems. Collaborative filtering is generally divided into user-based collaborative filtering, item-based collaborative filtering, and model-based collaborative filtering [2]. Collaborative filtering employs user ratings or click behaviors to calculate the similarity between users and items, and then provides recommendations based on similarity ranking. Scholars have proposed several improvement measures for collaborative filtering. Cao et al. proposed a hybrid e-commerce recommendation algorithm based on collaborative filtering and a content-based method [3], Liu et al. proposed a recommendation algorithm that combines social and collaborative filtering [4], Huang X et al. proposed an improved collaborative filtering recommendation algorithm based on user learning behavior and ratings [5], and Hong B et al. proposed a new collaborative filtering algorithm based on correlation coefficients (CORs) [6]. Researchers have found that considering a greater number of user behaviors and attributes improves the effects of collaborative filtering.

In short video apps, users currently exhibit rich behaviors such as clicking, liking, commenting, and sharing. The combination of a user’s different behaviors comprehensively reflects his/her degree of interest in a current video. For example, for two videos A and B, suppose a user only clicks on A, but clicks on and likes B. This reflects the different degrees of interest the user has in videos A and B. To
the best of the authors’ knowledge, current collaborative filtering algorithms do not take such behavior combinations into consideration when calculating similarity.

Therefore, this paper proposes a comprehensive evaluation of user interest in a set of items based on the analytic hierarchy process considering multi-dimensional user behavior. The similarity required for collaborative filtering is calculated based on the degree of user interest output by the analytic hierarchy process. In this manner, the similarity between user-user and item-item can be calculated more accurately, thereby achieving a more accurate personalized recommendation.

2. Relevant literature

2.1. Analytic hierarchy process
The Analytic Hierarchy Process (AHP) is a commonly used analysis method in the decision-making field. AHP is a systematic and hierarchical analysis method that combines qualitative and quantitative analysis. The characteristic of AHP is that, on the basis of in-depth research on the nature, influencing factors and internal relations of complex decision-making problems, using less quantitative information to mathematicize the thinking process of decision-making provides simple decision-making methods for multi-objective, multi-criteria or complex decision-making problems [7]. Liu Y et al. analyzed the degree of passenger satisfaction with bus routes based on the analytic hierarchy process [8]. Lina S et al. proposed an impact analysis method for top management teams and investment efficiency based on the analytic hierarchy process and support vector machines to improve the accuracy of the impact analysis and investment efficiency of top management teams under international accounting standards [9]. Ren Y et al. proposed an integrated network public opinion evaluation algorithm combining a bp neural network and the analytic hierarchy process [10]. These studies show that the analytic hierarchy process is useful for comprehensive evaluation.

2.2. Collaborative filtering
As one of the most widely applied algorithms in recommendation systems, collaborative filtering is associated with several theories and methods. Collaborative filtering can be divided into three main categories. The first category is user-based collaborative filtering, which calculates the similarity of users, finds the K users most similar to the current user, and recommends based on the preferences of the neighbor user set [11]. The second category is item-based collaborative filtering, which calculates the similarity of items and recommends similar items after a user expresses interest in an item [12]. Item-based collaborative filtering algorithms are currently the most widely used algorithms in the industry, i.e., the recommendation systems of Amazon, Netflix, Hulu, and YouTube are based on item-based collaborative filtering [2]. The third category is model-based collaborative filtering. In a nutshell, model-based collaborative filtering applies different modeling methods to a collaborative filtering algorithm, obtains a predictive model through training historical data, and makes recommendations. Model-based collaborative filtering generally includes matrix decomposition [13], neural networks [14], and clustering [15].

3. Improved collaborative filtering algorithm based on the analytic hierarchy process
To increase performance, this paper proposes an improved collaborative filtering algorithm based on the analytic hierarchy process. Its implementation framework is shown in Figure.1.

3.1. User behavior weight calculation based on the analytic hierarchy process
We adopt the analytic hierarchy process to assign weights to indicators and assign different weights to different user behaviors. The accumulation of behaviors and weights indicates the degree of user interest in a video.

\[ U = \{u_1, u_2, \ldots, u_n\} \]
The \( n \) behaviors of the user are represented by (1). For each behavior, we compare its importance in pairs, and get a consistency judgment matrix (2):

\[
A = \begin{pmatrix}
    r_{11} & \cdots & r_{1n} \\
    \vdots & \ddots & \vdots \\
    r_{n1} & \cdots & r_{nn}
\end{pmatrix}
\]

(2)

Where \( r_{ii} = 1 \), and \( r_{ij} + r_{ji} = 1 \). A description of \( r_{ij} \) is given in Table 1. Next, we calculate the weight of user behavior:

\[
CI = \frac{\lambda - n}{n-1}
\]

(3)

\[
CR = \frac{CI}{RI}
\]

(4)

\[
A_w = \lambda w
\]

(5)

Where \( \lambda \) is the largest characteristic root of matrix \( A \), and \( RI \) is a random consistency coefficient. A \( CR \) of less than 0.1 indicates that the constructed judgment matrix is reasonable. Then, we calculate the weight of each user's behavior according to (5).

Finally, we obtain the user's comprehensive interest in each item via (6).

\[
COM_{\text{rating}} = U_w
\]

(6)

**Figure 1.** Framework of improved collaborative filtering algorithm based on the analytic hierarchy process.
Table 1. Description of $r_{ij}$ values.

| Value | Description                        |
|-------|------------------------------------|
| 0.1   | $i$ is absolutely not important to $j$ |
| 0.2   | $i$ is very unimportant to $j$       |
| 0.3   | $i$ is less important to $j$         |
| 0.4   | $i$ is slightly less important to $j$|
| 0.5   | $i$ is equally important to $j$      |
| 0.6   | $i$ is slightly more important to $j$|
| 0.7   | $i$ is more important to $j$         |
| 0.8   | $i$ is very important to $j$         |
| 0.9   | $i$ is absolutely important to $j$   |

3.2. Personalized recommendation based on collaborative filtering

We use cosine similarity to calculate the similarity among users/items. To calculate user similarity when the user behavior is ‘whether to click or not’, we use (7):

$$
\text{CosSim}(u,v) = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u)||N(v)|}}
$$

(7)

where $N(u)$ represents the collection of items for which user $u$ had positive feedback, and $N(v)$ represents the collection of items for which user $v$ had positive feedback.

Using the degree of comprehensive interest defined in this paper to calculate user similarity, we use equation (8) to calculate:

$$
\text{CosSim}(u,v) = \frac{\sum_{i \in I_u} r_{ui} r_{vi}}{\sqrt{\sum_{i \in I_u} r_{ui}^2 \sum_{i \in I_v} r_{vi}^2}}
$$

(8)

where $I_u$ represents the set of items scored jointly by user $u$ and user $v$, $I_v$ represents the set of items rated by user $u$, $I_v$ represents the set of items rated by user $v$, $r_{ui}$ is the rating of user $u$ for items $i$, and $r_{vi}$ is the rating of user $v$ for items $i$.

For item similarity when the user behavior is ‘whether to click’, we use formula (9) to calculate:

$$
\text{CosSim}(i,j) = \frac{|N(i) \cap N(j)|}{\sqrt{|N(i)||N(j)|}}
$$

(9)

where $N(i)$ is the number of users who click item $i$, and $N(j)$ is the number of users who click item $j$.

Using the degree of comprehensive interest defined in this paper to calculate user similarity, we use equation (10) to calculate:

$$
\text{CosSim}(i,j) = \frac{\sum_{u \in I_u} r_{ui} r_{vj}}{\sqrt{\sum_{u \in I_u} r_{ui}^2 \sum_{u \in I_v} r_{vj}^2}}
$$

(10)
where $U_{ij}$ is the set of users who simultaneously click items $i$ and items $j$, $U_i$ is the set of users who click item $i$, $U_j$ is the set of users who click item $j$, $r_{ui}$ is user $u$’s rating for item $i$, and $r_{uj}$ is user $u$’s rating for item $j$.

4. Results

4.1. Experimental data

We used the NetEase Cloud Music app cloud village for short video data [16]. This data set is provided by the Revenue Management and Pricing Association of INFORMS and NetEase Cloud Music to support operations data-driven research in management. The data contains 57750395 impression logs of 2085533 users and 252762 short videos randomly sampled from November 1, 2019 to November 30, 2019. We selected videos with a certain level of popularity and users with a specified degree of activity (videos played greater than 100 times and user clicks greater than 50 are used as standards). The experimental data contains 438350 click logs of 4912 users and 5247 short videos.

4.2. Comparison of experimental results

We score user behavior across five dimensions: click, like, comment, share, and video watching completeness. Using experts to score the relative importance of each dimension, we obtained the consistency judgment matrix shown in Table 2. Using $CR=0.0144$, the judgment matrix structure is reasonable. The weight vector of user behavior $w=(0.09, 0.2, 0.24, 0.24, 0.23)$ is obtained using the method described in Section 3. Through $w$, we obtain the user’s comprehensive interest rate, or $\_COM\_rating$, for each item. To improve the distinction between items with positive user feedback and items that users did not click, we add 1 to each $\_COM\_rating$.

We compare the performance of $\_COM\_rating$ and click-only (CLICK) methods on user-based and item-based collaborative filtering. By choosing a different number of K neighbors, we mainly compare the performance of the improved and traditional methods on recall, accuracy, and coverage. The results are shown in Figure 2 and Figure 3.

In the user-based collaborative filtering algorithm, the $\_COM\_rating$ method improves the coverage rate by two percentage points when the recall and accuracy are equivalent. In the item-based collaborative filtering algorithm, the $\_COM\_rating$ method conveys clear advantages under the three performance indicators. The results show that the method proposed in this paper effectively improves the effects of collaborative filtering.

Table 2. User behavior consistency judgment matrix.

|          | Click | Like | Comment | Share | Watch completeness |
|----------|-------|------|---------|-------|-------------------|
| Click    | 0.5   | 0.3  | 0.2     | 0.2   | 0.2               |
| Like     | 0.7   | 0.5  | 0.4     | 0.4   | 0.5               |
| Comment  | 0.8   | 0.6  | 0.5     | 0.5   | 0.5               |
| Share    | 0.8   | 0.6  | 0.5     | 0.5   | 0.5               |
| Watch completeness | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 |
5. Conclusions
Considering the complexity and uncertainty of short video user behaviors, we propose a personalized recommendation algorithm based on the analytic hierarchy process to give a comprehensive evaluation of users' interest in items. Collaborative filtering models constructed comprehensive user interest to effectively improve the performance of recommendation algorithms. Based on this paper and related research findings, the comprehensive consideration of multiple user behaviors is an effective way to improve the effectiveness of recommendation systems; however, the question of how to
comprehensively consider user behaviors lacks theoretical guidance. The analytic hierarchy process used in this paper has been fully applied to multi-criteria decision-making over the past few decades, and this paper represents a first attempt at applying it to recommendation systems. The results show that the proposed algorithm improves the performance of recommendation systems. In future research, fuzzy decision theory methods will be used in recommendation systems to better measure complex user behavior in mobile apps.

Acknowledgement
We acknowledge the financial support from Sichuan Provincial Department of Science and Technology Project NO. JDR0125.

References
[1] Ricci F, Rokach L, Shapira B. Recommender Systems: Introduction and Challenges [M]. In: Ricci F, Rokach L, Shapira B. (Eds) Recommender Systems Handbook (Second Edition), Springer US, 2015:1-36
[2] Xiang Liang. The actual combat of recommendation system[M]. Beijing, People's Posts and Telecommunications Press, 2012(In Chinese)
[3] Cao Y, He W. A Mixed Recommender Mode Based on content and collaborative Filtering[C]//The 2008 International Conference on Business Intelligence and Financial Engineering (BIFE 2008)
[4] Liu Xudong, Ge Junjie, Chen Deren. A Hybrid Recommendation Algorithm Based on Clustering and Collaborative Filtering[J]. Computer Engineering and Science, 2010, 032(012):125-127,133.( In Chinese)
[5] Huang X, Wang G. Learning recommendation based on hybrid collaborative filtering algorithm[C]//Journal of Physics: Conference Series. IOP Publishing, 2020, 1629(1): 012008.
[6] Hong B, Yu M. A collaborative filtering algorithm based on correlation coefficient[J]. Neural Computing and Applications, 2019, 31(12): 8317-8326.
[7] Wang Lianfen, Xu Shubai. Introduction to Analytic Hierarchy Process[M]. Renmin University of China Press, 1990. ( In Chinese)
[8] LIU Y, Min Y. Analysis of Passenger Satisfaction for Bus Line Based on Analytic Hierarchy Process[J]. DEStech Transactions on Social Science, Education and Human Science, 2018.
[9] Lina S. Analysis of factors affecting investment efficiency based on analytic hierarchy process and support vector machine (SVM) model[J]. Cluster Computing, 2019, 22(2): 4367-4374.
[10] Ren Y. A Rapid Evaluation Method for Network Public Opinion Based on Neural Network and Analytic Hierarchy Process [J].
[11] Dong J, Li J, Li G, et al. Research on User-based Normalization Collaborative Filtering Recommendation Algorithm[C]// 2nd International Conference on Advances in Mechanical Engineering and Industrial Informatics (AMEII 2016). 2016.
[12] Linden G, Smith B, York J. Amazon. com recommendations: Item-to-item collaborative filtering[J]. IEEE Internet computing, 2003, 7(1): 76-80.
[13] Ortega F, Hernando A, Bobadilla J, et al. Recommending items to group of users using matrix factorization based collaborative filtering[J]. Information Sciences, 2016, 345: 313-324.
[14] Feng Zou. Recommendation algorithm based on deep neural network and improved similarity measurement[J]. Computer Applications and Software,2019,36(11):286-293+300.
[15] Shen Y, Pedrycz W. Collaborative fuzzy clustering algorithm: Some refinements[J]. International Journal of Approximate Reasoning, 2017, 86: 41-61.
[16] Zhang D, Hu M, Liu X, et al. NetEase Cloud Music Data[J]. Available at SSRN 3554826, 2020.