



Barrage Video Recommendation Method Based on Convolutional Recursive Neural Network

Siyuan Ma and Ping He^(✉)

Institute of Information Technology, Hebei University of Economics and Business,
Hebei, China
11112082@bjtu.edu.cn

Abstract. Aiming at solving the problems of traditional video recommendation process, such as low time efficiency and low accuracy, this paper proposes a convolutional recursive neural network video recommendation method based on barrage. Firstly, according to the number of barrage in video, the method selects the preferable video fragments of users, and adopts the k-means clustering method to extract key frames. Secondly, we construct a convolutional recursive neural network model (RCNN) to classify the similar video fragments. Finally, the recommendation can be achieved by the similarity between video fragments. The experimental results show that the proposed recommendation method improves the accuracy of recommendation by 0.22 compared with CTR, 0.18 compared with CDL, and 0.31 compared with ConvMF. So it improves the accuracy of recommendation in a certain extent.

Keywords: Barrage · Convolutional neural network · Recursive neural network · Video

1 Introduction

With the rapid development of digital technology and network technology, the domestic and foreign appear many big video websites. After more than ten years of exploration and practice, major video sites (domestic video sites such as iQIYI, Tencent, Youku, foreign video sites such as YouTube and Netflix, etc.) have already occupied a majority position in the field of video website. And watching the video has been integrated into people's daily life. Nevertheless, it is an important problem that how to timely recommend video that the users may be

This work is supported by Key Project of Hebei Educational Committee (Higher Education Research Project of Hebei Province) No. ZD2019017, 2019 School Level Graduate Innovation Funding Project, Natural Science Foundation of Hebei Province under Grant No. F2019207061, Science and Technology Research Project of Higher Education of Hebei Province under Grant No. QN2018116, Research Foundation of Hebei University of Economics and Business under Grant No. 2018QZ04.

interested, moreover it also becomes a key problem that needs to solve actively for each big video website.

In recent years, video that the users can send the barrage has become the mainstream [1], and researchers pay more and more attention to the barrage video. The barrage text sent by the users can reflect the feelings of users about the current segment of video, and the number of barrage can reflect the degree that the current segment of video is favored by users. Therefore, this paper studies the number of barrage in the video to improve the accuracy of video recommendation.

At present, a larger number of domestic researches are based on the analysis of barrage content. For example, according to the topic model to classify the word of barrage, the literature [2] combined with the theme distribution of each word and emotional dictionary and proposed the algorithm that can measure the emotional vector of dynamic evaluation words. It uses the “global + local” context-related emotion similarity calculation method to calculate the emotional similarity, and finally obtain the recommended video segments by the emotional similarity scores, however, the emotional classification of bullet screen content without emotional color or irony is a little insufficient.

The literature [3] firstly processed the barrage text data, calculated the emotional value of these texts, improved the traditional k-means algorithm, calculated the distance between the data by using the dynamic time normalization algorithm. And then, the emotional values are classified so as to distinguish the similarities and differences in users’ emotions when they watching video, however, the literature only calculates the positive and negative emotions of the users who send a large amount of barrage data, and cancels the users who send a small amount of barrage data, so it is impossible to classify the users who prefer this video but send a small amount of barrage data. In the emotional value analysis of barrage text, there is a problem that it is seriously relying on the barrage content, therefore this paper proposed a video recommended model that is based on the number of barrages and combined with the recursive convolutional neural network. The model uses the number of barrages to analyze the fragments favored by users in video, and adopts the K-means clustering method to extract the key frames of video clips. Furthermore, the key frames can be processed to the valid static images that can be used to the input of the recursive convolutional neural network model. Finally, the important human behavior characteristics can be extracted, and on this basis the method can be recommended to the users by looking for video with similar video frames.

The remainder of this paper is organized as follows. In Sect. 2 we describe some related studies on video recommendation and video recommendation based on convolutional neural network. Section 3 defines the problem of video recommendation on the basis of barrage. Section 4 formally describes the work flow of video recommendation that contains the data preprocessing and RCNN model. Section 5 illustrates the data sources and experimental sets, and gives the experimental results. Finally, Sect. 6 concludes this paper and explains the next researching direction of video recommendation based on barrage.

2 Related Works

This section introduces the related research results of video recommendation and video recommendation method based on convolutional neural network.

2.1 Traditional Video Recommendation

Traditional video recommendations are generally divided into content-based video mixed recommendation method and collaborative filtering based video mixed recommendation method.

- (1) Content based video recommended methods: the literature [4] proposed a recommendation method based on label weight score that each label is set the corresponding score, where the score represents the weight of the item or the user on the label, in order to reduce the influence of objective factors on user ratings and improve the accuracy and authenticity of the score. On the basis of context, Lin [5] proposed a recommendation model based on user decision that selects the required features or feature combinations according to the factors affecting user decision, and takes the preference of users' interests as the direct influencing factor of decision. Tzamouisis [6] uses various machine learning algorithms to learn the efficient combination method of various recommendation algorithms, and can select the best hybrid method according to the given input. This method can be easily extended to the other recommendation methods.
- (2) Video based on collaborative filtering recommendation method: Nguyen [7] proposed a probability model that combines the displayed and hidden feedback. The method firstly adopts the users and projects based the displayed feedback to represent the matrix decomposition model, secondly adopts the embedding model of projects to find the representation of goods, so as to capture the relationships among the projects based on hidden feedback. In order to solve the problem that a single user is only interested in some fields, Zhang [8] proposed a collaborative clustering recommendation method that adopts clustering algorithm to group the users and items according to their interests or characteristics, and then makes corresponding recommendations on the basis of each group. Zhang [9] improved the original frequency-based and ranking-based information kernel extraction methods, and proposed a method that considers the similarity between users and fully utilizes the scoring information of users and goods when seeking the list of neighbors for optimizing the process of finding the most similar neighbors.

2.2 Video Recommendation Based on Convolutional Neural Network

Li [11] proposed a method that computes video correlation directly from the content, rather than as the sub conditions of user behavior matrix decomposition. The method uses the deeping convolution neural network to process the

video information (such as the pixels, audio, subtitles and metadata) in order to build a video link table, and reduce the behavior requirement of the new video user. In addition, in order to improve the applicability of the convolutional neural network, the literature [11] uses the condition convolution to extraction the behavior features of users. The method introduces the feature vector of goods as the convolution kernel and does not set training parameters. Moreover, this method only needs one convolution layer to get the higher order combination between N attributes of user and N attributes of items. On this basis, the method integrates user and item attribute information in the neural collaborative filtering.

Cai [12] proposed an improved recommended method that combines the matrix decomposition and cross channel convolution neural network, and add the influence factors of the user and the project to the traditional matrix decomposition. Then, the information matrix composed of word vectors is input into the convolutional neural network. Finally, the eigenvalue of evaluation information is combined with the regularization term of the improved matrix decomposition model. The literature [13] applies the dynamic convolution probability matrix decomposition model to group recommendation, integrates the text representation method of convolutional neural network into the potential factor model, and integrates the state space model into the potential factor model.

3 Problem Formulation

This section introduces the related background of this paper, including the characteristics and mechanism of barrage video, and the relevant definition s. When a user sees a scene in video, he or she may write some texts and send it to the video in the period of time. The other users may send the texts to this video when they watch the same scene of video later. So when the number of barrage in this time is more than a certain number during a period of time, and it can be concluded that those users who send the video barrage are more interested in this scene. By analyzing these interested video segments, we can find the similar video in order to recommend to the users. The above questions can be summarized as Definition 1.

Definition 1 (The video recommendation model based on barrage). Given the set of video is $V = \{v_1, v_2, \dots, v_n\}$, where $|V|$ denotes the number of video in V and the time of the video v_i is T_{v_i} , the set of users is $C = \{c_1, c_2, \dots, c_m\}$, the video set contains the number of barrages in various times for n video v_i is d_{i,t_j} and the number of views is g_{ij} et al. The model firstly takes the time point t_j when the barrage numbers $d_{ij} > \lambda$ (λ is obtained after testing) in the video set v_i . According to the t_j , secondly extracts the video fragments $f_i = \{q_{i,1}, q_{i,2}, \dots, q_{i,|v_i|}\}$, where $q_{i,k}$ denotes the k -th video segment in the i -th video. Next, uses K-means clustering scheme to extract the key frame in the extracted fragment set f_i . The extracted data set S is processed to the structured data set D by introducing the recursive convolutional neural network model for

improve the accuracy ε of the recommendation system. The mathematical model of the problem is shown in Eq. (1).

$$\begin{aligned}
 \text{Input :} \quad & V = \{v_1, v_2, \dots, v_n\} \\
 & C = \{c_1, c_2, \dots, c_m\} \\
 & v_i = \{d_{ij}, g_{ij}\} \\
 & d_m \subseteq V \\
 & T = \{t_1, t_2, \dots, t_e\} \\
 & f_i = \{q_{i,1}, q_{i,2}, \dots, q_{i,|v_i|}\} \\
 \text{Output :} \quad & S \\
 \text{Goal :} \quad & D, \max \varepsilon \\
 \text{s.t. :} \quad & (1) 1 < i < n \\
 & (2) 1 < j < e \\
 & (3) n, m, e \in N
 \end{aligned}
 \tag{1}$$

4 Video Recommendation Structure

Video recommended workflow consists of two sub-modules: (1) data preprocessing; (2) RCNN model, as shown in Fig. 1.

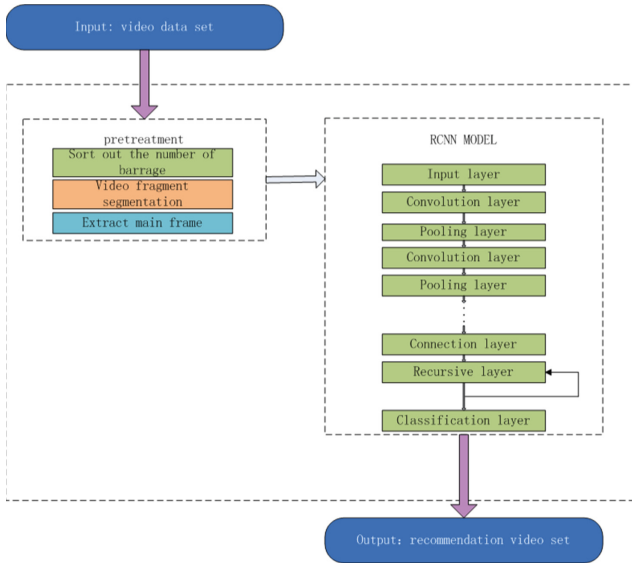


Fig. 1. The process of video recommendation

4.1 Barrage Data Preprocessing

The data preprocessing stage mainly includes three steps: sorting out the number of bullet screens, cutting video fragments, and extracting the main frame. The barrage is presented in the form of dynamic text, scroll or static, so the number of barrage in the same video playing video at the same time is the number of barrage on the same screen. The rolling time and stationary time of each barrage of Bilibili are 7s, and it is found by test that there are few cases when the number of barrages of the same screen is more than 50 [14]. Literature [15] drew the “time-barrage” polyline graph with coordinates (t_j, d_{i,t_j}) at an interval of 5s. This paper firstly counted the number of barrages d_{i,t_j} of each video, found the time point t_j of the barrage number $d_{i,t_j} > \lambda$ in video v_i , and then cut video segment f_i according to the time point t_j , where λ is simply computed in Eq. (2).

$$\lambda = \frac{1}{|T_{v_i}|} \sum_{t_j=1}^{T_{v_i}} d_{i,t_j} \tag{2}$$

Considering the interference at the beginning and end of video, it is necessary to remove the number of barrages in some times, such as when music plays at the start and end stage. The key frame extraction algorithm based on k-means clustering method is used to extract the main frames from video fragments and select the video frames that can fully represent video characteristics.

4.2 RCNN Model

The RCNN model is composed of convolutional neural network (CNN) and recursive neural network (Elman recursive neural network). The network structure is shown in Fig. 2.

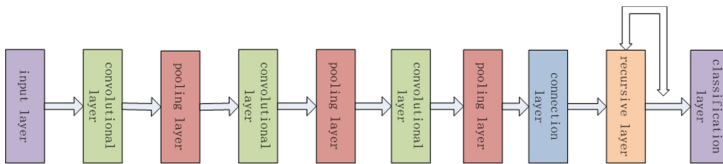


Fig. 2. RCNN model

CNN has three convolution layers and three pooling layers. After the network passes through the full connection layer, it does not go directly to the softmax function layer (classification layer) of the convolution layer, but directly to the added recursive neural network layer. Relu function is used as nonlinear mapping function in the network.

The construction steps of video recommendation model based on recursive convolutional neural network proposed in this paper are as follows:

Step 1: extract the necessary information in the spatial features of video frames that have been selected to represent video characteristics by using the convolutional neural network in the RCNN model. Then, convert the image into low-dimensional feature information, and output the low-dimensional feature information. The calculation form of convolution layer is generally shown in Eq. (3).

$$S_j^l = f \left(\sum_{i \in M_j} Kernel_{ij}^l \times S_j^{l-1} + b_j^l \right) \quad (3)$$

Where, the S_j^l is the result obtained by convolution at l layer, $f(\circ)$ is an activation function, S_j^{l-1} is the output feature of the previous layer, $Kernel_{ij}^l$ is the convolution kernel matrix, b_j^l is the bias of the feature graph after convolution, M_j is the number of feature inputs, that how many image features are selected as the input of the convolution layer.

Step 2: Input the low-dimensional feature information into the RNN model in the RCNN model. The output relation between input layer, hidden layer, connection layer and output layer of RNN network is calculated by Eqs. (4, 5 and 6).

$$v_i(k) = \sum_{j=1}^n w_{ij}^E(k-1) s_j^E(k) + w_i^u(k) u(k) \quad (4)$$

$$s_i(k) = f(v_i) \quad (5)$$

$$s_j^E(k) = s_j(k-1) \quad (6)$$

The output of the network $D(k)$ is calculated by Eq. (7).

$$D(k) = \sum_{j=1}^n w_i^D(k-1) s_i(k) \quad (7)$$

Where $v_i(k)$ represents all inputs of the i -th hidden layer unit, and $f(\circ)$ represents the activation function. w_{ij}^E, w_i^u, w_i^D represents the weights from the input layer to the hidden layer, the connection layer to the hidden layer, and the hidden layer to the output layer.

Step 3: according to the parameters set, the loss function model σ_j in this paper uses the error back propagation algorithm for the partial derivatives of the weight and bias of each layer in the network. Then update the parameters, and at the same time adjust the feedback weights of Elman respectively. When the model reaches the maximum number μ of iterations or loss function within a reasonable range, stop the training, or return to the second step.

Step 4: after ending the model train, get the feature vectors of specific video that users interests. According to these feature vectors, sort video similarity by descending. Finally, the users can be recommended some projects.

In summary, the recursive neural network recommendation algorithm based on barrage is described as Algorithm 1.

Algorithm 1. A recursive neural network recommendation algorithm based on barrage

Input :
video training set S_{train} , video test set S_{test}

Output :
video recommendation set D

- 1 : *for each video S_{train} do*
- 2 : *calculate S_j^l*
- 3 : *calculate σ_j*
- 4 : *calculate w_{ij}*
- 5 : *if $\sigma_j < \mu$ then*
- 6 : *update w_{ij} and b_j^l*
- 7 : *end if*
- 8 : *calculate $s_i(k)$*
- 9 : *calculate $s_j^E(k)$*
- 10 : *calculate σ_i*
- 11 : *if $\sigma_i < \mu$ then*
- 12 : *update w_{ij}^E, w_i^u, w_i^D*
- 13 : *end if*
- 14 : *end for*
- 15 : *return $D(k)$*

5 Experimental Results

This section briefly introduces the data sources and experimental environment used in the experiment. Video of this experiment is downloaded from the domestic barrage website “bilibili” by “jjji” Down software. From the classic video set to choose multiple types of video, such as the modern comedy-Home With Kids, ancient costume comedy-Bronze teeth Ji Xiaolan, humanities war drama-Bright sword, historical costume drama-Youth bao zheng, action comedy-World for the Monkey King, martial arts love drama-The Heaven Sword and Dragon Saber By Jin Yong respectively. We choose the first 5 episodes of each drama, so the total number of episodes is 30. Since bilibili.com maintains the barrage pool and limits the amount of barrage data, so video with more than 20 min has about

3,000 barrages in per episode, and video with more than 40 min has 6,000–8,000 barrages in per episode. The results is shown in Fig. 3, where <p = “when the barrage appears at video (second), the mode of barrage (such as 1..3 denotes rolling barrage, 4 denotes bottom barrage, 5 denotes top barrage, 6 denotes the reverse barrage, 7 denotes the specified bits, 8 denotes the advanced barrage), font size (px), font color (HTML color (decimal)), the generating time of barrage (Unix format), barrage pool, the sender id of barrage, barrage id in barrage database”> barrage content.

```
<chatid>3473629</chatid>
<mission>0</mission>
<maxlimit>6000</maxlimit>
<state>0</state>
<real_name>0</real_name>
<source>e</source>
<d p="466.52100,1.25,16777215,1537589892,0,eed5d550,5477613164298244">真是年轻啊</d>
<d p="695.31600,1.25,16777215,1537590121,0,eed5d550,5477733165432832">成吉思不会从</d>
<d p="758.83500,1.25,16777215,1537590185,0,eed5d550,5477766523256832">本来就是少女</d>
<d p="786.32700,1.25,16777215,1537590212,0,eed5d550,5477780895039488">和牌大头</d>
<d p="953.97900,1.25,16777215,1537590380,0,eed5d550,5477868729008128">真会玩马屁</d>
<d p="26.35200,1.25,16777215,1537590839,0,51671930,5478109172727808">真的停不下来 本来只打算看一集</d>
<d p="863.71700,1.25,16777215,1537593776,0,51671930,5479649300185090">和大人你都不是回去复习了哈哈哈哈</d>
<d p="930.55300,1.25,16777215,1537593861,0,51671930,5479693672775680">发姐汪斐高富帅</d>
<d p="1026.55900,1.25,16777215,1537593971,0,51671930,5479751557840898">和牌的功能不赖呢</d>
<d p="2385.80300,1.25,16777215,153760030,0,37380528,5487122667012100">撒子短很多</d>
<d p="1517.79900,1.25,16777215,1537613903,0,7186a5fe,5490201565593604">ooo</d>
<d p="1279.42500,1.25,16777215,1537614966,0,bf745471,5490758836551680">四姑娘，我看你是想做皇后啊</d>
<d p="1811.68900,1.25,16776960,1537616845,0,40ebd4fc,5491744182370308">任乎！</d>
<d p="1576.88700,1.25,16777215,1537619719,0,3cb9249f,5493250977169408">我才来本球</d>
<d p="1832.32400,1.25,16777215,1537620005,0,3cb9249f,5493400705433600">怎么感觉文豪语文首都是从名字开始的</d>
<d p="2003.00300,1.25,16777215,1537620192,0,3cb9249f,5493499065532416">老板好淡定</d>
<d p="2596.63900,1.25,16777215,1537620809,0,3cb9249f,5493822558044160">教你写字的时候你就写错了</d>
<d p="1855.50100,1.25,15338834,1537624402,0,b3e2015b,5495706352615424">熊</d>
```

Fig. 3. Barrage-list

By python program, it crawls the sending time, content and sender id of the barrage. Then, the barrage data is processed, and is visualized, as shown in Fig. 4.

Barrage transmission time	Barrage sender id	Barrage content
00:00:50	9.51E+07	oh亲爱的上帝 快看看我发现了什莫
00:00:54	1699353	下次整
00:00:54	65c0f76ab	我叫夏雷, 我叫夏雷, 我叫下冰雹
00:00:54	5baf1ca2	小紫霞!
00:00:55	b71fc21a	b站竟然有这个
00:00:56	849f240f	夏东海秒变刘新建
00:00:56	9acff6d	从组家庭
00:00:57	e89c7df3	合影!
00:00:57	211c0548	合影
00:00:58	552c84c2	合影
00:00:59	1eb025be	合影
00:00:59	d8f0a1b	好暖和
00:00:59	b8nda1cd	小雷类的勤
00:00:59	b749017e	居然已经15年了
00:00:59	c86a59fc	合影
00:00:59	66de1888	合影!
00:01:00	407b7aca	合影!
00:01:00	70bb9da8	合影
00:01:00	15818157	合影
00:01:00	14553631	合影!
00:01:01	45ec9115	合影
00:01:01	2a01391f	合影!
00:01:01	b693a27d	合影! !
00:01:01	74238a05	希望所有二婚家庭都这么美满鸭!
00:01:01	260e5012	合影
00:01:01	e533297c	合影
00:01:01	d8992e63	合影
00:01:01	2ab2ef26	合影
00:01:01	b3fbcc72	合影
00:01:01	6c1e4f9a	合影

Fig. 4. The list of processed barrage

In this paper, the number of barrage removed the barrage at the beginning and end of video, reducing the noise barrage, as shown in Fig. 5.

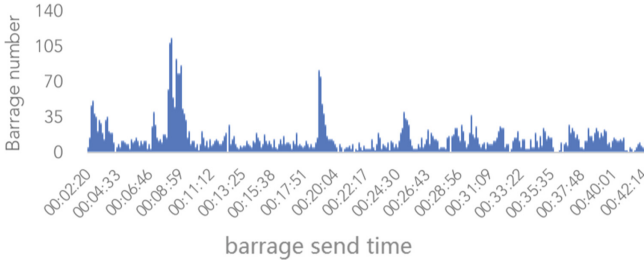


Fig. 5. The changing number of barrages

Besides data preprocessing in this paper, we set the training model and some parameters, as follows:

- (1) The operating system is Windows 7, Intel core i5 processor and 16 GB RAM.
- (2) The programming environment is python3.0, and the keras library environment is built. In order to train the weight of RCNN, the random gradient descending algorithm is used.
- (3) In order to evaluate the overall performance of the model, the data set was randomly divided into a training set(80 and the Mean Absolute Deviation (MAE) [16] is adopted as the evaluation merit, as shown in Eq. (8).

$$M_{MAE} = \frac{\sum_{i=1}^N |p_i - q_i|}{N} \tag{8}$$

Where N is the number of recommended video, p_i is the predicted results of the experiment, and q_i is the actual results. MAE reflects the difference between the predicted value and the actual value of the algorithm. The smaller MAE value is, the more accurate the recommended algorithm is. In order to verify the effectiveness of this model, the proposed algorithm of this paper is compared with collaborative filtering thematic regression algorithm (CTR) [17], collaborative filtering deep learning algorithm (CDL) [18] and convolutional matrix decomposition algorithm (ConvMF) [19] on the same data set of this paper.

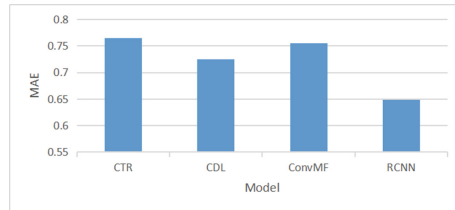


Fig. 6. The experiment results of model comparison

This paper uses the data set crawled from the bilibili website, In order to reduce herd effect in video barrage, we preprocess the data, the experimental results obtained in the experiment test set are shown in Fig. 6.

From the figure, we can see that the performance of proposed method compared with the CTR, CDL, ConvMF respectively increases by 0.22, 0.18 and 0.31. From the comparison results in this data set, the proposed method in this paper owns higher accuracy.

6 Conclusion

In order to improve the accuracy rate of the recommendation model, this paper proposes a video recommendation method that adopts the convolutional recursive neural network based on barrage. The barrage video can quickly capture the preference features of users by barrages, and improve the overall recommendation performance of prediction model with the help of convolutional recursive neural network. Experimental results show that the proposed convolutional recursive neural network recommendation method based on barrage can increase the selectivity of users. However, video without motion features has a weak recognition ability. Therefore, the next research based on the video key frame is to add video content-related descriptive statements, and improve the recursive neural network [20–22] for further improving the accuracy of the prediction model.

References

1. Ma, S., He, P.: Review of recommended methods for application. *Comput. Sci. Appl.* **9**(7), 1317–1327 (2019)
2. Deng, Y., Zhang, C., Li, J.: Video fragment recommendation model based on barrage emotion analysis. *Comput. Appl.* **37**(04), 1065–1070+1134 (2017)
3. Hong, Q., Wang, S., Zhao, J., Li, J., Rao, W.: Video user group classification based on barrage emotion analysis and clustering algorithm. *Comput. Eng. Sci.* **40**(06), 1125–1139 (2018)
4. Kong, X., Su, B., Wang, H., Gao, H., Li, J.: Research on recommendation model and algorithm based on tag weight score. *J. Comput. Sci.* **40**(06), 1440–1452 (2017)
5. Lin, X., Sang, Y., Long, C.: Personalized recommendation based on user decision mechanism. *Libr. Inf. Work* **63**(02), 99–106 (2019)
6. Tzamouris, E., Papadopouli, M.: On hybrid modular recommendation systems for video streaming (2019)
7. Nguyen, T.B., Aihara, K., Takasu, A.: A probabilistic model for collaborative filtering with implicit and explicit feedback data (2017)
8. Zhang, F., Zhang, L., Luo, T., Wu, Y.: A feature-based collaborative clustering model. *Comput. Res. Dev.* **55**(07), 1508–1524 (2018)
9. Zhang, W., Li, J., Yang, J.: An improved information kernel extraction method in collaborative filtering recommendation. *Comput. Appl. Res.* 1–6 (2019)
10. Li, Y., Wang, H., Liu, H., Chen, B.: A study on content-based video recommendation. In: *IEEE International Conference on Image Processing*, Beijing, 17–20 September 2017, pp. 4581–4585 (2017)

11. Li, N., Sheng, Y., Ni, H.: Conditional convolution implicit factor model for personalized recommendation [J/OL]. *Comput. Eng.* 1–10 (2019)
12. Cai, N., Liu, G., Cai, H.: A recommended model for the combination of improved matrix decomposition and convolutional neural network [J/OL]. *Comput. Eng. Appl.* 1–8 (2019)
13. Wang, H., Dong, M.: Potential group recommendation based on dynamic convolution probability matrix decomposition. *Comput. Res. Dev.* **54**(08), 1853–1863 (2017)
14. Liu, S.: Research on human-computer interaction efficiency of modern web barrage. Sun Yat-sen University (2015)
15. Guo, K.: Research on barrage data mining algorithm based on quantitative characteristics. Huazhong University of Science and Technology (2017)
16. Zhou, J., Huo, H.: Multi-dimensional video recommendation algorithm in social network service. *Comput. Eng.* **41**(01), 245–250 (2015)
17. Zhang, M., Ding, B.-Y., Ma, W.-Z., et al.: Mixed recommendation method based on deep learning enhancement. *J. Tsinghua Univ. (Nat. Sci. Ed.)* (10), 9–16 (2017)
18. Wang, C., Blei, D.M.: Collaborative topic modeling for recommending scientific articles. In: *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 448–456. ACM (2011)
19. Kim, D., Park, C., Oh, J., et al.: Convolutional matrix factorization for document context-aware recommendation. In: *ACM Conference on Recommender Systems*. ACM (2016)
20. Xiong, M.T., Feng, Y., Wu, T., Shang, J.X., Qiang, B.H., Wang, Y.N.: TDCTFIC: a novel recommendation framework fusing temporal dynamics, CNN-based text features and item correlation. *IEICE Trans. Inf. Syst.* **E102.D**(8), 1517–1525 (2019)
21. Liu, E., Chu, Y., Luan, L., Li, G., Wang, Z.: Mixing-RNN: a recommendation algorithm based on recurrent neural network. In: Douligieris, C., Karagiannis, D., Apostolou, D. (eds.) *KSEM 2019. LNCS (LNAI)*, vol. 11775, pp. 109–117. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29551-6_10
22. Hsieh, C.K., Campo, M., Taliyan, A., et al.: Convolutional collaborative filter network for video based recommendation systems (2018)