A Personalized Video Recommendation Algorithm Based on Complex Network

Shuxia Pang[*], Weifang Wang and Honglei Zhu

School of Computer and Communication, Lanzhou University of Technology, Lanzhou, China

[*]Corresponding author

Abstract—Since the current world is overwhelmed by internet, it becomes important and necessary to find the personalized video information to meet the growing demand for all kinds of users. Video recommendation is a very useful method for getting video data quickly. A personalized video recommendation algorithm based on complex network is presented in the paper, which can recommend the video users interest. Once the users give some features of the video information, taking the IMDB film as an example, they will get what they want. The experimental results confirm the effectiveness of this method.

Keywords—personalized; video recommendation; complex network

I. INTRODUCTION

In recent years, there has been a tremendous need to query and process large amount of data that cannot be easily described especially video data. It is very an urgent and essential to solve some problems that how to get the personalized video information. After that video recommendation emerges [1].

Text-based and content-based methods are considered as two fundamental methods for video recommendation. Text-based video recommendation is a process of searching the video by keyword based on the text of the video annotation. It adopts manual methods or computer auxiliary tool to add some text or tags to digital video image to find the needed information through query keywords. The main drawbacks of this method are as follows: the first is the keyword need to create text for the video, with the subjectivity and arbitrariness, different people or the same person might make different comments on the same video in different conditions, so it cannot objectively and accurately represent the video content, affecting the retrieval results [2]; the second is the manual annotation of video, with high cost and huge workload, need to browse and note the massive video data ,which is a very difficult thing; the third is, because of the content of video data rich and colorful, it is very difficult to carry on accurate and comprehensive summary only by some simple text.

Content-based video recommendation [3] allows users to give a picture and then search for video data that has the same or similar content according to the content of the picture. It mainly adopts the effective analysis of the scene, shot and image frames to extract the video image color, texture, shape, motion and other features, and then finds the video clips with maximum similarity and the most satisfactory from users by means of similarity match method. Content-based video includes generally shot boundary detection, key frame extraction, feature extraction and similarity matching.

Content-based video recommendation can meet the need of video retrieval to some extent, but it cannot meet the personalized demand with the different backward and different purpose. And so the personalized method is proposed, it can meet different users with different service, learning from the user's interesting and behavior by collecting and analyzing information from users. Since it can recommend video actively, the personalized recommendation can attract more visitors with the excellent quality and efficiency.

In the paper, the personalized video recommendation is proposed based on complex network. It can recommend quickly the most interesting video according to the sample video from users. And mean while it is also the content-based recommendation. The study takes IMDB as the example to demonstrate the result.

II. COMPLEX NETWORK

At the end of the twentieth Century, the complex network theory began to expand from the original mathematics. People began to consider the overall characteristics of the real network with large number of nodes and complex connection structures, thus raising the upsurge of studying complex networks, which is called "the new science of network".

In recent years, complex networks have developed rapidly, and the network structure has developed hundreds of hundreds of nodes to millions or even more. The changes in the network scale also promote the corresponding changes in the network analysis method.

In general, the main contents of the complex network theory are as follows:

(1) Discovery: the interpretation and characterization of the statistical properties of the structure of the system and the methods to measure these properties.

(2) Modeling: the establishment of a suitable network model to help people understand the significance and mechanism of these statistical properties.

(3) Analysis: analysis and prediction of network behavior based on the characteristics of a single node and the structural properties of the entire network.

(4) Control: proposes to improve the performance of the existing network and design new effective methods.

A complex network [4] is represented by a graph consisting of a set of vertices and edges. The complex network can be described by its adjacency matrix for undirected and unweighted networks in this study. The adjacency matrix is a symmetric matrix with the elements consisted by zero and one, where one represents that two nodes are linked by an edge and zero means there is no edges between two nodes. Each complex network model has its specific topological features that describe the characteristics of the complex network. Therefore, the discrimination and analysis of the complex networks rely on the measurements that can express the most relevant topological features and their subsequent classification.

III. Key Technologies of Video Recommendation

Video recommendation is mainly divided into the following modules:

(1) Acquiring video data, save data to the database; (2) Shot boundary detection. According to the correlation between the adjacent frames, the video is automatically divided into a series of independent lenses; (3) Key frame extraction. After the lens segmentation is completed, the corresponding key frames are selected according to the content of the lens; (4) Feature extraction. The features of key frames that meet the recommended requirements are stored in the feature library; (5) Feature matching. The same features are extracted from the sample query image of user input to match approximately with the features in the database, and to find the most similar video data; (6) Output results. The maximum similar video clips are returned to users.

A. Shot Segmentation

Video stream is made up of thousands of image frames. Frames are the smallest units of video. If each frame is processed separately, the efficiency of index and retrieval will be very low. Video is usually made up of many video shots. A shot is a short sequence of adjacent frames, which depict the same scene, representing an action of a camera, an event or a continuous action. In the index processing of video data, the video is first automatically segmented into a lens. When the

content of the video plot changes, there will be a shot switch. Usually, the switching modes between shots can be divided into two major categories: abrupt and gradual transition. The abrupt is a transition between two shots. It jumps directly from one scene to the next. The gradual refers to a gradual transition from one shot to another. There is no obvious skip, including fade in and fade out, dissolution and scrubbing. The shot segmentation of video is not only to detect the mutation between the lenses, but also to segment the gradient.

B. Extracting Key Frames

Key frames [6] are the representative frames that express the content of the shots. They are several static images that can be extracted from the video data and can generalize the content of the shots. Therefore, the key frame extraction is the key to establish the index of video data. It not only solves how to extract key frames that can represent the content of shots from the frame sequence of shots, but also precipitates the contents and categories from key frames.

X^2 histogram[7] is adopted in the study to extract key frames. The distance $d(I_i, J_j)$ between video images is calculated by the formula (1)

$$d(I_i, J_j) = \sum_{k=1}^{n} \frac{(H_i(K) - H_j(K))^2}{H_j(K)} \tag{1}$$

And I_i, J_j are two different video frames, $H(K)$ is the entropy of quantized video image, which is as formula (2)

$$H(K) = -\sum_{i=0}^{l-1} p(k_i) \log p(k_i) \tag{2}$$

$p(k)$ is probability density function of random variable k, k is one of the features grayscale levels, regional gray scale and gradient. Color is selected as the histogram feature. When to find the key frames, set a threshold T, while the distance d is greater than the threshold, the frame is extracted as the key frame to represent the main content of the video, otherwise the distance with other frames will carry over. The process of video recommendation is figured as figure I.
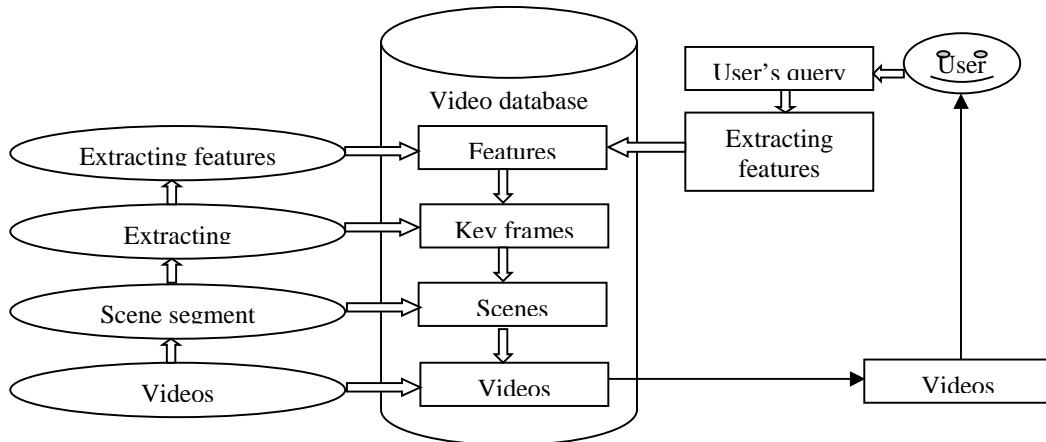

FIGURE I. PERSONALIZED RECOMMENDATION

IV. CONSTRUCTING VIDEO NET BASED ON COMPLEX NETWORK

A. Extracting Information Nodes of the Network

The nodes of complex networks form the research objects of complex networks, and the nodes of networks are made up of named video entities. Therefore, network point information extraction is an essential step to build complex networks. The essence of network point information extraction is to find out and identify the related named videos from a large amount of information on the Internet. The node of video network based on complex network is a video entity, and the node extraction of the video network is to extract the video information from the Internet.

Video is defined as a metadata of four tuples [8], including the name and attribute of the video, the keywords to describe the video and the video reviews. The attribute includes the director, leading actors, scenarist, type and other interested. So the acquisition of video entities includes the name of the video, the director, leading actors, scenarist, the type, the keywords and the video reviews.

B. Extracting Edge Information of the Network

The video entity constitutes the node of the complex network, and the relationship between the video entities should represent the edge of the video network node. Similar to video node extraction, edge is defined in advance to the similarity relation between videos, by calculating the similarity between videos, the video similarity value as the edge value. The network is shown as figure II.

In the system, the analysis of low level features is put aside, from the video name, leading actor, video type and video keyword to analyze the high-level semantic information of video. The similarity is calculated by the method in reference [9].
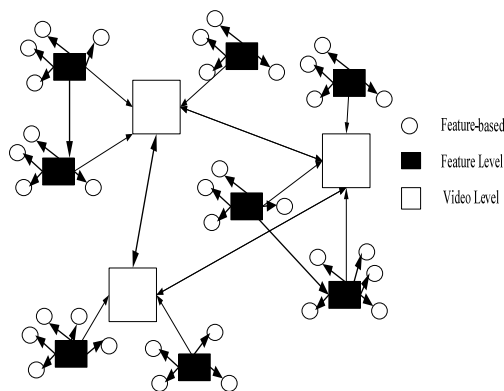

FIGURE II. THE NETWORK OF VIDEO

V. PERSONALIZED VIDEO RECOMMENDATION

In order to highlight the personalized video recommendation, we will execute second clustering [10] with video features, we take IMDB as the example, according to the online statistics, most users will select the film they interested

by the director, leading actor, the type of film or release time, so the four is used as the main attributes to find the similar film. Assuming there are n films, each film has these four properties, and it can form the film matrix as follows.

$$
v = \begin{pmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ ... & & & \\ f_{n1} & f_{n2} & f_{n3} & f_{n4} \end{pmatrix} \tag{3}
$$

And, $f_{ij}(i=1,2,…n, j=1,2,3,4)$ represents the four properties of the i-th film.

(1) The film matrix is normalized, and the data is compressed to the [0,1] closed interval by formula (4).

$$
f* = \frac{f_{ij} - \min\limits_{1<i<n} f_{ij}}{\max\limits_{1<i<n} f_{ij} - \min\limits_{1<i<n} f_{ij}} \tag{4}
$$

(2) Constructing similarity matrix S.

(3) Iterative updating attraction matrix R and ascription matrix a by formula (5),(6) and (7)

$$
r(i,k) \leftarrow s(i,k) - \max[a(i,k) + s(i,k)] \tag{5}
$$

If $i \neq k$:

$$
a(i,k) \leftarrow \min\{0, r(k,k) + \sum_{i'(i'\notin(i,k))} \max\{0, r(i',k)\} \tag{6}
$$

$$
a(k,k) \leftarrow \sum_{i'(i'\neq k)} \max\{0, r(i',k)\} \tag{7}
$$

Formula (5) renews attraction matrix R with similarity matrix S and ascription matrix a, and formula (6) renews ascription matrix a with attraction matrix R.

(4) Determining whether the default number of iterations is reached. If it is true, the fifth will be executed, otherwise measuring whether $a(i,i)+r(i,i)$ is greater than zero, and if false, the third and the forth will continue, otherwise the fifth will be executed.

(5) Finishing iterative update and selecting cluster center and then output clustering set.

Experiment is carried out in IMDB, all movies is featured as the director, leading actor, the type of film or release time.

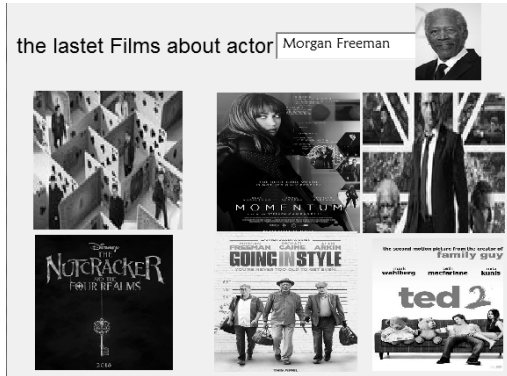Feature from Users is the actor Morgon Freeman, figure III is the recommendation.

FIGURE III. ABOUT ACTOR MORGAN FREEMAN

If users want to know the details about one of the recommendation, clicking the icon and the interface as figure IV will appear.
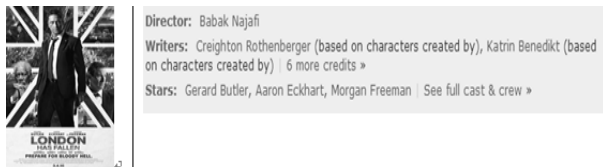

FIGURE IV. DETAILS ABOUT THE MOVIE LONDON HAS FALLEN

VI. CONCLUSION

Since there has been a tremendous need to query and to find some interesting video from video data on internet, it is very essential to solve some problems that how to get the personalized video information. In the paper, complex network-based method is proposed, taking IMDB as the example, firstly, constructing film network, including the information about the name and attribute of the video, the keywords to describe the video and the video reviews. And then extracting the feature information from the sample video given by users, some similar video information will be recommended for users. Experiments show that the result performs well in effectiveness.

REFERENCES

[1] H. Bai, "Research on key technology of content based video retrieval," Taiyuan University of Technology,2017..

[2] H.Asiyan, H.Abudurexiti, "A comparative study of text based image retrieval and content based image retrieval," Journal of Capital Normal University (Natural Science Edition), vol.04, 2012,pp.6-9..

[3] M.S.Lew, N. Sebe, C.Djeraba, "Content-based multimedia information retrieval:State of the art and challenges," ACM Transactions on Multimedia Computing Communications & Applications, vol.2, 2006, pp.1-19..

[4] D.Xu,"Complex networks-based texture extraction and classification method for mineral flotation froth images," Minerals Engineering, vol. 83,2015, pp. 105-116.

[5] Q. Li, "Research on complex network-based video network constructing measurement," East China normal university,2012..

[6] X.G. Pu, "A Survey on Key Techniques of Content-based Video Retrieval," Information Science, vol. 3,2010,pp.464-469

[7] S.Memar,L.S.Affendey, "An integrated semantic-based approach in concept based video retrieval," Multimed Tools Applicatioin ,vol.64, 2013,pp.77–95.

[8] S.Jaiswal, A. L.Rosenberg, "comparing the structure of power-law graphs and the Internet AS graph," 12th IEEE International Conference on network protocols,2004,pp.294-303.

[9] Y.F Yu, "Research on the method of film similarity calculation with multi feature fusion ,",2012.

[10] L.Jin,J.Yu, "Video recommendation algorithm based on clustering and hierarchical model," Journal of Computer Applications,vol.10,2017, pp.2828-2833.