CrossMark

REGULAR PAPER

Social-group-based ranking algorithms for cold-start video recommendation

Chunfeng Yang[1] · **Yipeng Zhou**[2] · **Liang Chen**[2] · **Xiaopeng Zhang**[3] ·
Dah Ming Chiu[1]

Abstract Video recommendation has become an essential part of online video services. Cold start, a problem relatively common in the practical online video recommendation service, occurs when the user who needs video recommendation has no viewing history. (Cold start consists of the new-user problem and the new-item problem. In this paper, we discuss the new-user one.) A promising approach to resolve this problem is to capitalize on information in online social networks (OSNs): Videos viewed by a user's friend may be good candidates for recommendation. However, in practice, this information is also quite limited, either because of insufficient friends or lack of abundant viewing history of friends. In this work, we utilize social groups with richer information to recommend videos. It is common that users may be affiliated with multiple groups in OSNs. Through members within the same group, we can reach a considerably larger set of users and hence more candidate videos for recommendation. In this paper, by collaborating with Tencent Video, we propose a social-group-based algorithm to produce personalized video recommendations by ranking candidate videos from the groups a user is affiliated with. This algorithm was implemented and tested in the Tencent Video service system. Compared with two state-of-the-art methods, the proposed

algorithm not only improves the click-through rate, but also recommends more diverse videos.

1 Introduction

Video recommendation has become an integral part of today's online video services, such as those provided by Netflix and YouTube. Good recommendation not only increases user engagement, but also improves user loyalty. Although the problem of recommending videos on the basis of users' viewing history has been well studied, it is still challenging to recommend videos for those users with little or no viewing history. In the literature, this is known as the data sparsity or cold-start problem.[1] Common recommendation strategies can be categorized into three types: collaborative filtering (CF), content-based filtering (CBF) and hybrid strategies combining CF and CBF [1].

CF [28], based on a user's preference and behaviors of other users with similar preferences, can accurately recommend videos of interest given sufficient historical records. It is therefore widely used in online video systems. However, as reported in [28], CF is not effective for users with little viewing history. CBF [21] is based on clustering items with similar descriptions and matching them to users' current selection. This strategy can exploit various advanced information retrieval techniques. However, a major weakness of the strategy is over-specification (keep on recommend-

✉ Chunfeng Yang
   yc012@ie.cuhk.edu.hk; yangchunfeng6731700@gmail.com

1 The Chinese University of Hong Kong, Hong Kong, China

2 Shenzhen University, Shenzhen, China

3 Precise Recommendation Center, Tencent Inc., Shenzhen, China

---

[1] Cold start can be seen as the extreme case of data sparsity. This study refers cold start to situations where available collaborative information is inadequate for effective recommendation.

ing items of the same type or with similar descriptions). The hybrid system combines the strengths of CF and CBF to overcome the cold-start and over-specification problems; however, its advantage over the other two strategies has so far been marginal [35].

An alternative strategy for solving the cold-start problem is to use social information [29]. By exploiting online social networks, videos viewed by a user's friend can be used for recommendation. Although this strategy is quite promising, it is not always effective because social information could be scarce for some users. In this study, instead of exploiting social information associated only with friends, we used social-group-based information for video recommendation. Our collaborator (Tencent Inc.) runs an online platform that provides multiple services, including online games, online videos and instant messaging (QQ), and facilitates the formation of QQ groups by QQ users. QQ groups allow users to easily communicate within a small circle of typically 50 to 100 users, sharing common interests; for example, there are classmate groups, colleague groups and interest groups of various types. We find that group affiliation is quite prevalent, and a user is typically affiliated with multiple groups. If we assume that a video viewed by a group mate is of potential interest, then QQ groups can provide considerably more candidate videos for recommendation compared with the number of candidate videos obtained by considering only the circle of friends.

Although QQ groups substantially increase the pool of candidate videos available for recommendation, the relevance of these videos depends strongly on the type of group the videos originate from. We propose an algorithm for ranking candidate videos from different groups and identifying the top R videos for recommendation.[2] This is done in two steps. First, videos from a single group are ranked. If this is the only group the user is affiliated with, then this ranking provides the order for recommendation. Otherwise, the ranked videos from multiple groups are (weighted) aggregated to obtain the final order for recommendation, where the weight of each group is learned by a supervised learning method based on several calculated group features.

Our objective was to recommend videos with a high hit rate and of high diversity because both these metrics are crucial to user satisfaction. To the best of our knowledge, this is the first study to develop a video recommendation method based on different group affiliations and merely implicit feedback data,[3] and the method was tested in an online video system. Our contributions are summarized as follows.

- We determined and analyzed the difference between the number of candidate videos obtained from social groups and that acquired from only a friend circle.
- We proposed a video recommendation method based on social group information; the method can rank candidate videos from a single group as well as multiple groups that a user is affiliated with.
- We evaluated the proposed social-group-based video recommendation algorithm by implementing it on the Tencent Video system and showed that it improved both click-through rate and video diversity.
- We further investigated the advantages of collaborative filtering and the proposed social-group-based algorithm in terms of click-through rate and analyzed the potential hybrid switching strategy.

The rest of this paper is organized as follows. We first discuss related works in Sect. 2 for identifying research gaps. We then describe (in Sect. 3) the studied system and demonstrate how the social-group-based approach can considerably increase the number of candidate videos, as the motivation of the social-group-based strategy. The ranking methodology and algorithms are detailed in Sect. 4. In Sect. 5, experimental results are shown, and further investigation of a hybrid strategy is discussed. Finally, Sect. 6 concludes the paper.

2 Related works

Collaborative filtering (CF) is a well-developed framework utilizing viewing history of different users to provide personalized recommendation. A detailed survey of collaborative filtering is provided by Su and Khoshgoftaar [28], and various enhancements have also been developed [16,32]. The main drawback of collaborative filtering is the cold-start problem, which has been discussed and studied in [5,36].

When we do not have users' behavioral data, such as for new users, users' interests can only be inferred based on other user information. Recent studies start to leverage social cues to enhance user interest modeling. Trust is shown to be positively correlated with interest similarity [37]. Besides, people tend to befriend others who share similar traits, known as homophily in sociology [19]. Wen and Lin [31] deduce user's interests by considering this user's social neighbors' interests. Also interests could be inferred to some extent from the users who share more demographic attributes [17].

As a promising approach to overcome the cold-start problem, social recommendation which capitalizes on social information has become more popular. Existing studies of social recommendation can be grouped into three types: (1) Extra social information is used to improve an existing recommendation system [13,14,22], for example, SCF [20] is a social collaborative filtering method. Unlike the traditional

---

[2] R is the number of videos recommended to a user, which depends on the system setting.

[3] Implicit feedback means only user viewing records (view or not) without explicit numerical rating data.

user-based collaborative filtering (CF) that considers the top-k similar users, SCF generates predictions based on users' direct friends; (2) social information is used to create or enable a recommender system [8]; and (3) studying user trust and item reputation [31]. Despite the popularity of social recommender systems, Tang et al. [29] points out that there are still some negative experiences in applying social recommender systems: (a) Social relations are too noisy and may have negative impact on recommender systems; (b) for cold-start users, they may also have few social relations; (c) different types of social relations may affect social recommender systems differently, and the success of one type of social relation may not be applicable to others.

While social recommender systems utilize social information of friends, others turn to communities of users for recommendation. For example, Sahebi and Cohen [24] utilize community detection approaches to find communities from different dimensions of social networks and then perform collaborative filtering within community members. Yang et al. [34] develop a circle-based recommender system that infers category-specific social trust circles from the available rating data combined with social network data, and it has also proposed several variants to weight friends based on their inferred expertise levels.

Different from those community-based approaches, which generate groups by virtue of community detection on top of interest similarity or the social relationship network, our work is based on explicitly defined groups formed by members autonomously due to common interest and other reasons. Furthermore, in our setting, users can be affiliated with multiple groups which lead to a new challenge of how to exploit these groups together for recommendation. In addition, our work is also different from the study of recommendation to groups, which either tries to construct group profiles through integrating individuals' profiles [15] or directly aggregates individuals' recommendation results [4]. Their goals are to help members arrive at a consensus about which recommendation to accept, either for some items or some group activities, such as restaurants or museum exhibits. In contrast, in our case, users' decisions are made independently as personal choices without compromising with others. Compared with the recommender systems based on social friends, our group-based approach shows its strengths in the following aspects: (1) Effective social relation is often quite limited, either because of insufficient friends or lack of abundant viewing history of friends; (2) moreover, the social-group-based approach is more privacy friendly when providing the explanations to users about the recommendation results.

For the design of recommender systems, there are four factors affecting recipients' advice-taking decisions [2,3], namely cognitive homophily [25], tie strength [18], trustworthiness [27] and social capital [10]. Cognitive homophily means the similarity of user behaviors, taste or interest. Tie strength can measure the intensity (frequency and duration) of the interaction between the recipient and source. Social capital of a source is the source's reputation or opinion leadership [2]. Various types of data collected online correspond to these factors: Profile similarity employed in traditional collaborative filtering quantifies homophily; communication records reflect tie strength, and social relationship data can be utilized to analyze trustworthiness.

3 Analysis of candidate video pool

Tencent Video[4] is one of the largest online video-on-demand (VoD) service providers in China, supporting more than 50 million active users on a daily basis. During peak hours, more than 2 million concurrent users are served. Tencent Video's video catalog includes movie, TV episodes, music video (MVs), news, user-generated content (UGC) and many other types.

The online social network used by us is Tencent QQ.[5] QQ is one of the most popular instant messaging services in China, through which one can make friends, chat with friends and join QQ groups. In June 2015, there are roughly 843 million active QQ accounts, with a peak of 233 million online QQ users [23]. Moreover, most users join multiple QQ groups (more than ten). For most QQ groups, there are more than 50 members.

To help discuss our ideas precisely, we first define some notations. All users, videos and groups are assigned unique IDs. For user i, his viewing record list for the last 30 days is represented by the set $\mathcal{V}_i$, i.e., if a video j is viewed by a user i, then $j \in \mathcal{V}_i$. The set $\mathcal{G}_i$ represents the groups that user i joins. For a group k, all its members are represented by the set $\mathcal{U}_k$. If user u is a group mate of user i, there exists a group k, such that $u, i \in \mathcal{U}_k$. The set $\mathcal{G}$ consists of all the groups, and the total number of groups is G. For a group k, the video pool, i.e., the set of all videos that have been viewed by any user within the group, is represented by $\mathcal{P}_k = \cup_{i \in \mathcal{U}_k} \mathcal{V}_i$. The group video candidate pool for user i is $\mathcal{P}_{\mathcal{G}_i} = \cup_{k \in \mathcal{G}_i} \mathcal{P}_k$. The set of friends for user i is $\mathcal{F}_i$, i.e., if user u is one of user i's friends, then $u \in \mathcal{F}_i$. The friend video candidate pool for user i is represented by $\mathcal{P}_{\mathcal{F}_i} = \cup_{u \in \mathcal{F}_i} \mathcal{V}_u$. We summarize important notations used in this paper in Table 1 for ease of reference (some of them will be described later).

Currently, recommender systems based on collaborative filtering are already deployed in Tencent Video to provide personalized recommendation. One primary challenge faced by the system is the new-user cold-start issue and data sparsity issue. Take the recommender system for movies as an example. According to system measurements, around 25% of the daily users have not watched any movie in Tencent

---

[4] http://v.qq.com/.

[5] http://www.qq.com/.

Table 1 Notations

$\mathcal{V}_i$	video records of user i	$\mathcal{G}_i$	groups joined by user i
$\mathcal{U}_k$	members of group k	G	total number of groups
$\mathcal{P}_k$	video pool of group k	$\mathcal{F}_i$	friends of user i
$\mathcal{P}_{\mathcal{G}_i}$	group video candidate pool of user i		
$\mathcal{P}_{\mathcal{F}_i}$	friend video candidate pool of user i		
$\mathcal{R}_i$	videos recommended to user i		
R	number of videos recommended to a user		
$\eta_{k,j}$	local popularity score of video j in group k		
β_j	representativeness score of group j		
$W_{k,j}$	score of video j in group k		
$\boldsymbol{l_k}$	ranked video list of group k		
S_k^a	social activeness of group k		
S_k^c	social conformity of group k		
I_k^a	interest activeness of group k		
I_k^c	interest conformity of group k		
J_k	group score of group k		
D_i	# of distinct videos in $\mathcal{P}_{\mathcal{G}_i}$		
$p_{\boldsymbol{l_k}}(j)$	the position of video j in list $\boldsymbol{l_k}$		
$W_{i,j}$	final score of video j for user i		
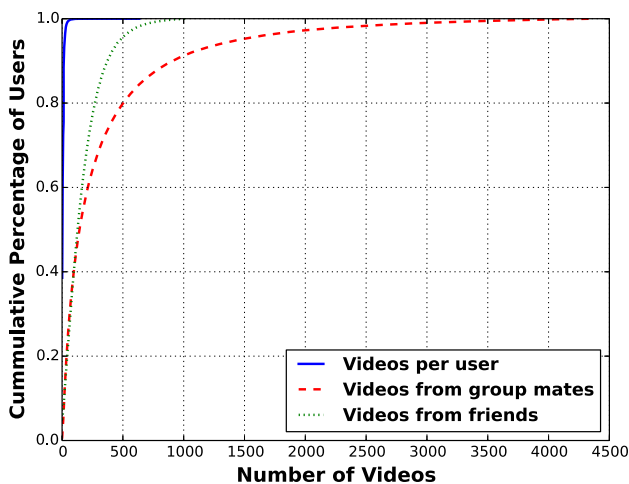
Fig. 1 The cumulative distribution of number of videos from individual users, their friends and group mates

Video in the last 30 days and thus can be treated as new users for a one-month window. In addition, the user viewing behavior is fairly sparse on the individual level, indicating data sparsity, as illustrated by the distribution curve of number of videos per user in Fig. 1. In other words, the nonzero entries in the user-item consumption matrix take less than 0.01 % of the whole matrix.

We can also analyze the candidate video pool for a given user in Tencent Video and Tencent QQ together. For a particular user, his viewing records, his friends' viewing records and his group mates' viewing records are jointly measured over 30 days. The results are summarized in Table 2 and

Fig. 1. The results show that for a user, there are more group mates than friends and more video records from group mates accordingly, which implies we may discover more interesting videos via groups. This motivates us to design the social-group-based recommendation strategy.

4 Ranking algorithms

4.1 Objective and challenges

The problem we want to solve in this paper is to recommend a set of relevant and diverse[6] videos $\mathcal{R}_i$ (with size R) for user i from the group video candidate pool $\mathcal{P}_{\mathcal{G}_i}$. Thus, we need to design algorithms to rank videos in user i's group video candidate pool. The main challenges lie in the availability of only implicit feedback data as well as different cases of group affiliation. Merely with implicit feedback, it is not easy to generate effective recommendation, because there is no negative feedback. For example, the reason a user did not watch a certain video might be because she disliked the video or she did not know the video at all. Moreover, users joined different number of groups, as shown in Fig. 2. Firstly, for users affiliated with a single group, we need an algorithm to rank videos viewed by users in the same group, which

---

[6] In this paper, relevant recommendation means the results conform to users' interest or the recommended items can attract users, which is similar to the accuracy measure. Diversity refers to sales diversity, which measures how unequally different items are chosen by users or recommended to users when a particular recommender system is used.

Table 2 Statistics of friends, group mates and video candidate pools

Per user	Mean value
Number of friends	35.9
Number of group mates	213.2
Size of friend video candidate pool	48.8
Size of group video candidate pool	142
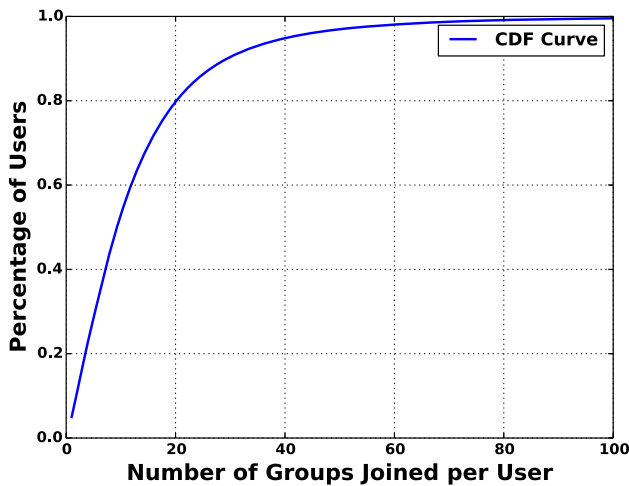
Fig. 2 Cumulative distribution function (CDF) and probability density function (PDF) *curves of the number* of affiliated groups per user

is the *single affiliation problem*. Secondly, if a user is affiliated with multiple groups, we need algorithms to rank videos from different groups, which refers to the *multiple affiliation problem*. In the real system, most users have joined multiple groups; however, we begin with the ranking algorithm in the single affiliation case for ease of clarifying the proposed algorithms.

4.2 Video ranking for single affiliation problem

In this case, we need to resolve the intra-group video ranking problem. Different from previous studies of group profiling [30] and recommendation with rank aggregation [4] that utilize explicit user preference, we only have implicit feedback data. Assume user i only joins group k, then $\mathcal{P}_{\mathcal{G}_i} = \mathcal{P}_k$.

Generally speaking, videos that are more representative of the group should be ranked higher. As argued by Doersch et al. [9], a representative item is supposed to be *frequent* and *discriminative*; in other words, it should be frequent so as to be a "pattern," and it can be used to distinguish one from others. Intuitively, the more the members of a group viewed a video, the more the video is likely to attract other users in the group. However, frequent items are not necessarily discriminative. For example, for a group comprised of sports fans, some very hot videos, such as breaking news, are less

discriminative than a certain sports video, although those hot videos were viewed by more members of this group.

To capture videos that are both frequent and discriminative, we firstly define scores to quantify these two characteristics of a video, respectively. We use a local popularity score to denote how frequently a video appears in the group video candidate pool, i.e., how many group members have viewed the video:

$$\eta_{k,j} = \sum_{i \in \mathcal{U}_k} I(j \in \mathcal{V}_i), \tag{1}$$

where $I(j \in \mathcal{V}_i)$ is an indicator with value 1 if video j was viewed by user i and 0 otherwise.

To measure how discriminative a video is for a group, we compare the total number of groups with the number of groups whose members have viewed video j, and define the discrimination score as

$$\beta_j = \log_2 \frac{G}{\sum_{k' \in \mathcal{G}} I(j \in \mathcal{P}_{k'})}, \tag{2}$$

where G is the total number of groups and $I(j \in \mathcal{P}_{k'})$ is an indicator with value 1 if video j is viewed by any user from group k' and 0 otherwise. Therefore, videos liked by a less number of groups will get higher discrimination scores.

Then, for each video j, we assess its representative level for group k by combining its local popularity score and discrimination score to generate a score $W_{k,j}$:

$$W_{k,j} = \eta_{k,j} * \beta_j, \tag{3}$$

where $\eta_{k,j}$ is the local popularity score and β_j is the discrimination score. The values of $\eta_{k,j}$ and β_j are affected by the number of members in the target group and the total number of groups in the system, respectively. For a certain video system, the values of $\eta_{k,j}$ and β_j can be scaled (such as max–min scaling) before being multiplied. In sum, we prefer videos that are locally popular in the target group rather than videos favored by most groups.

Furthermore, ranking videos by $\eta_{k,j}$ alone, i.e., local popularity, tends to favor those videos that are popular globally, while prioritizing β_j, namely discrimination, will generate more diverse and discriminative results. We could balance the local popularity and discrimination (diversity) by rescaling the range of $\eta_{k,j}$ and β_j. Since the video score is the product of these two values, it can be in a more general form:

$$W_{k,j} = \eta_{k,j}^p * \beta_j^q, \quad (p \geq 0, q \geq 0) \tag{4}$$

where $\eta_{k,j}$ and β_j are calculated as above; p and q are two constants to adjust the importance of local popularity score and discrimination score by changing the range of each score, i.e., the ratio of maximum and minimum values. A zero

value of p or q means ignoring the effect of the corresponding score, while a value between 0 and 1 will lessen one's impact and a value larger than 1 will enlarge one's importance, respectively. In our algorithm, we choose values of p and q so that local popularity and discrimination are well balanced, that is, their ratios of maximum and minimum values are comparable. Actually, some users are fond of hot videos, while others prefer videos of special interest. Thus, setting the values of p and q can be personalized for individual users, that is, for users who were active in the past, we can tune the values of p and q for adapting to their individual preference of video popularity learned from the historical records. We leave the exploration of other choices of p and q values as the future work.

Videos in group k's video pool $\mathcal{P}_k$ will be ranked according to their scores $W_{k,j}$ in the decreasing order. The ranked video list for group k is denoted by l_k. If user i is only affiliated with group k, then the top R videos from l_k, after removing videos viewed by user i, will be recommended to user i.

4.3 Video ranking for multiple affiliation problem

In this case, we can firstly apply the intra-group video ranking algorithm described in Sect. 4.2 to each of the affiliated groups. Then, we should consider how to merge the ranked video lists from those groups, which is a rank aggregation problem [4]. However, groups are of different values in video recommendation. For example, a highly interactive interest group may be more valuable than a colleague group because like-minded users are more likely to enjoy common videos. Thus, we need a group scoring algorithm to discriminate those groups.

4.3.1 Group scoring

To assess a group, we firstly define features that can distinguish groups and then use these features to calculate the group score.

Group features. Basically, we focus on two kinds of group features: social features and interest features. The social features we exploited comprise of social activeness and social conformity, which mainly take the social influence into account [25]. For example, good friends may share common interests in viewing videos. We detect the social influence from the density of friendship inside the group and the strength of interaction among group members. For interest features, we consider interest activeness and interest conformity. Preferred groups are those whose members are fond of viewing (especially representative) videos and like-minded with each other.

According to the advice-taking theory used in recommender systems [2], tie strength, trustworthiness and homophily are three significant factors to affect the likeli-

hood of users seeking and accepting someone's advice for decision making. In online social networks, tie strength can be measured by the frequency and duration of interaction, while trustworthiness corresponds to social relationship, and homophily means interest similarity.

With the knowledge of the group information and historical viewing records, we calculate the feature scores for group k as below:

- *Social activeness*: $S_k^a = \frac{\sum_{i \in \mathcal{U}_k} M_{k,i}}{|\mathcal{U}_k|}$, where $M_{k,i}$ is the number of group messages sent by user i in group k in last 30 days and $|\mathcal{U}_k|$ is the number of members in group k. Social activeness measures the intra-group interaction strength, which reflects the group-level tie strength.
- *Social conformity*: $S_k^c = \frac{\sum_{u,i \in \mathcal{U}_k, u \neq i} I_{u,i}}{|\mathcal{U}_k|(|\mathcal{U}_k|-1)/2}$, where $I_{u,i}$ is an indicator to show whether u and i are friends in QQ. Social conformity is the density of the friendship network inside this group, representing the trustworthiness on a group basis.
- *Interest activeness*: $I_k^a = \sum_{j \in \mathcal{P}_k} W_{k,j}$, where $W_{k,j}$ is the score of video j in group k calculated in Sect. 4.2. This feature measures how representative the group video candidate pool is in total.
- *Interest conformity*: $I_k^c = \frac{\sum_{u,i \in \mathcal{U}_k, u \neq i} \cos(u,i)}{|\mathcal{U}_k|(|\mathcal{U}_k|-1)/2}$, where $\cos(u,i) = \frac{|\mathcal{V}_u \cap \mathcal{V}_i|}{|\mathcal{V}_u|^{\frac{1}{2}}|\mathcal{V}_i|^{\frac{1}{2}}} \in (0,1)$. As shown in [4], the more alike the users in the group are, the more effective the group recommendations will be. This is a group-level video interest similarity measure corresponding to homophily. The larger this value, the higher the degree of common interest among this group.

Calculation of group scores. Those feature scores provide a four-dimensional comparison among different groups. However, in order to merge video lists from multiple groups, we need to combine these features to generate a single score for each group.

It is not easy to generate a group score with the feature scores. We should assign each feature a reasonable weight to combine them. We adopt the logistic function introduced in logistic regression to generate the group score, which is a value between 0 and 1 indicating the likelihood of a group to be effective for recommendation.[7]

$$J_k = \frac{1}{1 + e^{-(\theta_0 + \theta_1 S_k^a + \theta_2 S_k^c + \theta_3 I_k^a + \theta_4 I_k^c)}}, \tag{5}$$

where $\theta_0, \ldots, \theta_4$ are weights that can be tuned to combine feature scores.

---

[7] The reason we chose logistic regression is that it returns well-calibrated predictions by default as it directly optimizes log-loss.

We can use the supervised learning approach to learn the five feature weights. To prepare the training set, we choose in total m users randomly from all the online users. For each user i, we randomly select a group $k_i = \xi(\mathcal{G}_i)$ and use the proposed intra-group video ranking algorithm to recommend videos to the user. If any recommended videos are then selected and viewed by user i, then group k_i is effective to recommend videos for user i, denoted by $y_{k_i} = 1$; otherwise $y_{k_i} = 0$. For each instance in the training set, such as using group k_i to recommend videos for user i, the empirical error is calculated by the log-loss cost function $f_i(J_{k_i})$, which is defined as

$$f_i(J_{k_i}) = \begin{cases} -\log(J_{k_i}), & \text{if } y_{k_i} = 1 \\ -\log(1 - J_{k_i}), & \text{if } y_{k_i} = 0 \end{cases} \quad (6)$$

We can obtain optimal weights $\theta_0, \ldots, \theta_4$ by minimizing the regularized cost function shown below:

$$\min_{\theta_0,\theta_1,\theta_2,\theta_3,\theta_4} \frac{1}{m} \sum_{i=1}^{m} f_i(J_{k_i}) + \frac{\lambda}{2m} \sum_{j=1}^{4} \theta_j^2, \quad (7)$$

$$= -\frac{1}{m} \sum_{i=1}^{m} \left[y_{k_i} \log(J_{k_i}) + (1 - y_{k_i}) \log(1 - J_{k_i}) \right]$$

$$+ \frac{\lambda}{2m} \sum_{j=1}^{4} \theta_j^2. \quad (8)$$

Using the gradient descent method to train the parameters, the parameter updating in each iteration is

$$\theta_0 := \theta_0 \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^{m} (J_{k_i} - y_{k_i}), \quad (9)$$

$$\theta_1 := \theta_1 \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^{m} (J_{k_i} - y_{k_i}) S_k^a, \quad (10)$$

$$\theta_2 := \theta_2 \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^{m} (J_{k_i} - y_{k_i}) S_k^c, \quad (11)$$

$$\theta_3 := \theta_3 \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^{m} (J_{k_i} - y_{k_i}) I_k^a, \quad (12)$$

$$\theta_4 := \theta_4 \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^{m} (J_{k_i} - y_{k_i}) I_k^c, \quad (13)$$

where λ is the regularization parameter and α is the learning rate. The detailed training and testing results are shown in Sect. 5.

4.3.2 Video aggregation

With the ranked video lists from multiple groups as well as the score of each group, we should address the weighted ranking aggregation problem. Borda Fuse [4] is a widely known approach proposed to merge ranking lists. In Borda Fuse, each ranked video list is like a voter and each voter ranks a partial set of c video candidates. For each voter, the top-ranked video is assigned c scores, the second-ranked video is assigned $c - 1$ points and the like. If some videos left unranked by the voter, i.e., not in this ranked video list, the remaining scores are divided evenly among the unranked videos. In our case, we use a weighted Borda Fuse method to obtain the merged result R_i. For user i, the score of video j after merging multiple video lists is

$$W_{i,j} = \sum_{k \in \mathcal{G}_i} J_k * (D_i - p_{l_k}(j) + 1), \quad (14)$$

where J_k is the score of group k and D_i is the number of distinct videos in user i's group candidate video pool $\mathcal{P}_{\mathcal{G}_i}$. If $j \in l_k$, the variable $p_{l_k}(j) \in [1, D_i]$ is the position of video j in list l_k, otherwise, to allocate the remaining scores among the unranked videos,

$$p_{l_k}(j) = \frac{|l_k| + D_i + 1}{2}, \quad (15)$$

where $|l_k|$ is the length of list l_k. After ranking all the videos according to $W_{i,j}$ and removing videos viewed by user i, we can obtain the final recommendation list $\mathcal{R}_i$.

One thing to note is that the social-group-based algorithm in this paper does not utilize some model-based methods like those in CF,[8] for example, the proposed algorithm does not factorize the group information directly. Thus, the proposed algorithm is more like a memory-based one. Despite the higher accuracy in general achieved by model-based methods, memory-based methods show the superiority in terms of simplicity, interpretability and the ability of incremental updating [28], which makes them still prevalent in real systems.

5 Experimental results and discussion

5.1 Experiments and results

Instead of using historical data to conduct offline evaluations, we implement the social-group-based algorithm in the Tencent Video system to test it online. In our experiments, we focus on movie recommendation. In the current VoD service

---

[8] CF can be divided into two classes: memory-based CF and model-based CF.
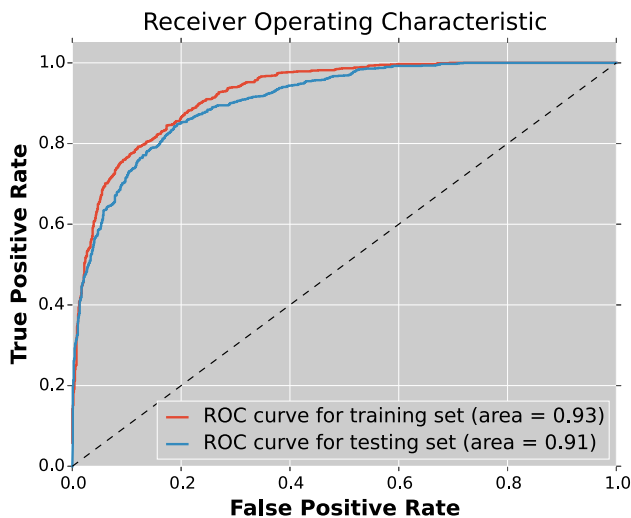
Fig. 3 The receiver operating characteristic curve of the feature weighting learning

of Tencent Video, there are more than 5 million daily views of movies.

To learn the feature weights in the group scoring algorithm, we randomly select 10 % of daily users along with random group selection and use the intra-group video ranking algorithm to recommend movies. Then, we collect the users' (implicit) feedback to conduct feature weighting training. To evaluate the performance of feature weighting, we collect feedback data in 2 days, using the former day's feedback data for training and the latter day's for testing.[9]

The receiver operating characteristic (ROC) curve of the feature weighting is shown in Fig. 3. Through our experiments for many times, the feature weighting result is relatively stable in terms of the feature weights, which implies that we could retrain it after a relatively long duration.

For online testing of the proposed social-group-based algorithm, we also implement two state-of-art approaches as benchmark for comparison.

- Implicit feedback-based collaborative filtering [12]: A matrix factorization model tailored for implicit feedback is utilized, where implicit feedback data are treated as indication of positive and negative preference associated with various confidence levels.
- Ontology-content-based filtering [26]: Each item profile is represented with a set of concepts taken from a video-related ontology. And each user's content-based profile, generated according to the user's implicit feedback, consists of a weighted list of ontology concepts representing

her interests. A cosine similarity measure is adopted to match users and items.

We conduct abundant A/B testing [11] online to evaluate the recommendation performance of the three algorithms.

In the A/B testing, users whose past viewing behaviors follow the distribution in Fig. 1 are diverted into several distinct sets evenly and randomly, where each set adopts a distinct setting for one targeted character and all other characters are fixed. These sets are then compared against one another over a set of predefined metrics. In our experiments, the targeted character is the adoption of different algorithms. The number of videos recommended to a user, i.e., R, is 16 in the online deployment. To evaluate the performances of these algorithms in terms of relevance and diversity,[10] we use two metrics in the experiments, namely click-through rate (CTR) [7] and Gini coefficient [6].

- CTR $= \frac{\# \text{ of click}}{\# \text{ of impression}}$, where "# of impression" is the total number of recommendation and "# of click" is the number of recommendation whose recommended video lists are clicked after they are shown to users.
- Gini $= \frac{2 \sum_{j=1}^{n} j*d_j}{n \sum_{j=1}^{n} d_j} - \frac{n+1}{n}$, where n is the number of distinct videos recommended and d_j is jth lowest frequency of occurrence in all recommended video lists. The smaller the Gini coefficient, the more diverse the recommendation results are.

We measured the performance over a period of 21 days. We normalize the second largest daily CTR to be 1 and the other two CTRs to be the ratio of it. The results of the normalized CTR and Gini coefficient are shown in Fig. 4, where each value is averaged over the same day of 3 weeks.[11]

The results show that the social-group-based algorithm achieved the highest CTR and smallest Gini coefficient, which means it can generate the most relevant recommendation in terms of hit rate and provide the most diverse results compared with the other two approaches. And this also indicates that the proposed algorithm can perform well in the cold-starting online video system.

5.2 Discussion and further investigation

In the online large-scale experiments, the social-group-based algorithm shows better performance in alleviating the cold-start issue in the Tencent Video system. Here we analyze

---

[9] We down-sampled the negative samples to maintain a positive/negative ratio of 1:2 in our dataset because the number of positive examples and negative examples is quite unbalanced.

[10] In recommender systems, the target is to meet the requirements of both the video Web sites (content providers) and the users. While accurately hitting the interest of users can improve the revenues of the websites (e.g., through advertisements), a diverse and even novel set of recommendations are also preferred by the users for better experience.

[11] Using a huge amount of users in a long period of experiments for the online tests could yield statistically significant results.
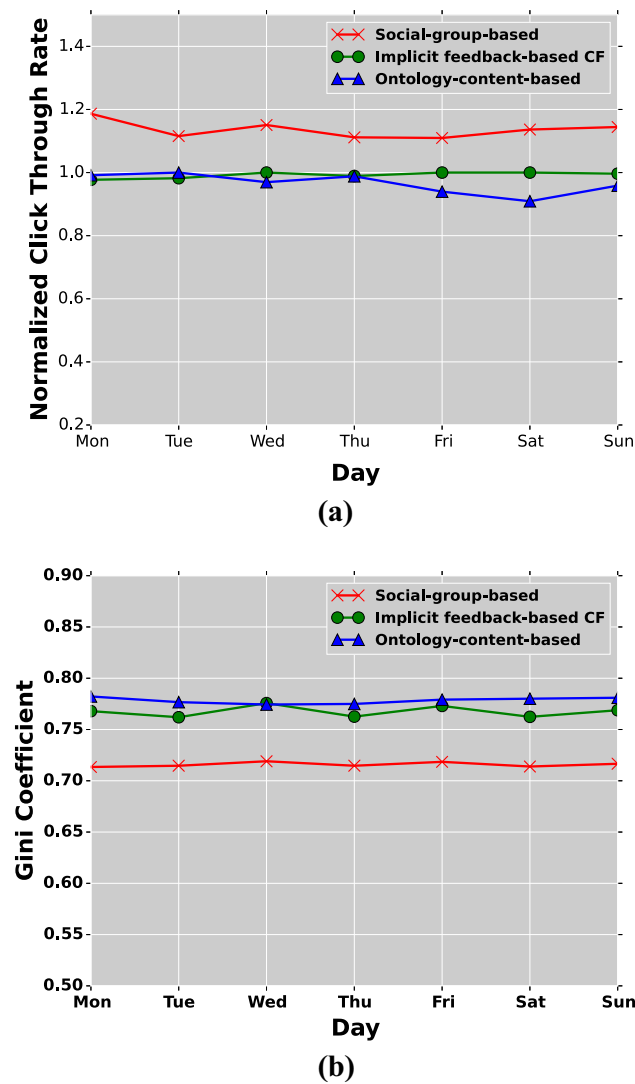
Fig. 4 Per-day average CTR and Gini coefficient for different algorithms over a period of 3 weeks

Fig. 5 CTR at each activeness value for comparison of collaborative filtering and group-based algorithms

the design rationale on how to address this issue. Firstly, the intra-group video ranking algorithm in the single affiliation problem tries to obtain a group-aggregated video list which aims at capturing preferences of the group and selecting representative videos for the group. For users with few or no viewing behaviors, i.e., in the cold-start or data sparsity scenarios, video candidates can still be obtained from their affiliated groups. Moreover, in the group scoring and video aggregation algorithms, four group features are chosen and calculated on the group level where social features can help infer groups' effectiveness to recommend videos. For cold-start users whose individual data (either social or interest) are missing, we could still infer their tastes from the groups they join.[12]

Compared with implicit feedback-based collaborative filtering, our proposed social-group-based algorithm achieves 15–20 % improvement concerning CTR metrics in the online test shown in Fig. 4. However, this does not mean that the social-group-based algorithm outperforms collaborative filtering in all cases. As we know, collaborative filtering does not perform well in cold-start scenarios, but it can accurately capture users' preference and make good recommendations given sufficient user behavior records since collaborative filtering is sufficiently personalized. As analyzed above, the social-group-based approach is not that dependent on individual data; thus, it is a partially personalized algorithm where users have different group affiliations. For those highly active users, it becomes harder for the social-group-based approach to generate relevant and diverse results since many videos in the group video candidate pool have been viewed previously by the active users. In order to illustrate the advantages of collaborative filtering and the social-group-based algorithm, respectively, we measure the performance of each algorithm regarding different types of users. Specifically, users are clustered according to their activeness, i.e., the number of videos viewed in the past few days.[13] Then, we plot the CTR over different activeness levels by averaging individual CTR values among users in each user cluster, as shown in Fig. 5.

Regardless of the specific values of CTR, we can observe the CTR trend of each algorithm with the increasing user activeness. Consistent with the above analysis, the social-group-based approach outperforms implicit feedback-based CF when users have low activeness. Note that in the Tencent Video system, the majority of users are inactive as shown by the distribution curve of user behavior amount in Fig. 1,

---

[12] This is also the reason why we did not consider individual features in the group scoring algorithm.

[13] In our case, we computed the number of movies viewed in the past 30 days.

which results in the better overall performance of the social-group-based approach. Also, large fluctuations are observed for high activeness levels due to the small number of very active users. With this comparative result, we can further design a hybrid switching algorithm. The scenarios when we can switch from social-group-based method to CF are (1) for users of activeness larger than the value corresponding to the intersection of the two curves[14] in Fig. 5; (2) for users that cannot get enough results due to lack of group affiliation or being too active in terms of number of viewed videos. The deployment and analysis of the hybrid algorithm will be in our future work.

6 Conclusion

In this paper, we propose a social-group-based video recommendation framework by virtue of explicitly formed groups (QQ groups). We elaborate on three algorithms of the framework, namely intra-group video ranking, group scoring and video aggregation. To validate the effectiveness of our approach, we deployed it in the online video system and compare it with two state-of-the-art algorithms. The evaluation results show that our design can produce recommendation results with both high relevance and diversity. In the future work, we will implement more algorithms for comparison, such as hybrid CF and CBF algorithms, and social-friend-based algorithm.

References

1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)
2. Arazy, O., Kumar, N., Shapira, B.: Improving social recommender systems. IT Prof. **11**(4), 38–44 (2009)
3. Arazy, O., Kumar, N., Shapira, B.: A theory-driven design framework for social recommender systems. J. Assoc. Inf. Syst. **11**(9), 455–490 (2010)
4. Baltrunas, L., Makcinskas, T., Ricci, F.: Group recommendations with rank aggregation and collaborative filtering. In: Proceedings of the Fourth ACM Conference on Recommender Systems, pp. 119–126. ACM (2010)
5. Bobadilla, J., Ortega, F., Hernando, A., Bernal, J.: A collaborative filtering approach to mitigate the new user cold start problem. Knowl.-Based Syst. **26**, 225–238 (2012)
6. Coefficient, G.: http://en.wikipedia.org/wiki/gini
7. Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B., et al.: The youtube video recommendation system. In: Proceedings of the Fourth ACM Conference on Recommender Systems, pp. 293–296. ACM (2010)
8. Dell'Amico, M., Capra, L.: Sofia: Social filtering for robust recommendations. In: IFIP International Conference on Trust Management, pp. 135–150. Springer (2008)
9. Doersch, C., Singh, S., Gupta, A., Sivic, J., Efros, A.: What makes paris look like paris? ACM Trans. Graph. (Proc. SIGGRAPH) **31**(4), 1–9 (2012)
10. Gilly, M.C., Graham, J.L., Wolfinbarger, M.F., Yale, L.J.: A dyadic study of interpersonal information search. J. Acad. Mark. Sci. **26**(2), 83–100 (1998)
11. evaluation at Google, S.: http://googleblog.blogspot.hk/2008/09/search-evaluation-at-google.html
12. Hu, Y., Koren, Y., Volinsky, C.: Collaborative filtering for implicit feedback datasets. In: Eighth IEEE International Conference on Data Mining, 2008. ICDM'08, pp. 263–272. IEEE (2008)
13. Jamali, M., Ester, M.: Trustwalker: a random walk model for combining trust-based and item-based recommendation. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 397–406. ACM (2009)
14. Jamali, M., Ester, M.: A matrix factorization technique with trust propagation for recommendation in social networks. In: Proceedings of the Fourth ACM Conference on Recommender Systems, pp. 135–142. ACM (2010)
15. Jameson, A., Smyth, B.: Recommendation to groups. In: The Adaptive Web, pp. 596–627. Springer (2007)
16. Koren, Y.: Factorization meets the neighborhood: a multifaceted collaborative filtering model. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 426–434. ACM (2008)
17. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **8**, 30–37 (2009)
18. Levin, D.Z., Cross, R.: The strength of weak ties you can trust: the mediating role of trust in effective knowledge transfer. Manag. Sci. **50**(11), 1477–1490 (2004)
19. Lewis, K., Gonzalez, M., Kaufman, J.: Social selection and peer influence in an online social network. Proc. Natl. Acad. Sci. **109**(1), 68–72 (2012)
20. Liu, F., Lee, H.J.: Use of social network information to enhance collaborative filtering performance. Expert Syst. Appl. **37**(7), 4772–4778 (2010)
21. Lops, P., De Gemmis, M., Semeraro, G.: Content-based recommender systems: state of the art and trends. In: Recommender Systems Handbook, pp. 73–105. Springer (2011)
22. Ma, N., Lim, E.P., Nguyen, V.A., Sun, A., Liu, H.: Trust relationship prediction using online product review data. In: Proceedings of the 1st ACM International Workshop on Complex Networks Meet Information & Knowledge Management, pp. 47–54. ACM (2009)
23. Report, T.: http://www.tencent.com/en-us/content/at/2015/attachments/20150812.pdf (2015)
24. Sahebi, S., Cohen, W.W.: Community-based recommendations: a solution to the cold start problem. In: Workshop on Recommender Systems and the Social Web, RSWEB (2011)
25. Shalizi, C.R., Thomas, A.C.: Homophily and contagion are generically confounded in observational social network studies. Sociol. Methods Res. **40**(2), 211–239 (2011)
26. Shoval, P., Maidel, V., Shapira, B.: An ontology-content-based filtering method. Institute of Information Theories and Applications FOI ITHEA (2008)
27. Smith, D., Menon, S., Sivakumar, K.: Online peer and editorial recommendations, trust, and choice in virtual markets. J. Interact. Mark. **19**(3), 15–37 (2005)
28. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. Adv. Artif. Intell. **2009**, 4 (2009)
29. Tang, J., Hu, X., Liu, H.: Social recommendation: a review. Soc. Netw. Anal. Min. **3**(4), 1113–1133 (2013)

---

[14] In the case that there are more than one crossing points due to the curve fluctuation, we can select any in-between point as the threshold.

30. Tang, L., Wang, X., Liu, H.: Understanding emerging social structures—a group profiling approach. School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Technical Report TR-10-002 (2010)

31. Wen, Z., Lin, C.Y.: On the quality of inferring interests from social neighbors. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 373–382. ACM (2010)

32. Xu, B., Bu, J., Chen, C., Cai, D.: An exploration of improving collaborative recommender systems via user-item subgroups. In: Proceedings of the 21st International Conference on World Wide Web, pp. 21–30. ACM (2012)

33. Yang, C., Zhou, Y., Chen, L., Zhang, X., Chiu, D.M.: Social group based video recommendation addressing the cold-start problem. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 515–527. Springer (2016)

34. Yang, X., Steck, H., Liu, Y.: Circle-based recommendation in online social networks. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1267–1275. ACM (2012)

35. Yao, L., Sheng, Q.Z., Segev, A., Yu, J.: Recommending web services via combining collaborative filtering with content-based features. In: 2013 IEEE 20th International Conference on Web Services (ICWS), pp. 42–49. IEEE (2013)

36. Zhou, K., Yang, S.H., Zha, H.: Functional matrix factorizations for cold-start recommendation. In: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 315–324. ACM (2011)

37. Ziegler, C.N., Golbeck, J.: Investigating interactions of trust and interest similarity. Decis. Support Syst. **43**(2), 460–475 (2007)