IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 27, 2019, accepted February 1, 2020, date of publication February 10, 2020, date of current version February 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2972920

Social- and Content-Aware Prediction for Video Content Delivery

YUQI FAN[1,2], BING YANG[1], DONGHUI HU[1], XIAOHUI YUAN[3], AND XIONG XU[2]

[1]Key Laboratory of Knowledge Engineering With Big Data (Ministry of Education), School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009, China

[2]State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, Luoyang 471003, China

[3]Department of Computer Science and Engineering, University of North Texas, Denton, TX 76203, USA

Corresponding author: Yuqi Fan (yuqi.fan@hfut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant U1836102, and in part by the Open Project of State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System in China under Grant CEMEE2018Z0102B.

ABSTRACT An ever-increasing number of videos in mobile social networks increase the waiting time of video downloading. Video prefetching is a viable way to save time during video downloading, while prefetching all the videos is resource wasteful if the videos are not watched. Therefore, careful prediction of whether a user will watch a video is critical for efficient video content delivery. Most existing work on social media recommendation focuses on the top-*N* problem to recommend multiple videos. In this paper, we deal with the problem of predicting whether a video will be watched by a user for efficient video content delivery in mobile social networks. We propose a Social- and Content-aware Video content delivery Prediction method (SCVP) for the problem by capturing the intrinsic relationship among users and videos. We design five metrics to estimate the factors of active degree of users, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity. We then use combined prediction for the video content delivery prediction to incorporate the impacts of the five factors on the prediction. Finally, we conduct experiments through simulations. Experimental results demonstrate that the proposed method SCVP can predict whether a video is watched by users with high accuracy.

INDEX TERMS Social video, video content delivery, social relationship, combined prediction.

I. INTRODUCTION

With the rapid development of mobile Internet and the popularity of mobile devices, such as smartphones, tablets, and personal computers, etc., an increasing number of users watch videos in online social networks (OSNs) on mobile devices. As the number of videos continues to increase, the waiting time of video downloading has increased dramatically [1]. Video prefetching is one of the effective approaches to reducing the waiting time of users to watch the videos. However, prefetching all the videos is resource wasteful if the videos are not watched, thereby generating a prediction problem of whether a video will be viewed by the user. The video prefetching prediction is used to decide whether to download the video to the user in advance, so as to reduce access delay and improve the user experience.

A problem related to the prediction of video content delivery is the social media recommendation problem which focuses on the top-*N* problem to recommend multiple videos. Social networks provide rich user information such as browsing history and social information, and most recommendation algorithms use the information for the correlation analyses between users and videos [2]–[5]. For example, it is highly possible that the users closely connected in their social circle have similar interests and watch the similar videos, and a user will potentially watch a video which is consistent to the user's interests. Note that the active degree of users, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity all have impacts on whether a user will watch a video as illustrated in Example 1.

Example 1: We consider 5 videos and 2 users in an OSN shown in Fig. 1. (1) Video v_1 is about the mental health of the elderly, which is not popular. User u_1 is a young man, and u_1 and his/her social friends seldom watch this kind of videos. However, user u_1 is so active on the OSN that he/she often kills time by browsing the OSN, and hence u_1 will watch video v_1. The video prefetching prediction will

The associate editor coordinating the review of this manuscript and approving it for publication was Xiping Hu.

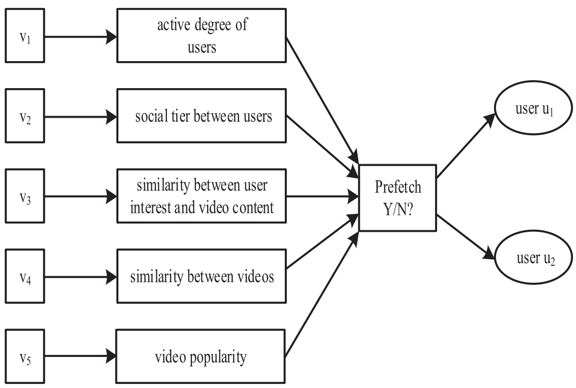


FIGURE 1. An example of 5 important factors.

not get the correct result if the four factors other than the active degree of users are considered during the prediction. (2) An unpopular video v_2 is about diving. User u_1 lives in an inland city. User u_1 is not interested in diving and has seldom watched the videos about diving. However, some of his/her social friends have watched v_2, and u_1 will also watch v_2, since he/she is curious why his/her friends watch the video. The video prefetching prediction will not work correctly without considering the social tiers between users. (3) Video v_3 is a skiing game which is not popular. User u_2 likes sports and he is interested in kinds of sports. Even if few of the user's friends have watched the skiing video, user u_2, a sports fan, will watch video v_3, although he seldom watches skiing games. However, we cannot obtain the correct prediction result by ignoring the similarity between user interest and video content. (4) Video v_4 introduces the place where the skiing game of video v_3 takes place. User u_2 is not interested in that place, and his/her friends have never watched the video about that place. However, user u_2 will watch video v_4, since videos v_4 and v_3 are similar in terms of the venue of the game. The video prefetching prediction will not generate the correct result without taking into account the similarity between videos. (5) Video v_5 is about the Amazon rainforest fire which is a hot topic. The friends of user u_2 do not care about the big issue, and user u_2 is not interested in natural disasters. However, user u_2 will watch the video, since the disaster is always among the top news. We cannot obtain the correct prediction without considering the video popularity.

The five factors are important for the video prefetching prediction accuracy. However, few social media recommendation methods integrate all the five factors. Furthermore, the existing recommendation algorithms focus on the top-N problem to recommend multiple videos, but rarely solve the prediction problem of whether a video will be viewed by the user. In this paper, we propose a Social- and Content-aware Video content delivery Prediction method (SCVP) for the problem by capturing the intrinsic relationship among users and videos. The main contributions of this paper are as follows:

(1) We deal with the problem of predicting whether a video will be watched by a user for efficient video content delivery in online social networks.

(2) We propose a Social- and Content-aware Video content delivery Prediction method for the problem by exploiting the intrinsic relationship among users and videos. We design five metrics to estimate the factors of the active degree of users, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity. We then adopt combined prediction for the video content delivery prediction to incorporate the impacts of the five factors on the prediction.

(3) We use the competition track 1 of 2012 KDD Cup provided by Tencent Weibo to verify the proposed method SCVP. The experimental results show that the proposed method SCVP can effectively improve the prediction performance in terms of precision, recall, and F1-measure.

The rest of the paper is organized as follows. In section II, we introduce the related work. In section III, we present the proposed method SCVP. In section IV, we illustrate the performance evaluation of the proposed method SCVP. Finally, the conclusions are detailed in Section V.

II. RELATED WORK

A. ONLINE SOCIAL NETWORK

An online social network is a network which reflects the social structure and the interdependence of users, and OSNs provide the medium through which users can disseminate information and ideas, such that the users can be influenced by friends' decisions and change their decision-making behavior [6]. With the rapid development of Web 2.0 technology, OSNs, i.e. Twitter, Facebook, Sina Weibo, etc., are now receiving more and more attention. According to the data collected in 2016, Facebook and Twitter have owned 1.59 billion and 329 million active users per month, respectively [1].

B. TRADITIONAL RECOMMENDATION SYSTEMS

Currently, many researchers adopt content-based recommendation algorithms, collaborative filtering recommendation algorithms, and hybrid recommendation algorithms to solve video recommendation problems. A content-based recommendation system computes the similarity among users based on historical records to achieve the recommendation purpose. For example, method CBF was proposed to use a multi-attribute network to measure the similarity between the linked items; method CBF adopts centrality and clustering techniques to consider the mutual relationship among items and the structural patterns of item interactions [7]. A Bayesian-inference based recommendation system was proposed to construct a Bayesian network to infer the rating of a querying user [8]. The content-based recommendation systems can solve the problem of structured information recommendation. However, for unstructured information, the recommendation performance needs to be improved.

Collaborative filtering recommendation algorithms [9] recommend videos to users based on the interest similarity between the target users and the other users by users' ratings [10], which has become one of the most successful methods for personalized recommendation service. A trust-based collaborative recommendation algorithm was proposed to use rating trust and preference trust [11]. A recommendation algorithm was proposed in [12] to apply matrix decomposition to the nearest neighbor recommendation system, so as to consider the relationship between users and items.

Hybrid recommendation algorithms combine collaborative filtering and content-based recommendation algorithms. A framework based on the combination of support vector machine and traditional collaborative filtering was proposed to improve the recommendation performance [13]. A joint content-based and collaborative filtering approach was proposed to generate a hybrid recommendation system; the approach adopts content-based predictors to enhance the existing user data and then provides personalized recommendations through collaborative filtering [14].

C. SOCIAL NETWORK BASED RECOMMENDATION METHODS

Users are influencing and influenced by the social friends in OSNs, and the social relationship can be used in recommendation. Therefore, some researchers have proposed recommendation methods by making use of the characteristics of social networks. A social recommendation method based on social interaction, trust relationship and product popularity was proposed to predict user preferences and recommend related products in social networks [2]. A video recommendation algorithm based on the combination of video content and social network was proposed; the proposed algorithm consists of trust friends computing model and video's quality evaluation model [3]. A model was proposed in [4] to contain the influence of social relationships to generate recommendations. A social recommendation system model called interest social recommendation (ISoRec) was proposed; based on probability matrix factorization (PMF), the model combines user-item rating matrix, explicit user social connection information, and implicit user interest social connection information to make a recommendation [15]. A recommendation system in social networks with Feature Transfer and Probabilistic Matrix Factorization (FTMF) was proposed; the auxiliary data and matrix factorization technique were integrated to learn a social latent feature vector of users, and an adaptive firm factor was introduced to balance the impacts from user's own factors and trusted people on purchasing behavior for each user [16]. A trust-based probabilistic recommendation model (TBPR) was proposed to combine the similarity among products and the trust of products, where the trust of products is obtained based on reputations and purchase frequencies [17]. An energy-efficient download scheduling algorithm was proposed for video streaming based on an aggregate model which is constructed through a personal retention model with users' personal viewing history and the audience retention on crowd-sourced viewing history [18].

The existing recommendation algorithms focus on the top-N problem to recommend multiple videos using social information and browsing history, but rarely solve the prediction problem of whether a video will be viewed by the user. Furthermore, the methods fail to fully capture the intrinsic relationship among users and videos. In this paper, we propose a Social- and Content-aware Video content delivery Prediction method to explore the impacts of five factors on video prefetching prediction accuracy, including the active degree of users, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity.

III. THE PROPOSED METHOD SCVP

The proposed method SCVP is illustrated in Fig. 2. We explore the characteristics of users and videos and the intrinsic relationship between users and videos, so as to provide video prefetching prediction. To be specific, we design five metrics to estimate the factors of active degree of users, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity. If a user is followed by a large number of social friends and watches a large amount of recommended videos, we think the user is active on the OSN. If a video is about a hot topic on the OSN, the video is very likely to be watched by the user. A user is influenced by the social friends whom the user trusts and who have the same interest as the user, and hence the user will potentially watch the video which those social friends have watched and shared. It is very possible that a user will watch a video which content the user is interested in. A user is also likely to watch a video which is similar or related to some videos the user has watched. After obtaining the five metrics, we adopt combined prediction to integrate

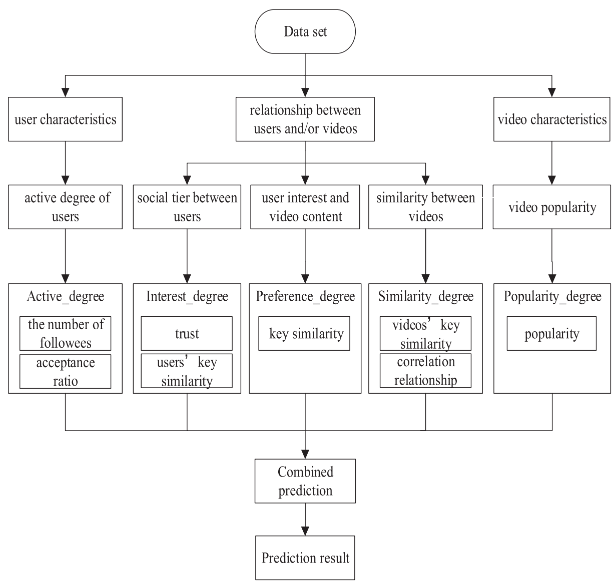


FIGURE 2. The proposed method SCVP.

the impacts of the five factors on video prefetching prediction by putting the calculated metrics as the input features of a classification algorithm. The users who need to prefetch videos and the videos that need to be prefetched are denoted as target users and target videos, respectively.

A. ACTIVE DEGREE OF USERS

The active degree of a user represents the active level of the user in the social network. An active user is likely to watch many videos.

In social networks, the users can freely share information. The users follow people who have similar interests so that the users are connected by the following behaviors. The user's following behavior potentially reflects the homogeneity phenomenon. Therefore, we introduce the following activity degree defined as Equation (1).

$$Foll\_active(i) = \frac{1}{1 + e^{\frac{-followee\_num(i)}{100}}} \tag{1}$$

where $Foll\_active(i)$ and $followee\_num(i)$ are the following activity degree of target user i and the number of users following target user i, respectively.

Among the videos recommended to the user, the user may not necessarily watch all the videos. The ratio of the number of videos watched to the number of videos recommended to the user can also provide a reference to whether the user will watch a video. Therefore, we introduce the user's acceptance ratio defined via Equation (2).

$$Accept\_ratio(i) = \frac{accept\_num(i)}{recomm\_num(i)} \tag{2}$$

where $Accept\_ratio(i)$, $accept\_num(i)$, and $recomm\_num(i)$ denote the acceptance ratio of target user i, the number of recommended videos which are watched by target user i, and the total number of videos recommended to target user i, respectively.

The following activity degree and acceptance ratio are combined as a user activity degree metric via Equation (3).

$$AD(i) = a_1 \cdot Foll\_active(i) + a_2 \cdot Accept\_ratio(i) \tag{3}$$

where $AD(i)$ represents the active degree of target user i, and a_1 and a_2 $(a_1 + a_2 = 1)$ are the weight coefficients of the following activity degree and acceptance ratio of target user i, respectively.

B. INTEREST DEGREE

Trust has become an indispensable element in social networks. A user is usually influenced by those the user trusts. If the trusted users of the target user watch the target video, it is of a high probability for the target user to watch the video. Some users post tweets, and some other users forward or comment on the tweets. The interactions between users reflect the trust between the users and the user's interests. The interaction level between target user i and user u is evaluated by mining their interactions including forwarding, commenting, and mentioning.

$$Interaction(i, u) = t_1 \cdot at\_num(i, u) + t_2 \cdot retweet\_num(i, u) + t_3 \cdot comment\_num(i, u) \tag{4}$$

where $Interaction(i, u)$ represents the interaction level of target user i and user u, at_num denotes the number of tweets in which user u is mentioned by target user i, $retweet\_num(i, u)$ is the number of tweets written by user u and forwarded by target user i, and $comment\_num(i, u)$ represents the number of tweets which are posted by user u and commented by target user i. Coefficients t_1, t_2 and t_3 $(t_1 + t_2 + t_3 = 1)$ are the weights of the interactions of mentioning, forwarding, and commenting, respectively.

The degree of how target user i trusts user u, $Tr(i, u)$, is evaluated through Equation (5).

$$Tr(i, u) = \frac{Interaction(i, u) - min_{Interaction(i,u)}}{max_{Interaction(i,u)} - min_{Interaction(i,u)}} \tag{5}$$

where $max_{Interaction(i,u)}$ and $min_{Interaction(i,u)}$ represent the maximum and minimum interaction levels between target user i and user u, respectively.

User's keywords indicate the user's potential interests. In general, a high similarity between two users' keywords indicates that the two users have common interests. We calculate the users' keyword similarity between target user i and user u, $Users\_ksim(i, u)$, with Equation (6) based on the Jaccard similarity coefficient [19].

$$Users\_ksim(i, u) = \frac{keyword_i \cap keyword_u}{keyword_i \cup keyword_u} \tag{6}$$

where $keyword_i$ and $keyword_u$ denote the keywords of target user i and user u, respectively.

We evaluate the intimacy between target user i and user u, $Intimacy(i, u)$, via Equation (7) based on the users' keyword similarity and the degree of target user i trusting user u.

$$Intimacy(i, u) = b_1 \cdot Users\_ksim(i, u) + b_2 \cdot Tr(i, u) \tag{7}$$

where b_1 and b_2 $(b_1 + b_2 = 1)$ are the weights of users' keyword similarity and trust degree, respectively.

In order to evaluate the impact of users' social tier on the correlation between the target user and the target video, the users' video watching behaviors and the intimacy between users are combined as the interest degree using Equation (8).

$$ID(i, m) = \frac{1}{p_{recomm\_num}} \sum_{j=1}^{N_1} Intimacy(i, u_j) \cdot R_j^m \tag{8}$$

where $ID(i, m)$ represents the degree of target user i being interested in target video m, p_{recomm_num} denotes the number of videos recommended to the social friends of target user i, and N_1 is the number of social friends of target user i. R_j^m indicates whether social friend u_j watches target video m; 1 represents that the video is watched, 0 means the video is not recommended, and -1 indicates that the video is not watched.

C. PREFERENCE DEGREE

If the user is interested in the video's content, it is highly possible for the user to watch the video. We analyze the similarity between user interests and video contents.

User's keywords indicate the user's interests, and the video's keywords reflect the content of the video. If the keywords of the user and the video are similar, the user is potentially interested in the video's content. We calculate the similarity between the keywords of target user i and target video m, $Key\_sim(i, m)$, via Equation (9) based on the Jaccard similarity coefficient [19].

$$Key\_sim(i, m) = \frac{keyword_i \cap keyword_m}{keyword_i \cup keyword_m} \tag{9}$$

where $keyword_i$ and $keyword_m$ represent the keywords of target user i and target video m, respectively.

The target user's social circle also has an influence on the target user's video watching behavior. If a user, one of the target user's social friends, is interested in the target video, the target user is also likely to watch the video. In order to evaluate the degree of target user i preferring target video m, we introduce the preference degree with Equation (10) by combining the similarity between the keywords of the target video and the target user and the similarity between the keywords of the target video and the social friends.

$$PD(i, m) = \partial_1 \cdot Key\_sim(i, m) + \frac{\partial_2}{N_1} \cdot \sum_{j=1}^{N_1} Key\_sim(u_j, m) \tag{10}$$

where $PD(i, m)$ represents the preference degree of target user i to target video m, $Key\_sim(u_j, m)$ is the similarity between the keywords of social friend u_j and target video m, and ∂_1 and ∂_2 ($\partial_1 + \partial_2 = 1$) respectively denote the weight coefficients of the similarity between the keywords of target user i and target video m and the similarity between the keywords of the social friends and target video m.

D. SIMILARITY DEGREE

If the videos similar to the target video are watched by the target user, the target video is also likely to be watched by the target user. Videos' keyword similarity reflects the similarity between video contents. The similarity between the contents of two videos m and v, denoted as $Videos\_ksim(m, v)$, can be calculated by Equation (11).

$$Videos\_ksim(m, v) = \frac{keyword_m \cap keyword_v}{keyword_m \cup keyword_v} \tag{11}$$

The correlation between videos can be evaluated via the users' video watching behaviors as well as the videos' contents. If a large number of users watch the same videos, the videos may potentially be closely correlated with each other. The correlation between video v and target video m is calculated via Equation (12).

$$Sr(m, v) = \frac{1}{1 + e^{-c_{mv}}} \tag{12}$$

where $Sr(m, v)$ is the correlation between target video m and video v, and c_{mv} denotes the number of users who have watched both videos m and v.

The similarity between videos m and v, $S\_sc(m, v)$, is evaluated according to the videos' keyword similarity and the correlation between the two videos.

$$S\_sc(m, v) = \kappa_1 \cdot Videos\_ksim(m, v) + \kappa_2 \cdot Sr(m, v) \tag{13}$$

where κ_1 and κ_2 ($\kappa_1 + \kappa_2 = 1$) are the weights of videos' keyword similarity and videos' correlation, respectively.

The similarity degree of target user i and target video m, $SD(i, m)$, is calculated via Equation (14).

$$SD(i, m) = \frac{\varepsilon_1}{q_{re\_num1}} \cdot \sum_{k_1=1}^{q_{re\_num1}} S\_sc(m, v_{k_1}) \cdot Q_i + \frac{\varepsilon_2}{q_{re\_num2}} \cdot \sum_{k_2=1}^{q_{re\_num2}} S\_sc(m, v_{k_2}) \cdot Q_u \tag{14}$$

where q_{re_num1} and q_{re_num2} represent the number of videos recommended to target user i and the social friends, respectively. Q_i indicates whether video v_{k_1} has been viewed by target user i; $Q_i = 1$ if the video is watched, and $Q_i = -1$ otherwise. Q_u indicates whether video v_{k_2} has been viewed by the social friends; $Q_u = 1$ if the video is viewed, and $Q_u = -1$ otherwise. ε_1 and ε_2 ($\varepsilon_1 + \varepsilon_2 = 1$) respectively denote the weights of the similarity between the target video and the videos watched by the target user and the similarity between the target video and the videos watched by the social friends.

E. POPULARITY DEGREE

The user's preference to a video is affected by not only the users' interactions and trust relationship, but also the video itself. A user may potentially watch a popular video. We define the popularity degree of a video by the ratio of the number of times that the video is watched to the number of social friends to whom the video is recommended. A high ratio indicates that the video is popular.

$$PE(m) = \frac{accept\_num(m)}{recomm\_num(m)} \tag{15}$$

where $PE(m)$ is the popularity degree of target video m, and $accept\_num(m)$ and $recomm\_num(m)$ denote the number of users who have watched target video m and the number of times that target video m is recommended, respectively.

F. COMBINED PREDICTION

We need to capture the intrinsic relationship among users and videos for effective prediction of whether the target video will be watched by the target user. Combined prediction incorporates different prediction metrics and comprehensively utilizes the information provided by various prediction metrics to improve the prediction accuracy. The prediction problem of whether the target video will be watched by the target user can be modeled as a binary classification problem. Therefore, we can use a classification algorithm, i.e. k-Nearest Neighbor (KNN), Random Forest (RF), Multilayer Perceptron (MLP), decision tree (DT), and Support Vector Machine (SVM), etc., to solve the prediction problem. We calculate the five metrics which can reflect the correlation between the target user and the target video, and use the calculated metrics as the input features of the classification algorithm. The output of the classifier is whether to prefetch the target video.

IV. EXPERIMENTS AND PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed method SCVP. We also investigate the impact of important factors on the performance of the proposed method.

A. EXPERIMENTAL SETUP

The experimental data set is the competition track 1 of 2012 KDD Cup provided by Tencent Weibo, which is publicly available and includes 1-month trace logs [20]. The data set contains user attributes, social relationships, interaction records in social networks, etc. as well as historical video recommendation records. All the data reflect the behaviors of whether the users watch the videos in one month, including acceptance and rejection behaviors with regard to the recommendations [3].

We take the data of user-watching videos in the first 20 days as the users' historical behaviors, and the data are used to estimate the five metrics introduced in this paper. The data of user-watching videos in days 21-27 are used as the training data set, which is used to train the classifier. We use the data of user-watching videos in the last 3 days as the testing data set, which is used to examine the prediction result. To evaluate the performance of the proposed method SCVP, the users to be recommended videos are those who have watched at least 10 videos and have at least 50 social friends.

The performance metrics are precision, recall, and F1-measure, which are commonly used in the prediction and recommendation systems. Precision is the prediction accuracy, which is calculated as Equation (16).

$$precision = \frac{N}{L} \tag{16}$$

where N and L denote the number of recommended videos watched by the target users and the number of videos recommended to the target users, respectively.

Recall is the ratio of the number of target users who watch the recommended videos to the total number of videos that the target users have watched. Recall is calculated via Equation (17).

$$recall = \frac{N}{B} \tag{17}$$

where B is the total number of videos that the target users have watched.

F1-measure is calculated as Equation (18), which combines the metrics of precision and recall.

$$F1 - measure = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{18}$$

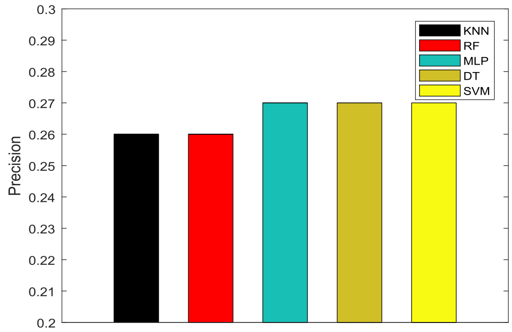


FIGURE 3. The precision performance of different classification algorithms.

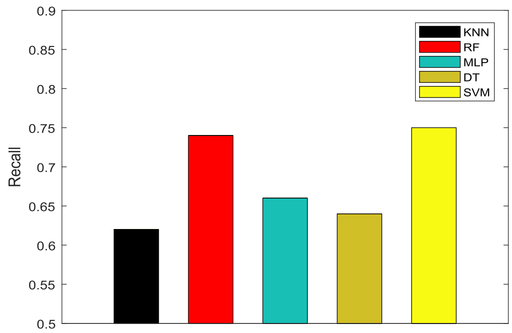


FIGURE 4. The recall performance of different classification algorithms.

The parameters used in the proposed method SCVP are listed in Table 1, and the parameters' values are determined through experiments so that SCVP can achieve good performance.

B. EXPERIMENTAL RESULTS

1) IMPACT OF DIFFERENT CLASSIFICATION ALGORITHMS ON THE PERFORMANCE

We adopt different classification algorithms, i.e. KNN, RF, MLP, DT, and SVM, to evaluate the performance of the proposed method SCVP. Fig. 3 shows the precision performance of the five algorithms. Algorithms SVM, MLP, and DT achieve close results, while algorithms KNN and RF obtain the similar results. SVM outperforms KNN and RF about 3.8%. Fig. 4 illustrates the recall performance of the five algorithms. SVM achieves the best results among the five classifiers. To be specific, SVM exhibits a performance improvement of about 20.9%, 17.2%, 13.6%, and 1.4% as compared with KNN, DT, MLP, and RF, respectively. Fig. 5 depicts the F1-measure performance of the five algorithms. It can be observed that SVM also obtains the best result among the five algorithms. SVM outperforms KNN, MLP, DT, and RF about 11.1%, 5.3%, 5.3%, and 2.6%, respectively. In general, SVM leads to the best results in the three performance metrics among the five classifiers as demonstrated in Figs. 3-5. Therefore, we adopt SVM as the classifier in the following experiments.

TABLE 1. The settings of parameters in our proposed algorithm.

Parameter	a_1	a_2	t_2	t_2	t_3	b_1	b_2	∂_1	∂_2	κ_2	κ_2	ε_1	ε_2
Value	0.3	0.7	0.3	0.4	0.3	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

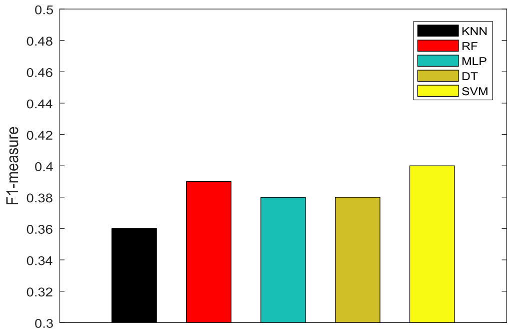


FIGURE 5. The F1-measure performance of different classification algorithms.

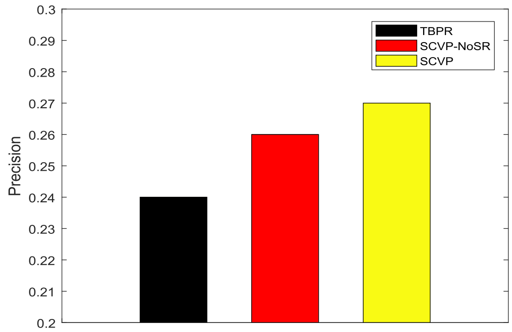


FIGURE 6. The precision performance of different methods.

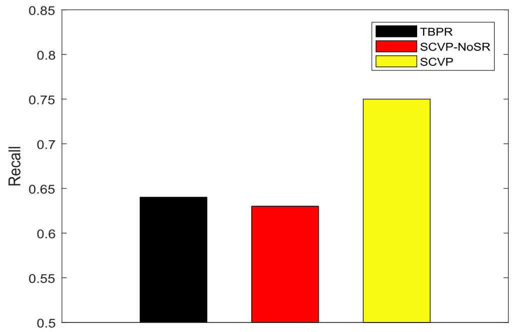


FIGURE 7. The recall performance of different methods.

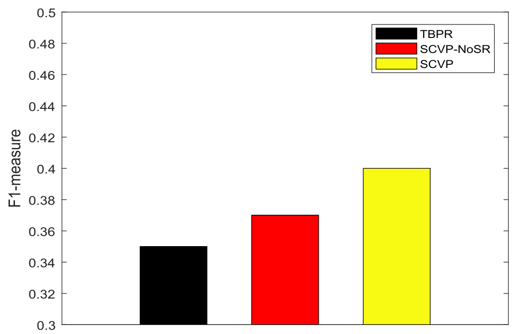


FIGURE 8. The F1-measure performance of different methods.

2) COMPARISON WITH DIFFERENT PREDICTION METHODS

We compare the proposed method SCVP with two methods: (1) Trust-Based Probabilistic Recommendation model (TBPR) [17] which is the state-of-art, and (2) SCVP-NoSR, where the proposed method SCVP ignores the social relationship, i.e. social tier between users, similarity between videos, and similarity between user interest and video content.

Figs. 6-8 illustrate the performance of the three methods in terms of precision, recall, and F1-measure, respectively. SCVP performs the best among the three methods in all the three performance metrics. Fig. 6 shows that SCVP exhibits an improvement of about 12.5% and 3.8% as compared with TBPR and SCVP-NoSR, respectively, in terms of precision. With regard to the recall performance, Fig. 7 demonstrates the results of SCVP are 17.2% and 19% better than those of TBPR and SCVP-NoSR, respectively. For the F1-measure performance shown in Fig. 8, SCVP outperforms TBPR and SCVP-NoSR 14.3% and 8.1%, respectively. SCVP obtains excellent performance by capturing the characteristics of users and videos and exploring the relationship between users and videos. TBPR only exploits the characteristics of and the similarity between the videos, ignoring the intrinsic relationship between users and videos. SCVP-NoSR generates the video prefetching prediction by only harnessing the characteristics of users and videos.

3) IMPACT OF DIFFERENT FACTORS ON THE PERFORMANCE

In order to evaluate the impact of different factors on the prediction results, we adopt different combinations of the factors to evaluate the proposed method SCVP. There are five cases with each ignoring a factor: (1) Case R1 represents the combination of interest degree, preference degree, similarity degree, and popularity degree, with user activity degree missing; (2) Case R2 denotes the combination of user activity degree, preference degree, similarity degree, and popularity degree, by ignoring interest degree; (3) Case R3 combines user activity degree, interest degree, similarity degree, and popularity degree, without considering preference degree; (4) Case R4 is the combination of user activity degree, interest degree, preference degree, and popularity degree, by ignoring similarity degree; (5) Case R5 combines user activity degree, interest degree, preference degree, and similarity degree, without incorporating popularity degree.

Figs. 9-11 depict the performance of different combinations of the factors in terms of precision, recall, and

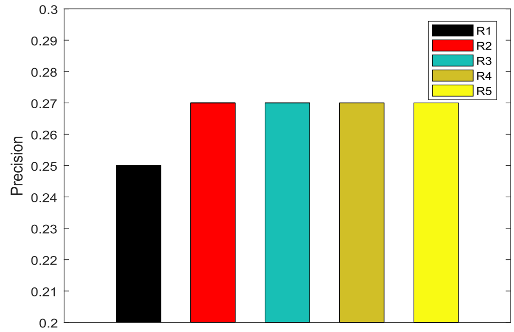


FIGURE 9. The precision performance of different factor combinations.

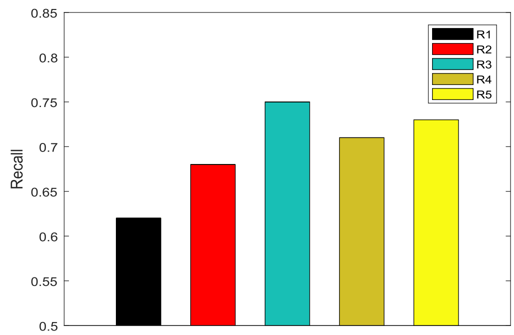


FIGURE 10. The recall performance of different factor combinations.

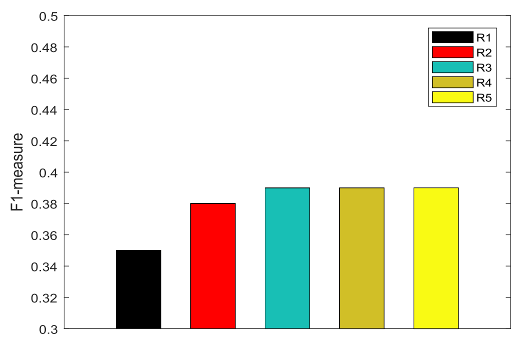


FIGURE 11. The F1-measure performance of different factor combinations.

F1-measure, respectively. For all the three performance metrics, case R1 results in the worst performance among the five combinations of the five factors, which demonstrates that the active degree of users is important for the prediction. Case R3 outperforms the other cases in terms of recall while achieving the similar performance to the other cases on precision and F1-measure, which shows that the preference degree has less impact on the prediction performance than the other four factors. It can be observed that all the five factors of user activity degree, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity have impacts on the prediction. Therefore, the prediction performance will be improved by harnessing all the five factors. In general, the user activity degree and the interest degree have more impacts on the prediction performance than the other 3 factors, which indicates that the active users benefit the most from the proposed method and the users are influenced significantly by the social friends.

V. CONCLUSION

Video prefetching is a viable way to save time during video watching in OSNs. However, prefetching all the videos is resource wasteful if the videos are not watched by the users. Therefore, careful prediction of whether a user will watch a video is critical for efficient video content delivery. In this paper, we studied the problem of predicting whether a video will be watched by a user for efficient video content delivery in mobile social networks. We proposed a Social- and Content-aware Video content delivery Prediction method (SCVP) which captures the intrinsic relationship among users and videos. We designed five metrics to estimate the factors of user activity degree, social tier between users, similarity between videos, similarity between user interest and video content, and video popularity. We used combined prediction for the video content delivery prediction to incorporate the impacts of the five factors on the prediction. The experimental results demonstrated that the proposed method SCVP could effectively improve the prediction performance in terms of precision, recall, and F1-measure.

The proposed method can be generalized to the OSNs which can obtain the characteristics of users and videos and the intrinsic relationship among users and videos. For example, the proposed method can be used in the OSNs of Facebook, Twitter, Weibo, Tecent Weibo, etc., since there exist a large amount of browsing history and social interaction data in these OSNs, which can be well utilized to estimate the five metrics which are the input to the classifier in the proposed method SCVP.

REFERENCES

[1] S. Peng, Y. Zhou, L. Cao, S. Yu, J. Niu, and W. Jia, "Influence analysis in social networks: A survey," *J. Netw. Comput. Appl.*, vol. 106, pp. 17–32, Mar. 2018.

[2] C.-H. Lai, S.-J. Lee, and H.-L. Huang, "A social recommendation method based on the integration of social relationship and product popularity," *Int. J. Hum.-Comput. Stud.*, vol. 121, pp. 42–57, Jan. 2019.

[3] L. Cui, L. Dong, X. Fu, Z. Wen, N. Lu, and G. Zhang, "A video recommendation algorithm based on the combination of video content and social network: Content and social network based video recommendation," *Concurrency Computat., Pract. Exper.*, vol. 29, no. 14, p. e3900, Jul. 2017.

[4] T. Yuan, J. Cheng, X. Zhang, Q. Liu, and H. Lu, "How friends affect user behaviors? An exploration of social relation analysis for recommendation," *Knowl.-Based Syst.*, vol. 88, pp. 70–84, Nov. 2015.

[5] S. Nepal, C. Paris, P. A. Pour, J. Freyne, and S. K. Bista, "Interaction based content recommendation in online communities," in *Proc. Int. Conf. User Modeling Adapt. Personalization*, Berlin, Germany, Jun. 2013, pp. 14–24.

[6] T. Zhu, B. Wang, B. Wu, and C. Zhu, "Maximizing the spread of influence ranking in social networks," *Inf. Sci.*, vol. 278, pp. 535–544, Sep. 2014.

[7] J. Son and S. B. Kim, "Content-based filtering for recommendation systems using multiattribute networks," *Expert Syst. Appl.*, vol. 89, pp. 404–412, Dec. 2017.

[8] X. Yang, Y. Guo, and Y. Liu, "Bayesian-inference-based recommendation in online social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 4, pp. 642–651, Apr. 2013.

[9] J. B. Schafer, F. Dan, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," in *Proc. Adapt Web*, Berlin, Germany, Jan. 2007, pp. 291–324.

[10] Y. M. Li, H. W. Hsiao, and Y. L. Lee, "Recommending social network applications via social filtering mechanisms," *Inf. Sci.*, vol. 239, no. 4, pp. 18–30, Aug. 2013.

[11] L. Li, Y. X. Dong, C. H. Zhao, and W. J. Cheng, "Collaborative filtering recommendation algorithm combined with user trust," *J. Chin. Comput. Syst.*, vol. 38, no. 5, pp. 951–955, May 2017.

[12] J. Wang and W. Xiong, "Collaborative filtering recommender system based on matrix factorization and its application," *Statis. Decis.*, vol. 35, no. 6, pp. 17–20, Jun. 2019.

[13] E. Rojsattarat and N. Soonthornphisaj, "Hybrid recommendation: Combining content-based prediction and collaborative filtering," in *Proc. Int. Conf. Intell. Data Eng. Auto. Learn.*, Berlin, Germany, Mar. 2003, pp. 337–344.

[14] P. Melville, R. J. Mooney, and R. Nagarajan, "Content-boosted collaborative filtering for improved recommendations," in *Proc. Nation Conf. Artifi. Intell.*, Edmonton, AB, Canada, Mar. 2002, pp. 187–192.

[15] Y.-S. Li, M.-N. Song, E. Hai-Hong, and J.-D. Song, "Social recommendation algorithm fusing user interest social network," *J. China Universities Posts Telecommun.*, vol. 21, pp. 26–33, Jul. 2014.

[16] Z. L. Zhao, C. D. Wang, Y. Y. Wan, J. H. Lai, and D. Huang, "FTMF: Recommendation in social network with feature transfer and probabilistic matrix factorization," in *Proc. Int. Conf. Neu. Net.*, Jul. 2016, pp. 847–854.

[17] Y. Wang, G. Yin, Z. Cai, Y. Dong, and H. Dong, "A trust-based probabilistic recommendation model for social networks," *J. Netw. Comput. Appl.*, vol. 55, pp. 59–67, Sep. 2015.

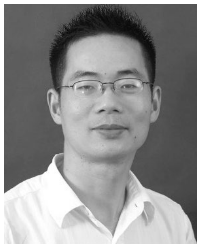
[18] Y. Li, D. Shen, and G. Zhou, "Energy optimization for mobile video streaming via an aggregate model," *Multimedia Tools Appl.*, vol. 76, no. 20, pp. 20781–20797, Oct. 2017.

[19] C. Wu and W. Bai, "Extracting topics based on Word2Vec and improved Jaccard similarity coefficient," in *Proc. IEEE Int. Conf. Data Sci. Cyberspace*, Jun. 2017, pp. 389–397.

[20] Apache. *KDD Cup 2012, Track 1, Predict Which Users (or Information Sources) one User Might Follow in Tencent Weibo*. Accessed: Feb. 3, 2020. [Online]. Available: https://www.kaggle.com/c/kddcup2012-track1

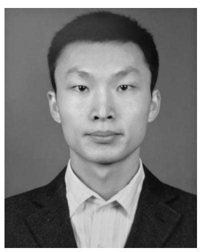
YUQI FAN received the B.S. and M.S. degrees in computer science and engineering from the Hefei University of Technology, China, in 1999 and 2003, respectively, and the Ph.D. degree in computer science and engineering from Wright State University, Dayton, OH, USA, in 2009. He is currently an Associate Professor with the Key Laboratory of Knowledge Engineering With Big Data (Ministry of Education), School of Computer Science and Information Engineering, Hefei University of Technology, China. His research interests include social networks, computer networks, SDN, cloud computing, the Internet of Things (IoT), and computer security.

BING YANG received the B.S. degree in information management and information system from Anhui Medical University, China, in 2017. He is currently a Graduate Student in computer science with the School of Computer and Information Engineering, Hefei University of Technology, Hefei, China. His research interests include social networks, data science, and machine learning.



DONGHUI HU received the B.S. degree from Anhui Normal University, Wuhu, China, in 1995, the M.S. degree in computer science and technology from the University of Science and Technology of China, Hefei, China, in 2004, and Ph.D. degree in information security from Wuhan University, Wuhan, China, 2010. He was a Visiting Researcher with UNCC IN 2013-2014. He is currently a Professor with the College of Computer Science and Information Engineering, Hefei University of Technology. His research interest includes network and information security.

XIAOHUI YUAN received the B.S. degree in electrical engineering from the Hefei University of Technology, China, in 1996, and the Ph.D. degree in computer science from Tulane University, in 2004. After his graduation, he was with the National Institutes of Health, until 2006. He is currently an Associate Professor with the University of North Texas. His research interests include computer vision, data mining, machine learning, and artificial intelligence. He was a recipient of the Ralph E. Powe Junior Faculty Enhancement Award, in 2008, and the Air Force Summer Faculty Fellowship, in 2011, 2012, and 2013.



XIONG XU received the B.S. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China, in 2007 and 2012, respectively. He is currently a Research Assistant with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (CEMEE). His research interests include signal processing and machine learning.