IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 18, 2019, accepted March 6, 2019, date of publication April 12, 2019, date of current version April 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2907494

CMBPR: Category-Aided Multi-Channel Bayesian Personalized Ranking for Short Video Recommendation

XICHEN WANG, CHEN GAO, JINGTAO DING, YONG LI, (Senior Member, IEEE), AND DEPENG JIN, (Member, IEEE)
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Corresponding author: Yong Li (liyong07@tsinghua.edu.cn)

This work was supported in part by The National Key Research and Development Program of China under Grant 2017YFE0112300, in part by the National Nature Science Foundation of China under Grant 61861136003, Grant 61621091, and Grant 61673237, in part by the Beijing National Research Center for Information Science and Technology under Grant 20031887521, and in part by the Research Fund of Tsinghua University–Tencent Joint Laboratory for Internet Innovation Technology.

ABSTRACT Due to the viewing convenience for social media users' fragmented time, the short video has become a new carrier for users' network demands on information spread, news reading, social contact, entertainment, and leisure. Therefore, short video recommendation is one of the most important research topics in social media. Current short video recommendation algorithms mainly focus on detecting user's social attributes, developing cross-domain information and so on, few researchers combine video category information and multi-behavior information together. This paper proposes a content-based recommendation algorithm **C**ategory-aided **M**ulti-channel **B**ayesian **P**ersonalized **R**anking (CMBPR) for short video recommendation, which integrates users' rich preference information by considering the difference among both different video categories and different user interactions. The experimental results demonstrate the effectiveness of the CMBPR video recommendation algorithm, which achieves a significantly higher recommendation accuracy than the traditional video recommendation algorithms and solves the influence of the "Long Tail" effect.

INDEX TERMS Video recommender system, Bayesian personalized ranking, long tail, sampling method.

I. INTRODUCTION

As a comprehensive information carrier, video always plays a leading role in the media industry. Especially with the popularity of mobile internet technology, there are 788 million Chinese mobile netizens by 2019 according to China Internet Network Information Center (CNNIC). Besides, short video applications are rising rapidly, 74.1% mobile netizens use short video applications to meet their entertainment needs in fragmented time [1]. Therefore, short video, especially a new form of user-generated short video, is gaining increasing enthusiasm due to its short-length and viewing convenience on mobile platforms. The short video recommender system is an important function in social media platform, for it not only assists the users to explore the new videos they may favor, but also helps the social network to enhance users' stickiness. Since short video recommendation is very challenging, a continuous effort has made for this task [2]–[5].

Compared with the traditional long video, such as videos of Netflix which is a website for user renting and rating movies and expert-generated videos of American Broadcasting Company, the short video has the following three characteristics:

- Short videos are often made by amateurs via video capturing devices, digital cameras and video editing software. Differing from professional videos, short videos are often found on content sharing website such as social networks, web search and YouTube [6].
- The length of short video is often within 10 minutes. Cheng *et al.* found that 97.9% short videos of YouTube are within 600 seconds, and 99% are within 700 seconds while traditional long videos are typically 1-2 hour [7].
- The low production cost of short videos greatly enriches their content, making them cover various categories. Besides, the distribution of video category is highly skewed compared with traditional long videos [7]. This indicates that the users' preference on short videos are more diverse.

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

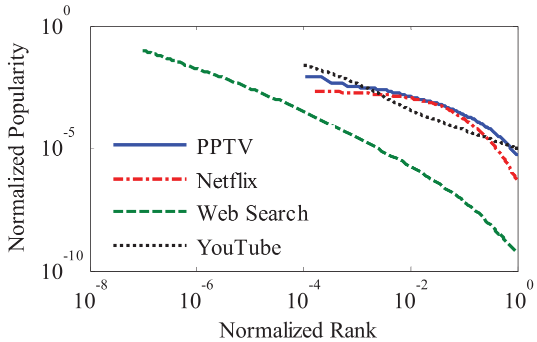

FIGURE 1. **The popularity distribution of PPTV, Netflix, Web Search and YouTube.**

The "Long Tail" effect describes the phenomenon that niche category short videos can grow to become a large share of total short videos to be recommended. The Long Tail effect is more important for short video recommendation problem compared with long video recommendation problem. FIGURE 1 plots the videos' popularity distribution of Netflix, PPTV, YouTube and Web Search [8]–[10] where the x-axis stands for normalized rank and the y-axis is the normalized popularity. The popularity of video is the total number of sessions referred to the video during the whole measurement period. Netflix, PPTV are websites featured as long videos such as expert-generated videos and movies while YouTube and Web Search are websites featured as user-generated short videos. We can clearly find that lines of PPTV and Netflix are more curved than those of YouTube and the web search. A more straight line here means a larger "long tail" which will accumulate more user demand. Therefore, the niche content of short videos have more significance than niche content of long videos [10]. Once this niche category is assembled, it can still attract a large scale of users. Therefore, the study on overcoming the Long Tail effect of short video recommender system is of importance and we need to propose a short video recommender system to tackle the Long Tail effect.

Although the Long Tail effect is very prevalent in short video recommendation scenario, few current recommender systems focus on this problem. This is because the key to solve the Long Tail effect is to identify users' special preference with high precision even the preference is very niche. However, differentiating users' demand with high precision in a large-scale dataset is very challenging for the following reason: the short videos in the "long tail" part usually belong to the niche content that only get a few accesses by the other users, therefore the latent vectors for these short videos of niche category will be highly imprecise due to the data sparsity and inadequate embedding, leading to incorrect personalized ranking. Besides, the users in social networking will produce complicated feedbacks to short videos, which greatly increases the difficulty of this problem. We need to make clear the correlation between users' preference degree embodied by multi-behaviors.

To address the above mentioned limitations, two types of information should be considered to better exploit the user preference behind the observed video view logs. First of all, users' personal preference is strongly related to video category. Therefore, video categories can effectively collect users' potential interested videos even if the potential interesting videos are accessed by a few users. Secondly, today's social media will collect multiple behaviors of users, and these behaviors encode different levels of user preference. In the case of Weibo, a Twitter-style social media, users usually have three kinds of operations to interact with short videos, including view, reply and retweet. Therefore, we intend to exploit different types of feedbacks in our proposed model. With the above two information, the recommender system will both consider user side information and item side information. Consequently, it will identify users' preference better.

Content based recommender system [11]–[15] can directly identify users' preference on video category via users' browsing history, however it lacks the ability to identify users' different level of preference on video categories that the user has browsed. Meanwhile, multi-behavior recommender system can differentiate the users' degree of preference via users' different feedbacks to short video but it lacks category information. Existing approaches for multi-behavior recommendation basically divide into two kinds, the first kind is collective matrix factorization (CMF) model and the second is based on Bayesian Personalized Ranking (BPR). We need to incorporate video category information with multi-behavior recommender system to better differentiate users' preference for different video. Despite there are CMF based methods that combining content information and multi-behavior information together [16], [17], CMF generates the user and item latent vectors based on pointwise method, making it difficult to interpret the latent vectors. Besides, CMF lacks the ability to define multi-behaviors' priority which is an import trait of how social networking users browse the short videos. BPR [18] utilizes pairwise ranking method based on preference orders and multi-channel BPR can assign different feedback level with different priority [19]. Pairwise ranking means each individual item is compared directly against the others so as to emerge with a ranking from highest to lowest [18]. Moreover, if we focus more on comparing the pairwise preferences relation of the same category videos, the computational efficiency will be largely enhanced by increasing the opportunities of sampling users' potential preferred videos as preferrence pairwise and getting trained as latent vector, consequently more detailed information about users' interesting degree for same kinds of videos can be obtained. Therefore, BPR based model is more suitable for short video recommender system. However, existing BPR models mainly focus on combining with multiple sampling criteria [20]–[23] or with relational data [24]–[28], few BPR algorithms combine video category information and users' multiple behavior information at the same time. For instance, despite Guo *et al.* propose GcBPR to utilize generated data to resolve the data sparsity issue, they do not incorporate the item information and user behavior information together into BPR framework [29]. The difficulty lies in that we need powerful modeling capability to capture the multi-modal data with the BPR model when we merge multi-behavior and video content information into it. Besides, with the increase of system complexity, we need to design reasonable hyperparameters to control algorithm performance while not to increase the difficulty of parameter adjustment process.

In this paper, we proposed a novel Category-aided Multi-channel Bayesian Personalized Ranking (CMBPR) approach, which combine content-based recommendation method and multi-channel BPR method together. The Category-aided sampling approach is for more precise preference comparison of users' potential interesting category of short videos, and multi-channel sampling approach of three different implicit feedback is for more detailed information of users' preference degree of a specific video. In summary, the advantage of our algorithm is that we convert a linear process of detecting users' preference to a comprehensive and nonlinear process. The users' preference is obtained from different dimensions (video category and multi-behavior) and then we decide the users' preference degree on the basis of the multi-behavior's correlations. In this way, once a viewed video information is collected, our CMBPR can fit the users' preference with high precision and prioritize the potential favored video as preference pairwise even if they belong to niche category. Correspondingly, niche category short video will get more accurate latent vector and CMBPR can resolve the Long Tail effect.

The main contributions of this paper can be summarized as follows:

- For the modeling of video category information, we propose a Category-aided sampling approach based on Bayesian Personalized Ranking to capture user's preference. Through biased sampling between users' inner-class category preference and outer-class category preference, we exploit more detailed preference information and greatly facilitate in solving the "Long Tail" effect.
- For the modeling of users' different behavior when browse short videos, we propose the 2MC-BPR algorithm, extending the BPR sampling model by utilizing the impacting difference among users' different behavior. This is more suitable for real-world short video recommendation of social media for users will not only browse short videos, but also reply and retweet.
- We are the first to our knowledge that merge the Category-aided sampling strategies and 2MC-BPR algorithm together into BPR based framework as CMBPR, greatly improving the algorithm performance for it can better capture users' preference with multi-modal data while the algorithm complexity and parameter adjustment difficulty do not obviously increase.

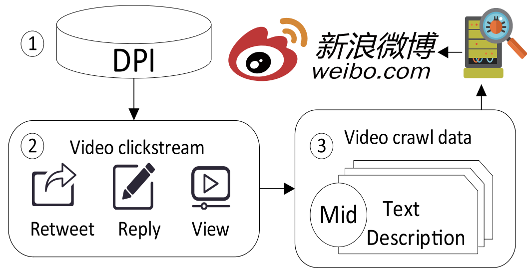

FIGURE 2. Illustration of the raw data collection process.

- We conduct extensive experiments on a large-scale real-world short video dataset of social media, demonstrating that our CMBPR method outperforms state-of-the-art methods by a large margin. Moreover, we show that CMBPR can cope with the "Long Tail" effect in short video recommendation scenario for it shows higher prediction accuracy on niche category short videos.

The rest of the paper is organized as follows: In Section II, we give a thorough description of our dataset. Section III introduces some preliminaries and presents our proposed method. In Section IV, experimental settings and evaluation protocols are presented. Section V briefly reviews the related work. Finally, Section VI provides some concluding remarks.

II. PRELIMINARY STUDY

In this section, we will give a thorough description of our utilized dataset. In addition, we will demonstrate the relationship between video category and user preference, and validate the Long Tail effect in video browsing.

A. DATA DESCRIPTION

Our data is from Sina Weibo, one of the largest Chinese Twitter-style social network platforms in China. Weibo is the most popular online social network, of which monthly active users reached 411 million, and daily active users reached 184 million by the end of the fourth quarter of 2018 [30]. Weibo strengthens the function of social media and brings new eras of short videos. For instance, Miaopai as a short video sharing app embedded in Weibo like Vline, Keek, and Instagram, dominating Weibo' s traffic via competitive strategy such as external video sharing, inviting celebrities to create original videos, and to provide valuable content for Weibo.

We collected our raw data from one of the largest Internet Service Providers in Shanghai, spanning a period between April 21st and April 26th, 2016. The records in the raw data contain detailed information about users' Weibo access logs in the 3G cellular network, including anonymous user IDs, timestamps, HTTP headers and base station IDs. From these HTTP headers, we can extract the users' video browsing records.

We find that there are substantial Weibo video interactions records in the dataset and some of them can be extracted by special URLs. Through packet capture, we identify the

TABLE 1. Classifying video interactions into Retweet, Reply and View.

Behavior	Keywords in URLs	User	Record
Retweet	statuses/repost_timeline; statuses/repost	6k	7.5k
Reply	comments/show; comments/allowed	11.2k	20k
View	statuses/video_adext; statuses/video_timeline; statuses/show	61k	76k

TABLE 2. Top 10 words of video classification.

Category	Top10 words
Fun	laugh; haha; jokes; funny; boy friend; game; friendship; what the hell; cold joke; absolute
Music	sun; lie; sing nicely; descendant; G-Dragon; youth;version; concerts; Shanghai; Jay Chou
Baby&Pet	baby; master; a cat; dog; Husky; puppies; catman; little sister; golden retriever;adorable
Star&Show	Jacky Xue; exo; video; Chinese Running Man; endeavor; interview; shot; Jackson Yi; run; fans
Society	subway; student; male; teacher; candid camera; remember; cellphone; police; pervert; Nanjing
Life	simple; how; foodie; make up; like; everybody; tutorial; crayfish; delicious; greedy
Movie	Song Joong Ki; movie; movie trailer; exposure; The Mad Monk; captain; shoot; exclusive; Benchang You; actor
Sports	move; body building; train; playoffs; byebye; lose weight; advice; today; everyday; keep

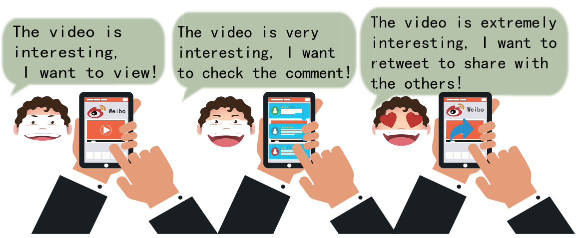

FIGURE 3. Illustration of three different feedback levels.

URLs of Retweet, Reply, and View. TABLE 1 lists the basic statistics of the Weibo video interactions. Besides, there are unique MIDs representing different videos in corresponding URLs. Therefore, we can crawl text description of each MID's matching video to obtain the category of videos, this process is illustrated in FIGURE 2. TABLE 2 lists the categories of videos and top 10 descriptive words of each category. We need to pay extra attention that there are 0.8k videos without any text description, we categorize them as Others. In order to improve the data quality, we remove the users who watch short videos no more than twice, then we obtain a dataset consisting of about 10k users and 5k short videos.

From TABLE 1, we find that there are three interactions representing three different feedback levels. Considering the scenario of browsing video illustrated in FIGURE 3, we first view the video if we find it interesting, then we check the reply of the short videos to find the others' opinion. Hence, the interesting level embodied by reply is higher than view. Similarly, if we find the video extremely interesting and thus want to share with the others, we will retweet the video to expand the video's influence. Therefore, interesting level embodied by retweet is higher than reply. Our dataset can also be an indication because the amount of retweet is lower than reply, and the amount of view is the highest. Besides, for each user, the videos he retweets are often included in the videos he replies. In summary, the priority order of the different feedback levels is Retweet>Reply>View.

B. USER INTEREST AND VIDEO CATEGORY

The user's interaction process is the user's real demand for the social network. Therefore, we need to collect user historical behavior data to better understand users' interest. According to Section II-A, Weibo classifies the video into 9 categories and the users typically have three levels of feedback on short videos, including Retweet, Reply, and View. Here, we define that the level of feedback and the ratio represent users' interest.

To validate our assumption, we randomly choose five users and demonstrate their percentage of visited short video categories in FIGURE 4. It is easy to find that different users have different preferences on video category. For instance, Fun videos occupy 60% for User 1, while User 2 prefers to view Baby&Pet videos. User 3 has more interest in Life video, and User 4 is fond of Movies and Others. User 5 concentrates most on Other videos. It is apparent that different users do have different interest and can be reflected by the category distribution of videos he has seen.

FIGURE 5 lists the number of all the categories over all interactions and short videos. First, we find that distributions of different level actions and videos along categories looks similar. Then, we can find that the distribution of videos with different categories is highly skewed, ranging from 5.82% @Sports to 25.2% @Fun in FIGURE 5[a]. Thirdly, we find that the distribution of the total number of videos does not necessarily proportionate to the distribution of View, Reply and Retweet. For instance, for some video categories of larger number, such as Fun, Others, Life, their videos' proportions are higher than or equal to corresponding proportion of View, Reply and Retweet. For instance, percentage of Fun video is higher than percentage of Fun for View by 0.3%, percentage of Fun for Reply by 7% and percentage of Fun for Retweet by 8%. Meanwhile, for some smaller amounts of videos, such as Baby&Pet, Star&Show, Sports and Music, their videos' proportions are lower than corresponding proportion of View, Reply and Retweet. This phenomenon is especially evident on the distribution of Reply and Retweet. Because different level of actions stand for varying degrees of preference, this phenomenon can demonstrate that the users will show more their special preference over higher level interactions. This also indicates that for some niche categories of short videos, minority of users will demonstrate their distinct tendentiousness to the certain category.

To find whether each category has a group of enthusiasts, we conduct BiCluster on user-category matrix [31], the row

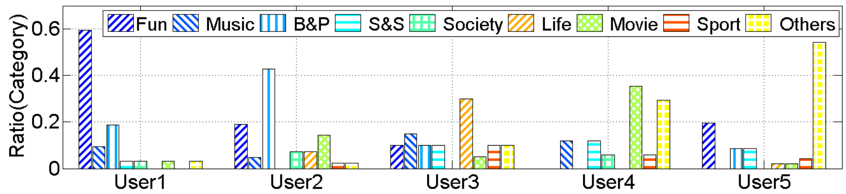

FIGURE 4. Short videos' category distribution of different users based on several random samples.

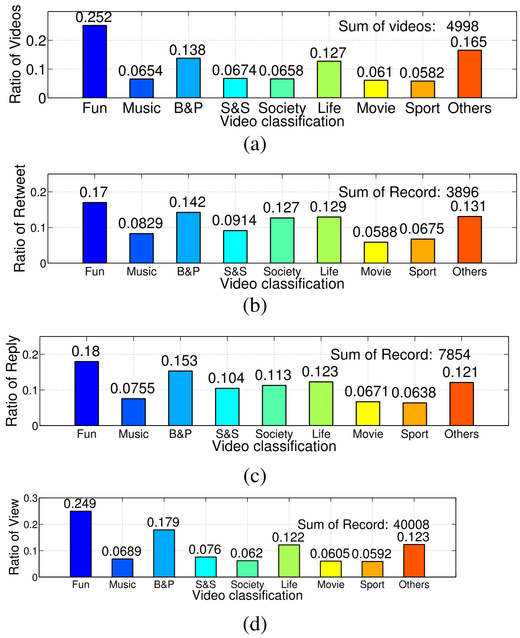

FIGURE 5. Ratios of total number distribution of videos over different categories and ratios of total number distribution of Retweet, Reply and View over different video categories. (a) Ratio of video number along different categories. (b) Ratio of Retweet along different categories. (c) Ratio of Reply along different categories. (d) Ratio of View along different categories.

of which stands for the user's category distribution of viewed videos and the column stands for the index of users. The result is shown in FIGURE 6. In this algorithm, if each row and column belong to the same BiCluster, the rearranged matrix will render the BiCluster as diagonal blocks. The diagonal BiClusters of this matrix have higher average value than the other ranks. We can observe from FIGURE 6 that almost each category has a group of users that favor this category most, despite that Fun block has a wider group.

With the above analysis, we expound the relationship of user interest and video category. Except for user interest, another challenge for recommender system is the Long Tail effect. We will discuss this problem in the following.

C. THE LONG TAIL EFFECT OF SHORT VIDEOS

The Long Tail effect means that a long tail of some distributions of numbers is the portion of the distribution having a large number of occurrences far from the "head" or central part of the distribution. In other words, a relatively small number of items account for a disproportionately large fraction of total consumption while the tail is relatively heavy in aggregate [32]. Since the Long Tail effect often forms power law, we usually depict it with complementary cumulative distribution function(CCDF).

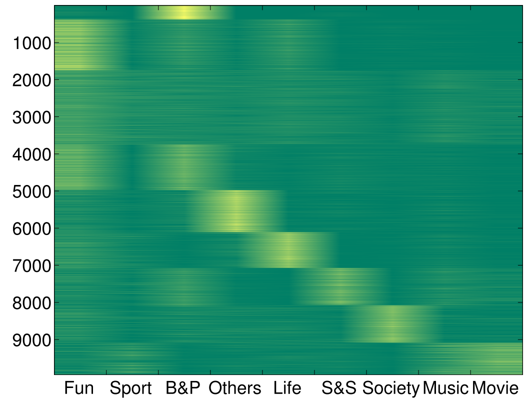

FIGURE 6. BiCluster of users and video category.

The Long Tail effect is in consistent with two hypotheses. One is that most users prefer popular videos while only a minority want niche categories; the other is that most users have their special preference, therefore they will consume both popular and special content [8]. The key to tackle the Long Tail effect is to differentiate popular demand and special demand of users. For a recommender system, this means to identify users' special preference even it is niche content and then recommend popular items and special need items together.

Previous analysis in Section I shows that short videos have larger "tail" than the long video. Therefore, we want to depict our dataset with Complementary Cumulative Distribution Function (CCDF) to visualize the "long tail" of our dataset. The x-axis stands for number of videos and the y-axis measures the proportion which is greater than x. If the CCDF reflects the characteristic of power law distribution, then we will claim that the dataset possesses the "long tail". Apparently, according to FIGURE 7, the dataset has large, straight "tails" and thus keeps the power law across all the interactions and all categories. Besides, there are other interesting observations. As shown in FIGURE 7(a), the lines stand for the videos of Fun and of Baby&Pet have larger "tail" than the others. A larger "tail" means some users view video of Fun and of Baby&Pet with higher repetition. However, this may be caused by the large total number proportion of Fun and Baby&Pet, which are 25.2% and 13.8%

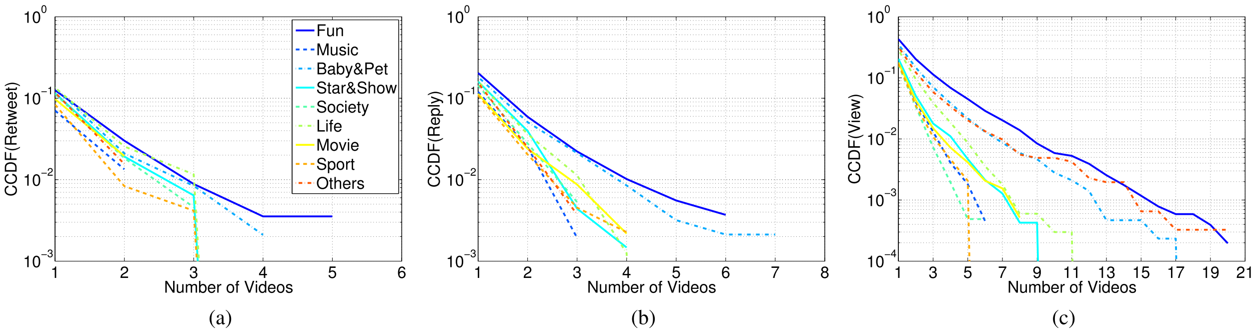

FIGURE 7. CCDF of Retweet, Reply and View times. (a) CCDF of Retweet. (b) CCDF of Reply. (c) CCDF of View.

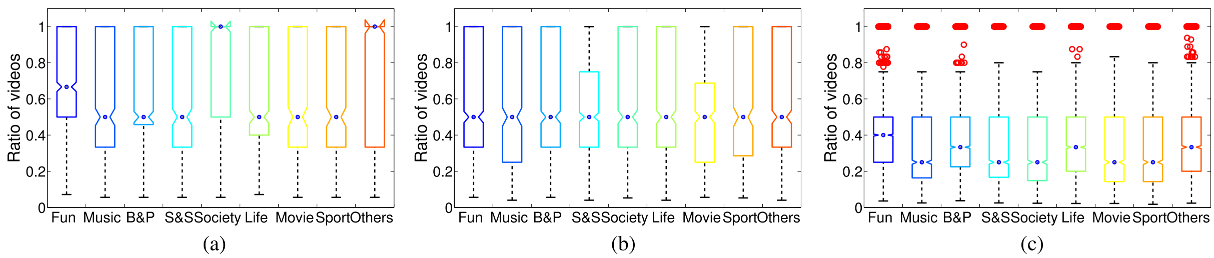

FIGURE 8. Ratio distribution of Retweet, Reply and View times over different categories. (a) Ratios of retweeted videos over different categories. (b) Ratios of replied videos over different categories. (c) Ratios of viewed videos over different categories.

respectively. We can also find that the "tails" have different gradients in FIGURE 7(c). A steeper "tail" refers that the smaller proportion of users will pay attention to this category. Therefore, categories of Sport, Movie, Music and Star&Show are all niche categories that attract the attention of the minority while categories of Fun, Other and Baby&Pet are popular categories. According to the definition of the Long Tail effect, the niche categories are the main component of the "long tail".

After demonstrating our dataset also have "long tail", we want to look further on whether the users' preferred video category is highly skewed when they retweet, reply and view. Therefore we count the total number of each category's videos when the users retweet, reply and view and normalize them as ratios. We then demonstrate the ratio of every user in terms of the three level of interactions in FIGURE 8. The boxes here can represent the skewness of a certain video category. In other words, the higher Quartile and Median of a box, the users show more interests in this category. For instance, the median of the box for Music in FIGURE 8[a] is 0.5, standing for that median level users' retweet videos of Music occupy 50% of all their retweeted videos. Considering that the total number of Music category only occupies 6% of all videos, the tendentiousness presented by users who retweet Music short videos is very clear. We have to acknowledge that the boxes of Fun and Other have higher median and quartile, this may due to the fact that the Retweets in the dataset are very sparse, and most users only retweet a few times. Besides, the total number of Fun and Other short videos are much higher than the niche categories. The skewness of video category distribution can also be found in FIGURE 8(b). For instance, despite the videos of Fun and Others occupies 25% and 16.5% of all videos respectively, their medians are almost equal with the other categories. This indicates that users show strong tendentiousness over niche categories when they reply short videos. Music, Movie and Sport get least quantities in all categories, and their boxes have smaller lower quartiles compared with Fun and Other. However, we can still find that some users who reply videos of Music and Sport show strong viscosity for their upper quartiles are 100%. In FIGURE 8(c), the distribution of medians of all box roughly coincides with the proportion of total number of videos in each category. However, the upper quartiles of each category are almost same. This indicates that each category has a group of users prefer to see this kind of videos most.

In summary, there are three characteristics of our dataset:

- The dataset has obvious Long Tail effect.
- Categories of Sport, Movie, Music and Star&Show are all niche categories that attract the attention of the minority while categories of Fun, Other and Baby&Pet are popular categories. The niche categories are the main component of the "long tail".
- The users can show their distinguishing preferences through the ratio distribution of different categories of Retweet, Reply and View.

Therefore, differentiating the users' preference over categories and interactions in finer-grain will improve personalized ranking accuracy of short videos and solve the Long Tail effect.

III. METHODOLOGY

In this section we will firstly give a brief view of the problem definition and proposed solution, then we review the BPR algorithm, thirdly we discuss its drawbacks in motivation part, and finally we propose the CMBPR and present the algorithm's superiority.

A. PROBLEM DEFINITION

We define U as the user set and V as the video set. In our learning scenario, the users have three levels of implicit feedbacks except for non-viewed videos, which are $L = (L_1, L_2, L_3 ... L_n)$ kinds of different feedbacks. Besides, we categorize our video dataset V as $\{V_1, V_2 ..., V_l\}$, where l is the video's category. Our goal is to optimize the sample procedure, biased sample potential preferred video category and give different level feedback according to users' multi-level behavior, so as to differentiate the users' preference with fine-grain, recommend top-N videos that can solve the Long Tail effect.

B. BAYESIAN PERSONALIZED RANKING

BPR is a real ranking algorithm by maximizing the likelihood of observed pairwise preference ranking relations and seamlessly combines the existing MF and KNN models to achieve personalized ranking. As we define U as the user set and V as the video set, for item $j \in V$ which the user $u \in U$ has viewed (refers as Positive) while item $i \in V$ that the user u has not viewed (refers as Negative), we can define $p(j >_u i)$, which means the user u prefers item j over item i.

We assume Θ as the model parameter, BPR is to find the mostly likelihood parameters that meet the $P(\Theta| >_u) \propto p(>_u |\Theta)\, p(\Theta)$. where $>_u$ is the pairwise preference ranking of users' preference. We define p as the user' s feature vector and q as the item' s feature vector in Θ, α as the learning rate and λ as the regularization rate, correspondingly, the model for the iterative solution is as follows:

$$p_u \Leftarrow p_u - \alpha \left(\frac{1}{1 + e^{p_u q_i - p_u q_j}} (q_i - q_j) + \lambda p_u \right), \quad (1)$$

$$q_i \Leftarrow q_i - \alpha \left(\frac{-p_u}{1 + e^{p_u q_i - p_u q_j}} + \lambda q_i \right), \quad (2)$$

$$q_j \Leftarrow q_j - \alpha \left(\frac{p_u}{1 + e^{p_u q_i - p_u q_j}} + \lambda q_j \right), \quad (3)$$

where $p(j >_u i) = p_u q_i - p_u q_j$.

As we can see, the accuracy of the algorithm depends much on the ranking process of selecting preference pairwise $j >_u i$ and the learning rate α. Although we can make sure that the learning rates are chosen suitable enough for the right direction gradient descent, inappropriate preference pairwise i, j will lead to slow convergence and low Hit Rate (HR). It might also be noted that the author also suggests that the stochastic gradient descent (SGD) algorithm based on bootstrap sampling performs better. Therefore, how to choose the appropriate preference pairwise $j >_u i$ will greatly impact on the prediction result.

Despite we can make sure that j stands for the user's preference, which usually belongs only a very sparse subset compared with V, while i is usually chosen by random, the individualized need of the users is neglected. Some paper proposes that BPR can sampling by the items' popularity [33], nonetheless, this method ignores the specific needs of a minority to some extent. Consequently, the videos of the "Long Tail" are neglected. Therefore, we need to develop an effective sampling strategy to select meaningful preference pairwise to obtain better personalized ranking performance.

C. MOTIVATION

BPR assumes that the dataset only has two kinds of data for a user, which is the videos that the user has seen and has not seen. However, the dataset can be given a much more comprehensive definition.

From the view of interaction type, in the real world short video browsing scenario, users often make tendentious browse with multi-relational behavior (Retweet, Reply and View). This results in a variety of feedbacks and biased data. Sampling methods of BPR, which learn pairwise preference of items only from positive and negative level, are unable to utilize the information of multi-channel feedbacks, therefore we need to give multilevel pairwise preference. Besides, from the view of video category, different interests make great varieties in the preferred video category between the users. The sampling methods of BPR, which only learn pairwise preference of items by random, is unable to utilize the information of video category, therefore we need to give non-uniform pairwise preference. For the above reasons, traditional BPR sampling method needs to be updated. We should propose a sampling method that combines multi-feedback solution and multi-category solution together.

In order to overcome these shortcomings of BPR, we define the following principles:

- Users do not change their preferences of video category for a short time.
- What category of a video that the user has seen stands for the user's potential interest in this kind of video.
- Intensive comparison of video of the same category will lead to better ranking.

Based on these assumptions, we propose Category-aided Multi-channel Bayesian Personalized Ranking to combine multi-relational feedbacks sampling and category-aided sampling together to overcome this problem. This method learns the preference pairwise by extending to multi-channel of interactions and selective learning what the user most interest in, implicitly enhances the probabilistic of comparing the pairwise that users are most likely interested in. Therefore, our proposed algorithm is expected to have better performance.

D. CMBPR

In this part, the paper proposes the Category-aided Multi-channel Bayesian Personalized Ranking to improve the performance of top-N recommendation. We explain the principle that how CMBPR takes advantage of multi-level feedback and fine granularity of video classification. In addition, we investigate the key factors that result in the success or

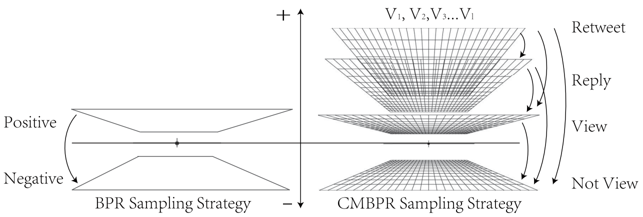

FIGURE 9. Illustration of the change of sampling method.

failure of the algorithm. Beyond that, we give the definition of CMBPR+, an extension of CMBPR based on CMBPR's sampling character to achieve better performance.

1) THE ADVANTAGE OF CMBPR

We now express formally how CMBPR makes use of multi-feedback levels and multi-category. Like shown in FIGURE 9, in our learning scenario, the users have three levels of implicit feedbacks except for non-viewed videos, and we categorize our video dataset V as $\{V_1, V_2 ..., V_l\}$, where l is the video's category. Therefore, we have to deal with a two-tier problem: the first is how to combine three levels of implicit feedbacks; the second is how to utilize the category information to solve the Long Tail effect.

For the first problem, we propose 2MC-BPR, which is the updated version of MC-BPR [19]. Each level of feedback stands for users' interest level in varying degrees, which can be extracted from the user's operation to the short videos. In BPR, L only has two levels. The higher level is for short videos that the user has viewed and the lower level is for the short videos that the user has not interacted with. In our scenario, we extend L to four different levels according to the characters of our dataset, the highest level is Retweet, Reply, View and Not View are in descending order. In this way, we extend pairwise preference items to 6 different kinds of combination. This can greatly improve the iterative update information for Equation 1, 2, 3.

In addition to extending richer update information of pairwise preference items, CMBPR also engages in more accurate update information of pairwise items. For the convenience of understanding, we first describe two definitions before introducing our new assumption in the following.

Definition 1: Inner-class Pairwise The inner-class pairwise is defined as $p\left(i\left(l\right)\right) >_u j\left(l\right))|\Theta)$. It is the preference probability of a user on item $i\left(l\right)$ over item $j\left(l\right)$, where l is the category of item i and j.

Definition 2: Outer-class Pairwise Correspondingly, the outer-class pairwise $p\left(i\left(l\right)\right) >_u j\left(\bar{l}\right))|\Theta\right)$ stands for preference probability of a user on item $i\left(l\right)$ over item $j\left(\bar{l}\right)$, where i and j are from different categories.

We assume that the inner-class pairwise preference is more important than the outer-class pairwise. Therefore, when we choose the video i, we judge the video's category first, and then we choose the video j of the same category with higher probability.

The next important part of the algorithm is to define the sampling ratio of inner-class pairwise and outer-class pairwise. We want to testify whether the change of ratio will affect the HR of CMBPR. The algorithm is demonstrated in Algorithm 1. In this algorithm, for a certain item i with category l_i, the sampling distribution can be described as:

$$p\left(j|u\right) = \underbrace{\beta_{l_i} p\left(j|l_i, u\right) p\left(l_i|u\right)}_{inner-class} + \underbrace{\sum_{l \in L_u|\{l_i\}} \beta_l p\left(j|l, u\right) p\left(l|u\right)}_{outer-class}, \tag{4}$$

where $L_u|\{l_i\}$ is the set of items rated by user u, excluding the item i, which is deemed as inner-class. In this equation, $\beta_{l(i)}$ is set to be hyperparameter described as "*paralabel*" and$\beta_{l(i)} + \sum_{l \in L_u|\{l_i\}} \beta_l = 1$. Then the formalization of the sampling distribution is as follows:

1) Random choose a user u.
2) Random choose a feedback channel U of the user u.
3) Sample item i from this feedback channel U.
4) Judge the category of item i.
5) Generating a random number below 1, if the random number is below than hyperparameter *paralabel*, then sample from inner-class $p\left(j|l_i, u\right) p\left(l_i|u\right)$, otherwise sample uniformly from outer-class.

The step 1 to step 5 all can be performed in O(1), therefore, the computational complexity of the original bootstrap stochastic gradient descent algorithm does not increase.

2) THE STOCHASTIC GRADIENT DESCENT METHOD CMBPR

However, when we do experiments on this algorithm, we find that there is no obvious increase of HR. We go a step further to make out the sampling procedure and find it owe to the bootstrap stochastic gradient descent method. Here we will give a detailed definition of several stochastic gradient descent of BPR [18].

Definition 3: Full stochastic gradient descent It will have the full gradient over all training data computed and then the model parameters are updated with the learning rate α in each step.

Definition 4: Stochastic gradient descent It will traverse the positive item and each $i >_u j$ an update is performed.

Definition 5: Bootstrap stochastic gradient descent It will abandon the idea of full cycles through the data and chooses $i >_u j$ randomly.

The author of BPR suggests that choosing bootstrap stochastic gradient descent instead of stochastic gradient descent will lead to faster convergence. However, in our algorithm, we will sample more negative items from inner-class, which means the preference pairwise is more efficient, therefore the bootstrap stochastic gradient descent will not work and the stochastic gradient descent will improve the HR as it will transverse all items of the same level feedback and give effective gradient descent direction.

Algorithm 1 *CMBPR*

Input: Retweet feedback U_{Rt}, Reply feedback U_{Rp}, View feedback U_V, No feedback $\tilde{U}$, Video classification V, iteration N

Output: The learned model parameter Θ

Randomly initialize Θ
iteration $T \leftarrow 0$
for $T \leftarrow 1$ to N **do**
    $u \leftarrow$ draw a random user from U
    $U \leftarrow$ draw a random feedback level
    for $i \leftarrow 1$ to $|U|$ **do**
        $l \leftarrow$ acquire the category of i
        $\bar{U} \leftarrow$ random draw a feedback level that is lower than U
        if $\bar{U} == \tilde{U}$ **then**
            $rq \in (0, 1)$
            if $rq < paralabel$ **then**
                $j \leftarrow$ draw a random video from V_l of $\tilde{U}$
            else
                $\bar{l} \leftarrow$ get category different from l
                $j \leftarrow$ get item from $V_{\bar{l}}$ of $\tilde{U}$
            end if
            Compute gradients and updated the above parameters
        else
            $j \leftarrow$ draw a random video from $\bar{U}$
            Compute gradients and updated the above parameters
        end if
    end for
end for

3) CMBPR+

Based on the analysis of the previous section, we have a deeper look at our dataset. Naturally, we find that a user may prefer several categories of the short videos, and CMBPR cannot measure the different interest level if the user prefers several categories of videos. To solve this problem, we define the following principles:

- The number of view times of a category stands for the user's preference extent.
- The more view times of a category, the more chance to sample from negative items from the same category can enhance the HR of the algorithm.

Then the formalization of the sampling distribution is as follows:

1) Random choose a user u.
2) Calculates the user's viewing times of each category and normalize the result as $[p_1, p_2, \ldots, p_l]$.
3) Random choose the feedback level U of the user u.
4) For items in this feedback level, begin iteration T.
5) Judge the category of item i as l.
6) Generating a random number below 1, if the random number is below than $p_{i(l)} * paralabel$, then sample from inner-class $p(j|l_i, u)\, p(l_i|u)$, otherwise sample uniformly from outer-class. Noting that we still use *paralabel* as hyperparameter to control sampling ratio.

We define this algorithm as CMBPR+ in Algorithm 2. In this procedure, the time-consuming step is (2), and the other steps can be performed in O(1). However, the users' Retweet, Reply and View behavior are all very sparse. Therefore, the computational complexity does not significantly increase.

Algorithm 2 *CMBPR+*

Input: Retweet feedback U_{Rt}, Reply feedback U_{Rp}, View feedback U_V, No feedback $\tilde{U}$, Video classification V, iteration N

Output: The learned model parameter Θ

Randomly initialize Θ
iteration $T \leftarrow 0$
for $T \leftarrow 1$ to N **do**
    $u \leftarrow$ draw a random user from U
    $p \leftarrow [p_1, p_2, \ldots, p_l]$ draw user's normalize preference distribution.
    $U \leftarrow$ draw a random feedback level
    for $i \leftarrow 1$ to $|U|$ **do**
        $l \leftarrow$ acquire the category of i
        $\bar{U} \leftarrow$ random draw a feedback level that is lower than U
        if $\bar{U} == \tilde{U}$ **then**
            $rq \in (0, 1)$
            if $rq < p_{i(l)} * paralabel$ **then**
                $j \leftarrow$ draw a random video from V_l of $\tilde{U}$
            else
                $\bar{l} \leftarrow$ get category different from l
                $j \leftarrow$ get item from $V_{\bar{l}}$ of $\tilde{U}$
            end if
            Compute gradients and updated the above parameters
        else
            $j \leftarrow$ draw a random video from $\bar{U}$
            Compute gradients and updated the above parameters
        end if
    end for
end for

IV. EXPERIMENTS

In this section, we conduct a serial of experiments to validate our algorithm CMBPR. To comprehensively evaluate our proposed method, we conducted experiments to answer the following research questions:

RQ1: Can CMBPR and CMBPR+ outperforms the other state-of-the-art recommendation methods for multi-channel implicit feedbacks?

RQ2: How do different hyperparameter settings affect CMBPR and CMBPR+?

RQ3: Are CMBPR helpful for the Long Tail effect by recommending more niche short videos to specific user?

A. EXPERIMENTAL SETTINGS

1) DATA PREPARATION

In order to evaluate the performance of CMBPR and CMBPR+ in the scenario of short video recommendation in Weibo, we select users who have done at least one of the three kinds of typical Weibo interactions, including Retweet, Reply, and View. We recommend top-N videos instead of scoring the video due to Retweet, Reply and View are all implicit feedbacks and the data do not have an effective rating for the short videos.

2) TRAINING SET AND TEST SET

The aim of our algorithm is to predict the video that the user most likely to watch. For each user, videos in his View records usually include of videos in his Reply and Retweet records for a user has to view before reply and retweet in the process of browsing Weibo. Besides, considering the logical sequence of browsing Weibo, we should predict the future with the past interaction. Therefore, we apply Leave One Out Method to generate training set and test set.

3) EVALUATION CRITERIA

In the evaluation process, for each user u, we score his preference of all items with $s_{u,i}$. Then we acquire the score list. After removing the videos in the user's training set, we obtain the top-N recommendation list by finding the maximum N scores. We will describe top-N as "@N" in the following paper. Besides, we use Hit Rate(HR) and Normalized Discounted Cumulative Gain (NDCG) to evaluate all the algorithms.

HR@N: *HR@N* is defined as the ratio of the number of users for which the model successfully recall from the test item in the top-N list and the number of all users [34].

$$HR@N = \frac{\#hits}{\#users} \tag{5}$$

where #users is the total number of test users and #hits is the number of users that our algorithm successfully hit with the size-N recommendation list.

NDCG@N: *NDCG@N* is the ratio of *DCG@N* to the Ideal *DCG@N*. *NDCG@N* is a measure of ranking quality where the higher *NDCG* value indicates better learning performance, Given the ranking list $[v_1, v_2, ..., v_N]$, where v is the recommended video and k is the ranking position [35].

$$NDCG@N = \frac{1}{Z}\sum_{i=1}^{N}\frac{2^{t(i)}-1}{log\,(1+i)} \tag{6}$$

here $DCG@N = \sum_{i=1}^{N}\frac{2^{t(i)}-1}{log(1+i)}$, and $t\,(i)$ stands for the given value of the video at position i, meanwhile Z is the Ideal $DCG@N$, a normalization value of the top N videos of a perfect ranking. In NDCG, $t\,(i)$ usually ranks from 0 to r where r can be set. In our paper, we set r as 1.

B. BASELINES

We use the following algorithm to demonstrate CMBPR and CMBPR+'s performance. We divide the control group as two kinds: comparison with BPR's extensions and comparison with state-of-the-art multi-relational algorithms and several basic algorithm.

1) BPR'S EXTENSIONS

- **BPR:** This pairwise method was introduced by Rendle *et al.* [18] and is the state-of-the-art of personalized ranking for implicit feedback dataset.
- **MC-BPR:** This is a multi-relation algorithm based on BPR. MC-BPR [19] extends the standard BPR sampling model by exploiting the difference in strength among user feedback "channels". It is also designed for implicit feedback datasets.
- **2MC-BPR:** This is an improved version of MC-BPR, from two feedback channels to three feedback channels. This is specially designed for our dataset, which conclude of Retweet, Reply and View, three kinds of interactions as feedback.

2) MULTI-BEHAVIOR ALGORITHM AND OTHER BASIC ALGORITHMS

- **itemKNN:** This is the standard item-based collaborative filtering method. We followed the setting of [36] to adapt it for implicit data.
- **Popularity:** This algorithm is based on the prevalence of items. The more popular products are easier to be recommended, and the unpopular products can not be recommended [37].
- **PMF:** Matrix factorization can factorizes the user-item preference matrix to latent vectors to obtain the user's score on items [38].
- **CMF:** The Collective Matrix Factorization [39] simultaneously factor several matrices, sharing parameters among factors when an entity participates in multiple relations. In our paper, CMF consist of three different relation, which is Retweet, Reply and View.

C. PARAMETERS SETTING

There are six parameters in CMBPR. They are learning rate α, regularization λ, *paraw* and *parav*, *factors*, *maxIter* and *paralabel*. The relation of learning rate α and regularization λ is shown in FIGURE 10(a). We can find that learning rate around 0.01 have the best performance. Meanwhile, the corresponding regularization is around 0.1.

paraw and *parav* control the probability distribution of different feedback level. The equation is described as below:

$$p\,(L) = \frac{\omega_L\,|S_L|}{\sum_{Q\in L}\omega_Q\,|S_Q|}, \tag{7}$$

where *paraw* and *parav* are correspond to ω_Q and ω_L in Equation (7). According to FIGURE 10(b), we set *paraw* to be 0.8 and *parav* to be 0.9. *factors* is the dimension of latent vector of user and item, we define it as 128. *maxIter* is set

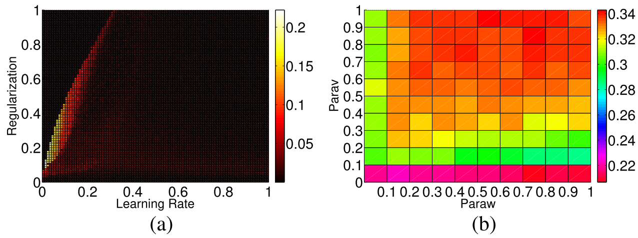

FIGURE 10. Parameter of category-aided Multi-channel BPR. (a) The change of HR affected by learning rate α and regularization λ. (b) The change of HR affected by *paraw* and *parav*.

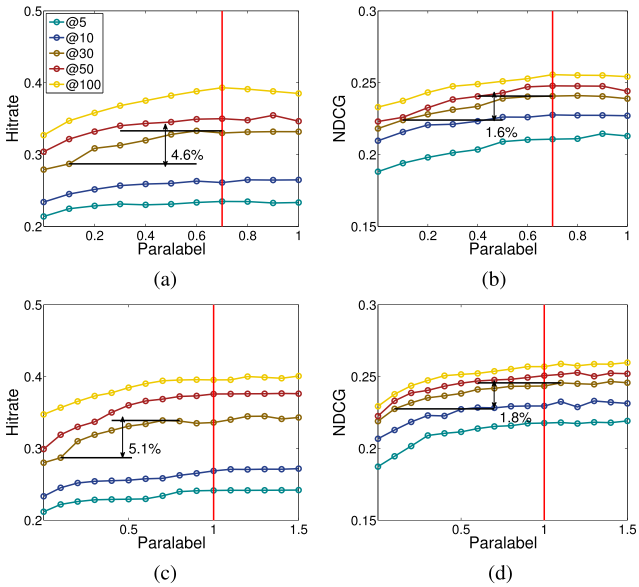

FIGURE 11. HR and NDCG of CMBPR and CMBPR+. (a) HR of CMBPR. (b) NDCG of CMBPR. (c) HR of CMBPR+. (d) NDCG of CMBPR+.

to be 200. Apparently, parameter relate to our algorithm is *paralabel*, which controls the sample ratio of inner-class preference pairwise and outer-class preference pairwise. Therefore, we testified the influence of *paralabel* on HR and NDCG of CMBPR and CMBPR+ in FIGURE 11.

FIGURE 11[a][b] demonstrates the influence of *paralabel* on CMBPR. The x-axis is sampling ratio of inner-class, and the y-axis is HR and NDCG. We find that *paralabel* do have an influence on the algorithm. With the increase of *paralabel*, HR and NDCG are also improving with positive correlation. However, this tendency stops at around 0.7. This indicates that although a high ratio of inner-class sampling will lead to better performance, excessive sampling ratio will make convergence tend to be stabilizing. We can also observe *paralabel*'s influence on CMBPR+ in FIGURE 11[c][d]. *paralabel* in CMBPR is within 1, while *paralabel* in CMBPR+ do not have this restriction. We can find that *paralabel* also affect the result of CMBPR+.

D. ALGORITHM PERFORMANCE

We compare the recommendation performance of BPR's extensions and other methods at first. Then we testify whether CMBPR can exploit the Long Tail effect.

TABLE 3 and TABLE 4 demonstrate the HR and NDCG of all the algorithms. The common feature of CMBPR+ and CMBPR, 2MC-BPR, MC-BPR, BPR is they all based on BPR. The shared character of CMBPR+ and CMBPR, 2MC-BPR, MC-BPR, CMF is they are all multi-behavior algorithms. Except for MC-BPR only consists of two feedbacks, View and Retweet, the other algorithms all consist of three different feedbacks. We can have the following observations:

- BPR is much better than other traditional recommendation algorithm like Popularity, itemKNN and PMF, indicating preference pairwise information can effectively help improve HR.
- Although CMF performs a little better than BPR due to it has multi-behavior information, it is much lower than 2MC-BPR and MC-BPR, which are also featured as multi-behavior algorithm. 2MC-BPR has better performance than MC-BPR, demonstrating the assumption of multi-channel is very effective.
- CMBPR and CMBPR+ substantially outperform the other baselines, for it not only adopts multi-behavior information, but also the video category information.
- HR and NDCG of CMBPR+ are a little better than CMBPR. It is reasonable for it take more detailed user's preferences into consideration.

To analyze whether CMBPR can solve the Long Tail problem, we analyze the HR of each category of CMBPR and 2MC-BPR in FIGURE 12. The distinction between CMBPR and 2MC-BPR is whether they adopt category-aided sampling method. The bars in FIGURE 12 are the HR of CMBPR and 2MC-BPR of each category and the HR value refers to the left y-axis. Meanwhile, the stair step plot depicts the ratio of each category videos and the ratio value refers to the right y-axis. We rearrange the order of categories from small ratio to larger ratio for the convenience of observing. We can find that significantly improved categories are Sport, Movie, Music and Star&Show, which are all niche categories while the popular category videos such as Other and Fun can not significantly improve precision by CMBPR.

This is because CMBPR have knowledge on user's potential interested categories and then acquire more pairwise information of these categories based on biased sampling inner-class method. Under this condition, CMBPR can mine more details of user's interested category even if it is niche category. Besides, the top-N ranking result depends much on how sufficient the target video is sampled as preference pairwise. The more pairwise effective information will lead to better ranking result. Correspondingly, CMBPR can largely enhance sampling effectiveness and achieves the purpose of precision recommendation while BPR can not fully considerate of sampling negative pairwise from niche category. The above reasons lead to the higher precision of CMBPR than of 2MC-BPR over niche categories. However, we also find that Baby&Pet's HR also enhances and it is a relative large category. We infer that it is because the ratio of total number of Baby&Pet are lower than corresponding ratio of View, Reply and Retweet, which can be find in Section II-B. In other

TABLE 3. HR of different algorithms.

HR	CMBPR+	CMBPR	2MC-BPR	MC-BPR	BPR	CMF	PMF	Popularity	itemKNN
@5	**0.243**	0.235	0.214	0.112	0.049	0.052	0.034	0.031	0.035
@10	**0.271**	0.265	0.237	0.138	0.073	0.070	0.056	0.049	0.051
@30	**0.336**	0.333	0.275	0.198	0.122	0.112	0.108	0.096	0.082
@50	**0.385**	0.350	0.301	0.231	0.158	0.141	0.129	0.142	0.098
@100	**0.407**	0.393	0.344	0.280	0.199	0.180	0.173	0.191	0.122

TABLE 4. NDCG of different algorithms.

NDCG	CMBPR+	CMBPR	2MC-BPR	MC-BPR	BPR	CMF	PMF	Popularity	itemKNN
@5	**0.220**	0.215	0.190	0.095	0.0295	0.038	0.024	0.016	0.025
@10	**0.233**	0.228	0.206	0.105	0.049	0.043	0.034	0.021	0.030
@30	**0.247**	0.243	0.220	0.121	0.059	0.053	0.049	0.032	0.038
@50	**0.253**	0.250	0.223	0.130	0.068	0.059	0.051	0.041	0.041
@100	**0.260**	0.258	0.245	0.136	0.072	0.065	0.057	0.049	0.041

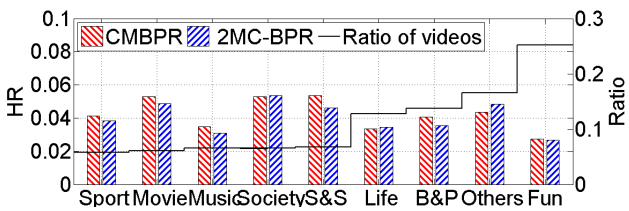

FIGURE 12. HR of CMBPR and 2MC-BPR of each video category.

words, users who favor Baby&Pet videos show strong viscosity on this category, therefore the CMBPR enhances sampling effectiveness for pairwise information of this category.

From the above results we can draw the conclusion that in addition to implicit feedback View, multi-feedback Retweet and Reply can provide multi-level of interest tendencies, in turn, the accuracy can be improved. Based on that, biased sampling inclined to the inner-class category will also increase the HR and NDCG. Besides, the Long Tail effect can be effectively solved. Therefore, CMBPR and CMBPR+ are fit for multi-behavior video recommendation in our scenario.

V. RELATED WORK

Recommender system is a very hot research direction, which has received extensive attention in industry and academia. We will introduce recent recommender system from three aspects: BPR Model' extensions, cross domain recommender system and content-based recommender systems.

A. BPR MODEL'S EXTENSIONS

BPR uses maximum posterior estimator and follows pairwise learning-to-rank framework. It can learn users' preferences in recommender system and have a good performance [18]. Therefore, there are various of extensions based on BPR. We will review the related work in detail in this section.

1) BPR MODEL WITH MULTIPLE SAMPLING CRITERIA

Loni *et al.* [19] propose multi-feedback Bayesian Personalized Ranking to exploit different types of feedback with an extended sampling method, and demonstrate that sampling the negative item with a level-based method can influence the the quality of recommendations. Pan and Chen [20] design a novel algorithm called Group Bayesian Personalized Ranking which introduces richer interactions among users via group preference. Rendle and Freudenthaler [21] find that tailed item distributions lead to noneffective SGD updates, and propose a non-uniform sampling distribution that adapts both to the context and the current state of learning. He *et al.* [22] proposed Listwise Bayesian Personalized Ranking Next Point-of-Interest (POI) recommendation. This method first predicts user's next destination type then filter the POI candidates by the predicated destination type. From the above work we can draw conclusion that sampling distribution of BPR will affect the algorithm performance. It is important to design an efficient sampling method to adapt to the model. However, the above models do not achieve significant performance improvement on personalized recommendation of short video. Lerche and Jannach [23] derive additional pairwise preferences and in the optimization phase bias the optimization procedure to draw a certain amount of samples from these additionally available data points.

2) BPR MODEL WITH NEURAL NETWORK

Ding *et al.* [40] propose the BayDNN model combining Bayesian Personalized Ranking and Deep Neural Networks for friend recommendation. They use one-dimensional convolutional neural network to extract latent deep structural feature representations and then use BPR to captures users preference based on the extracted deep features. He and Mcauley [41] propose a scalable factorization model to incorporate visual signals extracted from pretrained deep CNN into BPR. In summary, with the introduction of DNN, latent deep structural feature representations of social relation or image are extracted, and then can be applied to the Bayesian personalized ranking learning to captures user preference based on the extracted deep features. However, the above Neural Network have limitation on detecting individualized need on short video recommendation.

3) BPR MODEL FOR RELATIONAL DATA

Relational learning is an important part for recommender system. Not only it can improve accuracy of recommender system results by making full use of auxiliary information, but also can predict another relation by collective matrix factorization. As a matrix factorization method, BPR is also applied in relational learning. For instance, Singh and Gordon [24] propose a hierarchical Bayesian model and develop a block Metropolis-Hastings sampler and Hessian of the likelihood to accelerated the convergence. Manotumruksa *et al.* [25] proposed PRFMC model that incorporate geographical and social probabilistic models in the same time for the first time to enhance the performance of the BPR model for venue recommendation. Liu *et al.* [26] also propose BPR with multi-relation data such as social relations and item contents. Krohngrimberghe *et al.* [27] present a multi-relational factorization, which extended BPR to social-relational case and solve cold-start problem correspondingly. Coincidentally, Zhao *et al.* [28] also propose social Bayesian personalized ranking to leverage social connections to improve personalized ranking for collaborative filtering. From the above work we can see, as matrix factorization method, Multi-matrix factorization BPR provide a scalable and effective approach for multi-relational learning tasks. Peska [42], [43] proposes method incorporates content alignments based on multiple similarity matrices into the BPR' s optimization criterion and is able to learn the weight of each similarity matrix. However, despite the auxiliary information, these algorithms are not designed for personalized recommendation of short videos in social networking.

In conclusion, the above literature focuses on relational data, combining users' social relation or items' feature and bias the optimization procedure. There are few special designed BPR algorithms coping with personalized recommendation problem of short videos in social networking which combines detecting user preference and utilize multi-behavior information at the same time.

B. CROSS DOMAIN RECOMMENDER SYSTEM

Online platforms are generally divided into two types, one is information-oriented, such as some e-commerce websites, emphasizing user-item interaction; the other is social-oriented, such as Twitter, which provides social network services with rich user-user links. Although these two domains are heterogeneous, they will share some users, called bridge users, through which we can make cross-domain social recommendation.

Liu *et al.* [26] assume that the user hyperparameters and item hyperparameters are different for each user vector and item vector and sample user hyperparameters and item hyperparameters according to the social relations and item contents to fuse social relations and item contents with ratings. Wang *et al.* [44] present a novel Neural Social Collaborative Ranking approach to seamlessly sews up the user-item interactions in information domains and user-user connections in social networks. Zhao *et al.* [28] develop a model, Social Bayesian Personalized Ranking, based on the simple observation that users tend to assign higher ranks to items that their friends prefer. Wang *et al.* [45] use neighbourhood overlap to approximate tie strength and extend the popular Bayesian Personalized Ranking (BPR) model to incorporate the distinction of strong and weak ties. Cui *et al.* [46] represent videos with users' social attributes, and represent users with content attributes of videos for video recommendation. Deng *et al.* [47] proposes a dynamic user modeling strategy to tackle personalized video recommendation issues in the multimedia sharing platform YouTube, by transferring knowledge from the social textual stream-based platform Twitter. Sun *et al.* [48] proposes a social-aware group recommendation framework that jointly utilizes both social relationships and social behaviors to infer a group's preference for video recommendation.

Despite cross domain recommender system can improve the algorithm accuracy by enriched users' information, it has limitation on items' content information and users' specific operational behavior when watch videos. Therefore, recommender system only considering the social features is not enough.

C. CONTENT-BASED RECOMMENDER SYSTEMS AND HYBRID APPROACHES

Due to the advantage of having no cold start problem, content-based recommendation algorithm attract lots of attetion.

Mei *et al.* [11] propose a novel video-driven recommender called VideoReach, which is able to recommend a list of the most relevant videos according to a user' s current viewing without his/her profile. Mihai Gabriel [12], Mehdi Elahi [13] design a recommender system algorithm which needs to be optimized in order to use a wealth of information encoded in the actual image and audio signals. Users and textual content are embedded into latent feature space. Chen *et al.* [14] apply text embedding function that can be learned end-to-end to predict user interactions with items. Deldjoo *et al.* [49] extract a number representative low-level visual features and then generate personalized content-based recommendations based on this feature. Cui *et al.* [3] propose an algorithm consists of trust friends computing model and video's quality evaluation model to enhance prediction precision. Liu *et al.* [15] propose an improved Item Category aware Conditional Restricted Boltzmann Machine Frame model for recommendation by integrating item category information as the conditional layer, aiming to optimise the model parameters. Choi *et al.* [50] propose a method that constructs genre correlation, and apply our proposed method to the GroupLens movie database.

However, content-based recommendation algorithm usually have limitations of complex content feature extraction procedure and hard to take account of users' watching behavior. Therefore, we propose an method to combine content-based recommendation algorithm and BPR together to conquer the above problems.

VI. CONCLUSION

In this paper, we study the relationship between user preference and video category. Based on the conclusion, we design a state-of-the-art algorithm Category-aided BPR to introduces more detailed relationship of users potentially interested video category. Besides, we extend multi-channel BPR from two different feedbacks to three different implicit feedbacks. In addition, we propose a content-based recommendation algorithm Category-aided Multi-channel Bayesian Personalized Ranking, its sampling method combines Category-aided BPR and Multi-channel BPR. With a real-world short video dataset of one of the largest social network Weibo, we perform a series of experiments. And we demonstrate that CMBPR can recommend short videos with more accuracy compared with BPR and other multi-behavior algorithms, and CMBPR can effectively solve the influence of the "Long Tail" effect.

For future works, we will extend BPR with more social network information, for instance, we will optimize the short video's classification method, besides, we will combine social relation to improve short video recommendation.

REFERENCES

[1] CNNIC. *Statistical Report on the Development of Internet in China*. Accessed: Jan. 18, 2019. [Online]. Available: http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201808/t20180820_70488.htm
[2] R.-C. Chen, H.-Y. Shih, Y.-S. Lin, and H. Hendry, "Video recommendation system based on personalized ontology and social network," in *Proc. ASE Bigdata Socialinformatics*, 2015, Art. no. 6.
[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," *Concurrency Comput. Pract. Exper.*, vol. 29, no. 14, p. e3900, 2017.
[4] C. Bhatt, M. Cooper, and J. Zhao, "Seqsense: Video recommendation using topic sequence mining," in *Proc. Int. Conf. Multimedia Modeling*, 2018, pp. 252–263.
[5] X. Du and H. Yin, "Challenging personalized video recommendation," Tech. Rep., 2017.
[6] J. Guo and C. Gurrin, "Short user-generated videos classification using accompanied audio categories," in *Proc. ACM Int. Workshop Audio Multimedia Methods Large-Scale Video Anal.*, 2012, pp. 15–20.
[7] X. Cheng, C. Dale, and J. Liu, "Statistics and social network of YouTube videos," in *Proc. Int. Workshop Qual. Service*, Jun. 2008, pp. 229–238.
[8] S. Goel, A. Broder, E. Gabrilovich, and B. Pang, "Anatomy of the long tail: Ordinary people with extraordinary tastes," in *Proc. 3rd ACM Int. Conf. Web Search Data Mining*, 2010, pp. 201–210.
[9] F. Guillemin, B. Kauffmann, S. Moteau, and A. Simonian, "Experimental analysis of caching efficiency for YouTube traffic in an ISP network," in *Proc. 25th Int. Teletraffic Congr.*, Sep. 2013, pp. 1–9.
[10] X. Tan, Y. Guo, Y. Chen, and W. Zhu, "Characterizing user popularity preference in a large-scale online video streaming system," in *Proc. Int. Conf. Wireless, Mobile Multi-Media*, Nov. 2016, pp. 246–249.
[11] T. Mei, B. Yang, X.-S. Hua, and S. Li, "Contextual video recommendation by multimodal relevance and user feedback," *ACM Trans. Inf. Syst.*, vol. 29, no. 2, 2011, Art. no. 10.
[12] Y. Deldjoo, M. G. Constantin, H. Eghbal-Zadeh, B. Ionescu, M. Schedl, and P. Cremonesi, "Audio-visual encoding of multimedia content for enhancing movie recommendations," in *Proc. 12th ACM Conf. Recommender Syst.*, 2018, pp. 455–459.
[13] Y. Deldjoo, M. Elahi, M. Quadrana, and P. Cremonesi, "Using visual features based on MPEG-7 and deep learning for movie recommendation," *Int. J. Multimedia Inf. Retr.*, vol. 7, no. 4, pp. 207–219, 2018.
[14] T. Chen, L. Hong, Y. Shi, and Y. Sun. (2017). "Joint text embedding for personalized content-based recommendation." [Online]. Available: https://arxiv.org/abs/1706.01084
[15] X. Liu, Y. Ouyang, W. Rong, and Z. Xiong, "Item category aware conditional restricted Boltzmann machine based recommendation," in *Proc. Int. Conf. Neural Inf. Process.*, 2015, pp. 606–616.
[16] S. K. Bu, H. Kim, J. Lee, and J.-H. Lee, "Improving a recommender system by collective matrix factorization with tag information," in *Proc. Int. Symp. Soft Comput. Intell. Syst.*, Dec. 2015, pp. 980–984.
[17] H. Bang and J.-H. Lee, "Collective matrix factorization using tag embedding for effective recommender system," in *Proc. Joint Int. Conf. Soft Comput. Intell. Syst.*, Aug. 2016, pp. 846–850.
[18] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian personalized ranking from implicit feedback," in *Proc. 25th Conf. Uncertainty Artif. Intell.*, 2009, pp. 452–461.
[19] B. Loni, R. Pagano, M. Larson, and A. Hanjalic, "Bayesian personalized ranking with multi-channel user feedback," in *Proc. ACM Conf. Recommender Syst.*, 2016, pp. 361–364.
[20] W. Pan and L. Chen, "GBPR: Group preference based Bayesian personalized ranking for one-class collaborative filtering," in *Proc. Int. Joint Conf. Artif. Intell.*, 2013, pp. 2691–2697.
[21] S. Rendle and C. Freudenthaler, "Improving pairwise learning for item recommendation from implicit feedback," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2014, pp. 273–282.
[22] J. He, X. Li, and L. Liao, "Category-aware next point-of-interest recommendation via listwise Bayesian personalized ranking," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, 2017, pp. 1837–1843.
[23] L. Lerche and D. Jannach, "Using graded implicit feedback for Bayesian personalized ranking," in *Proc. ACM Conf. Recommender Syst.*, 2014, pp. 353–356.
[24] A. P. Singh and G. J. Gordon, "A Bayesian matrix factorization model for relational data," in *Proc. 26th Conf. Uncertainty Artif. Intell. (UAI)*, Catalina Island, CA, USA, Jul. 2010, pp. 556–563.
[25] J. Manotumruksa, C. Macdonald, and I. Ounis, "A personalised ranking framework with multiple sampling criteria for venue recommendation," in *Proc. ACM Conf. Inf. Knowl. Manage.*, 2017, pp. 1469–1478.
[26] J. Liu, C. Wu, and W. Liu, "Bayesian probabilistic matrix factorization with social relations and item contents for recommendation," *Decision Support Syst.*, vol. 55, no. 3, pp. 838–850, 2013.
[27] A. Krohn-Grimberghe, L. Drumond, C. Freudenthaler, and L. Schmidt-Thieme, "Multi-relational matrix factorization using Bayesian personalized ranking for social network data," in *Proc. 5th ACM Int. Conf. Web Search Data Mining*, 2012, pp. 173–182.
[28] T. Zhao, J. Mcauley, and I. King, "Leveraging social connections to improve personalized ranking for collaborative filtering," in *Proc. ACM Int. Conf. Conf. Inf. Knowl. Manage.*, 2014, pp. 261–270.
[29] G. Guo, H. Qiu, Z. Tan, L. Yuan, M. Jing, and X. Wang, "Resolving data sparsity by multi-type auxiliary implicit feedback for recommender systems," *Knowl.-Based Syst.*, vol. 138, pp. 202–207, Dec. 2017.
[30] Tonghuashun. (May 9, 2018). *Weibo's Financial Report of the 2018*. [Online]. Available: http://stock.10jqka.com.cn/usstock/20180509/c604358605.shtml
[31] R. Keralapura, A. Nucci, Z.-L. Zhang, and L. Gao, "Profiling users in a 3G network using hourglass co-clustering," in *Proc. Int. Conf. Mobile Comput. Netw. (MOBICOM)*, Chicago, IL, USA, Sep. 2010, pp. 341–352.
[32] Y.-J. Park and A. Tuzhilin, "The long tail of recommender systems and how to leverage it," in *Proc. ACM Conf. Recommend Syst.*, 2008, pp. 11–18.
[33] Z. Gantner, L. Drumond, C. Freudenthaler, and L. Schmidt-Thieme, "Personalized ranking for non-uniformly sampled items," *J. Mach. Learn. Res.*, vol. 18, pp. 231–247, Jun. 2013.
[34] S. Kabbur, X. Ning, and G. Karypis, "FISM: factored item similarity models for top-n recommender systems," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 659–667.
[35] M. Yan, J. Sang, and C. Xu, "Mining cross-network association for YouTube video promotion," in *Proc. ACM Int. Conf. Multimedia*, 2014, pp. 557–566.
[36] D. P. Kingma and J. Ba. (2014). "Adam: A method for stochastic optimization." [Online]. Available: https://arxiv.org/abs/1412.6980
[37] X. Mao, Q. Li, H. Xie, and Y. Rao, "Popularity tendency analysis ranking-oriented collaborative filtering from perspective loss function," in *Proc. 19th Int. Conf. Database Syst. Adv. Appl. (DASFAA)*. Springer, 2014.
[38] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2007, pp. 1257–1264.

[39] A. P. Singh and G. J. Gordon, "Relational learning via collective matrix factorization," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2008, pp. 650–658.
[40] D. Ding, M. Zhang, S.-Y. Li, J. Tang, X. Chen, and Z.-H. Zhou, "BayDNN: Friend recommendation with Bayesian personalized ranking deep neural network," in *Proc. ACM Conf. Inf. Knowl. Manage.*, 2017, pp. 1479–1488.
[41] R. He and J. Mcauley, "VBPR: Visual Bayesian personalized ranking from implicit feedback," in *Proc. 13th AAAI Conf. Artif. Intell.*, 2016, pp. 144–150.
[42] L. Peska, "Linking content information with Bayesian personalized ranking via multiple content alignments," in *Proc. ACM Conf. Hypertext Social Media*, 2017, pp. 175–183.
[43] L. Peska, "Hybrid recommendations by content-aligned Bayesian personalized ranking," *New Rev. Hypermedia Multimedia*, vol. 24, no. 2, pp. 88–109, 2018.
[44] X. Wang, X. He, L. Nie, and T.-S. Chua. (2017). "Item silk road: Recommending items from information domains to social users." [Online]. Available: https://arxiv.org/abs/1706.03205
[45] X. Wang, W. Lu, M. Ester, C. Wang, and C. Chen, "Social recommendation with strong and weak ties," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, 2016, pp. 5–14.
[46] P. Cui, Z. Wang, and Z. Su, "What videos are similar with you?: Learning a common attributed representation for video recommendation," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 597–606.
[47] Z. Deng, M. Yan, J. Sang, and C. Xu, "Twitter is faster: Personalized time-aware video recommendation from Twitter to YouTube," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 11, no. 2, 2015, Art. no. 31.
[48] L. Sun, X. Wang, Z. Wang, H. Zhao, and W. Zhu, "Social-aware video recommendation for online social groups," *IEEE Trans. Multimedia*, vol. 19, no. 3, pp. 609–618, Mar. 2017.
[49] Y. Deldjoo, M. Elahi, M. Quadrana, and P. Cremonesi, "Toward building a content-based video recommendation system based on low-level features," in *Proc. Int. Conf. Electron. Commerce Web Technol.*, 2015, pp. 45–56.
[50] S.-M. Choi, S.-K. Ko, and Y.-S. Han, "A movie recommendation algorithm based on genre correlations," *Expert Syst. Appl.*, vol. 39, no. 9, pp. 8079–8085, 2012.

XICHEN WANG received the B.S. degree in optoelectronic information engineering from the Chongqing University, Chongqing, China, in 2011. She is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University, Beijing, China. Her work focuses on social networks and mobile data mining.

CHEN GAO received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2016, where he is currently pursuing the Ph.D. degree with the Department of Electronic Engineering. His work focuses specifically on recommender systems and mobile data mining.

JINGTAO DING received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2015, where he is currently pursuing the Ph.D. degree with the Department of Electronic Engineering. His research interests include mobile computing, mobile data mining, and user behavior modeling.

YONG LI (M'09–SM'16) received the B.S. degree from the Huazhong University of Science and Technology, in 2007, and the Ph.D. degree from Tsinghua University, in 2012, where he is currently a Faculty Member with the Department of Electronic Engineering. From 2012 to 2013, he was a Visiting Research Associate with Telekom Innovation Laboratories and the Hong Kong University of Science and Technology, respectively. From 2013 to 2014, he was a Visiting Scientist with the University of Miami. His research interests include mobile computing and social networks, urban computing and vehicular networks, network science, and the future Internet. He has served as the General Chair, the Technical Program Committee (TPC) Chair, and a TPC Member for several international workshops and conferences. He is currently an Associate Editor of the *Journal of Communications and Networking* and the *EURASIP Journal of Wireless Communications and Networking*.

DEPENG JIN (M'09) received the B.S. and Ph.D. degrees in electronics engineering from Tsinghua University, Beijing, China, in 1995 and 1999, respectively, where he is currently an Associate Professor and the Vice Chair of the Department of Electronic Engineering. He received the National Scientific and Technological Innovation Prize (second class), in 2002. His research interests include telecommunications, high-speed networks, ASIC design, and the future Internet architecture.