IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 4, 2019, accepted March 23, 2019, date of publication March 28, 2019, date of current version April 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2907929

Exploring Popularity Predictability of Online Videos With Fourier Transform

YAN ZHOU[1,2], ZHANPENG WU[1,2], YIPENG ZHOU[3], MIAO HU[1,2], (Member, IEEE), CHUNFENG YANG[4], AND JING QIN[5], (Member, IEEE)

[1]School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China
[2]Guangdong Province Key Laboratory of Big Data Analysis and Processing, Guangzhou 510006, China
[3]Department of Computing, Faculty of Science and Engineering, Macquarie University, Sydney, NSW 2109, Australia
[4]Tencent Inc., Shenzhen 518057, China
[5]School of Nursing, The Hong Kong Polytechnic University, Hong Kong

Corresponding author: Miao Hu (humiao5@mail.sysu.edu.cn)

This work was supported in part by the National Key R&D Program of China under Grant 2018YFB0204100, in part by the Nature Science Foundation of China under Grant 61572538 and Grant 61802452, in part by the Guangdong Special Support Program under Grant 2017TX04X148, in part by the Natural Science Foundation of Guangdong under Grant 2018A030310079, in part by the China Postdoctoral Science Foundation under Grant 2018M631025, in part by the Australian Research Council under Grant DE180100950, and in part by the Hong Kong Innovation and Technology Commission under Grant ITS/319/17.

ABSTRACT The prediction of video popularity is of significant importance to online video service providers in terms of resource provisioning, online advertisement, and video recommendation. Traditional approaches normally utilize videos' historical popularity traces to make such a prediction. However, it is still uncertain whether the future popularity of a video is sure to be associated with its past popularity. In this paper, we explore the problem of video popularity predictability by analyzing videos' view count traces in the frequency domain with Fourier transform. We observe that sharp turns (e.g., peaks and valleys) of view count traces cause the inaccuracy in popularity prediction, which can be seized and quantified by high-frequency components in the frequency domain. Based on the ratio of high-frequency energy, videos can be classified as the fluctuating group, which is hard for prediction, and the smooth group, which is friendly for prediction. The result is further verified via experiments with state-of-the-art predictive algorithms. Inspired by our findings, we propose a strategy to improve prediction performance by removing out-of-date traces before each sharp turn because it is highly possible that the popularity evolution trend has been altered at each sharp turn. To the end, we compare the prediction issue between videos and microblogs. Surprisingly, most microblog traces are smooth. We conjecture that video providers' recommendation and promotion strategies are prone to causing sharp turns in view count traces. In contrast, there is no such initiative counterpart on microblog platforms changing trace evolution of microblogs frequently.

INDEX TERMS Popularity predictions, Fourier transform, video classification.

I. INTRODUCTION

The knowledge of a video's popularity (defined as daily view count of a video) is of significant commercial value for online video providers. With known future popularity, we can adjust video caching, video recommendation, video procurement and video advertisement decisions accordingly [1]. Thereby, predicting popularity has attracted significant attention from both industry and academia recently. Similar to predict a stock's price from its past prices, predicting popularity by learning historical view count traces is the most widely adopted approach, based on which a great number models have been developed [2]–[7].

However, it is uncorroborated that future popularity must be associated with past popularity. Imagine that video providers can easily change the popularity evolution trend of a video by removing it from the recommendation list in the front page. Such events can cause sharp turns (including peaks and valleys) on view count traces, which in fact are very difficult to predict in advance. Without clarifying this impact, it is risky to design a trace based prediction model by merely exploring knowledge from past popularity.

In this work, we explore the predictability of video popularity by analyzing video view count traces with

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

Fourier Transform. In frequency domain, we can detect the change of popularity evolution trends by the ratio of high frequency energy, which is determined by the weights of high frequency components. The intuition is that the components of view count traces with smooth movement trends will concentrate on low frequency. In contrast, the trace with numerous sharp turns is difficult for prediction and has to be recovered with heavy high frequency components. For convenience, videos with low ratio of high frequency energy are grouped as the *smooth group*, while the rest ones are grouped as the *fluctuating group*.

By classifying videos into two groups, we can discriminate videos of which the evolution trends have been significantly altered. We temporarily assume it is difficult to predict videos with many sharp turns, which will be verified with experiments. Inspired by the comparison study of two video groups, we propose to improve prediction performance by removing out-of-date view count traces with each sharp turn. Our conjecture is that it is highly possible that the popularity evolution trend of a video is altered at a sharp turn due to some event unknown to the public. It is necessary to learn the new evolution trend with new traces instead of past traces.

To the end, we conduct a comparison study between video traces and microblog traces, which shows that traces of most microblogs are smooth, and hence it is easy for prediction. We conjecture that video providers' recommendation strategies will significantly alter videos' popularity evolution trends by causing sharp peaks and valleys on view count traces, and thereby impede accurate prediction. In contrast, there is no such initiative counterpart on microblog platforms that can easily change popularity evolution trends.

To verify our study, we collected the daily view count traces from Tencent Video, which is similar to the YouTube system [8]. It is one of the largest online video providers in China with tens of millions of monthly active users, and millions of concurrent users during peak hours. For comparison, we use samples obtained from a public dataset of Sina Weibo (a platform similar to twitter) with 0.3 million messages crawled over one-month period [9].

The contributions of our work are summarized as follows.

- To the best of our knowledge, the idea of applying Fourier transform for video predictability classification has not been investigated in previous works. We explore video predictability with Fourier transform and conduct a statistical study to unveil the characteristics of videos that are difficult for prediction;
- We propose a strategy to improve prediction performance for trace-based algorithm by removing out-of-date data in time. Its effectiveness is verified by extending RPP-ex algorithm;
- We further conduct a comparison study using microblog traces, which shows that video popularity is more fluctuating and hence more difficult for prediction. Furthermore, we explore the possible reason from video recommendation perspective.

The rest content of the paper is organized as follows. We discuss state-of-the-art relevant works in Sec. II. Preliminary knowledge concerning dataset and predictive models is introduced in Sec. III. Following that, we present video classification criteria and the results of statistical study. Sec. V introduces the strategy to refine trace-based prediction algorithms. A comparison study using microblog traces is illustrated in Sec. VI before we conclude our work in Sec. VII.

II. RELATED WORK

In previous studies, many papers focused on the evolution in the popularity of online videos. For example, Figueiredo *et al.* [10] analyzed three video data sets from YouTube: TOP, YouTomb, Random. TOP dataset refers to those videos that appear in top lists. YouTomb dataset refers to those videos that removed from YouTube due to copyright violation. The random dataset includes videos that are randomly sampled from search engines. They found that videos in different datasets have different popularity growth patterns. Zhou *et al.* [11] proposed a methodology to evaluate video recommendation mechanisms, and explained how the recommendation mechanism affects view count evolution. Wu *et al.* [12] used two processes to describe the evolution pattern of online videos and considered several factors (direct recommendation, word-of-mouth recommendation and the intrinsic popularity) that affect evolution processes. Borghol *et al.* [13] found that not all videos will have the same popularity even if their content is the same. The difference in popularity is not only because of the content of the video, but also because of the content-agnostic factors. For videos with the same content, the authors found that there is a ''rich-get-richer'' phenomenon. Cha *et al.* [1] did a data analysis of the popularity distribution, popularity evolution, and content duplication of user-generated content. They found that video popularity follows a power-law distribution, and it is mostly determined at the early stage of video ages. Zhou *et al.* [14] found that cross-platform information such as release date, actors, director has a big impact on video view counts.

Predicting online video popularity can help estimate the traffic load, which can improve existing video delivery models [15]–[20]. There are a lot of research focusing on predictive models that can predict online video popularity. The basic predictive models are to predict future popularity based on historical popularity series. Szabo and Huberman [6] found that there is a strong linear correlation between the logarithmically transformed popularity at early and later times. Based on this, they designed a simple linear prediction algorithm. Pinto *et al.* [7] improved Szabo-Huberman Model. They introduced multiple input values and considered the influence of different evolution patterns of popularity. Shen *et al.* [5] introduced reinforced Poisson process (RPP) that can model and predict popularity of online videos. Gao *et al.* [21] extended the RPP model with a more general rate factor. Gürsun *et al.* [22] divided the videos into two types: rarely-accessed and frequently-accessed videos. Then they designed different forecasting methods based on the two types. Tan *et al.* [23] proposed a novel time series approach to predict video popularity. This model is based on the correlation between early and future popularity. There are also some studies that use time series related techniques to predict the important parameters in Video on Demand (VoD) systems, such as the online population, bandwidth volatility [24], [25]. These time series related techniques are also well suited for predicting video popularity.

At the same time, many predictive models incorporated some factors that affect popularity. Li *et al.* [3] proposed a propagation-based popularity prediction solution for videos shared in OSNs. This solution considers both the intrinsic attractiveness of a video and the influence from the underlying propagation structure. Lerman and Hogg [26] proposed stochastic models to predict popularity based on early user reaction to new content. Wu *et al.* [27] took multiple time-scale dynamics into consideration when predicting popularity. Chen *et al.* [2] proposed a transductive multi-modal learning method (TMALL) to predict the popularity of videos. Yu *et al.* [28] showed that we can use user and content information from Twitter to predict popularity on YouTube. Mazloom *et al.* [29] added a set of parameters in the field of Marketing which influence popularity of brand-related social media posts, not just relying on textual and visual features. Roy *et al.* [4] proposed a transfer learning framework to predict sudden popularity bursts by utilizing knowledge from social streams. Ma *et al.* [30] presented an approach to predict the long-term popularity on YouTube videos by leveraging the knowledge of video lifetime in the early stage. Ding *et al.* [31] proposed a novel video popularity prediction model which uses Hawkes process as its framework. Xu *et al.* [32] proposed a systematic methodology and an associated online algorithm for predicting the popularity of videos promoted by social networks. However, these studies only implemented prediction algorithms and failed to analyze circumstances trace-based prediction approaches is inefficient. Jing *et al.* [33] proposed a low-rank multi-view regression framework for micro-video popularity prediction. Differently, our work analyze the predictability of online videos and explore the effectiveness of traces in popularity prediction.

Some papers used methods in signal processing to analyze time series data. Rafiei and Mendelzon [34] used the Discrete Fourier Transform (DFT) to extract features of time series similarity measures. Chan and Fu [35] used Haar Wavelet instead of Fourier transform for time series indexing, and found that wavelet can outperform DFT through experiments. Popivanov and Miller [36] used wavelet transformations as a dimensionality reduction technique to do similarity search in time series. Inspired by these efforts, we used Fourier transforms to analyze view count traces of online videos, which is never investigated by existing works.

TABLE 1. Category distribution of videos collected from tencent video.

Category	Movie	TV series	MV	News	UGC
#	3155	358	754	3928	760

III. DATASET AND PREDICTIVE BASELINE MODELS

This section introduces preliminary knowledge including predictive problem, dataset description, predictive models and evaluation metric.

A. PREDICTIVE PROBLEM

A predictive model can use the previous t_r days of popularity to predict the popularity at day t_p, t_r is the reference time in the video popularity prediction, and t_p is the target time. The predicted value of popularity at day t_p is denoted by $\hat{N}_i(t_p, t_r)$, the true value is represented by $N_i(t_p)$. So we have the objective of the predictive model as:

$$\min \quad \frac{\left|\hat{N}_i(t_p,t_r)-N_i(t_p)\right|}{N_i(t_p)}. \tag{1}$$

B. DATASET

In our study, the dataset in use is collected from Tencent Video, a very representative online video system in China. In 2017, Tencent video app has over 150 million active daily users and more than 540 million monthly active users. On average, each user contributes 437 minutes of monthly views. More than 20 video categories are provided such as movies, TV series, Music Video(MV), User Generated Content (UGC).

We collected the view count traces of all videos uploaded to Tencent Video from January 1, 2015 to December 31, 2015, belonging to Movie, TV series, MV, News and UGC. For each video, we track its view count trace for over 1 year with the following information: upload time, video id, daily view count and type. The distribution of video types of our dataset is shown in Table 1.

C. PREDICTIVE BASELINE MODELS

We select five typical trace-based predictive models for our study. They are Szabo-Huberman (SH) Model [6], Multivariate Linear (ML) Model [7], MRBF Model [7], RPP-ex Model [21] and ARIMA Model [37]. Note that our framework is general and can include a broader range of predictive models. However, due to limited space, we only use five of them for this study. In addition, predictive models involving side information such as social information are not involved either because of the difficulty to compare them fairly.

1) SZABO-HUBERMAN (SH) MODEL

Through experiments with the real dataset, Szabo and Huberman [6] detected a linear relationship between videos' view counts (in logarithmic form) at early and later stages. Inspired by experimental outcomes, they proposed a predictive model as follows:

$$\hat{N}_i\left(t_p, t_r\right) = \alpha_{t_p,t_r} \cdot N\left(t_r\right). \tag{2}$$

Here, α_{t_p,t_r} is a parameter learned from dataset. It means for any given t_r and t_p, we learn the linear relation α_{t_p,t_r} from sampled videos before we apply the model for prediction.

2) MULTIVARIATE LINEAR (ML) MODEL

Pinto *et al.* [7] extended SH model by expanding a single parameter α_{t_p,t_r} to a parameter vector Θ_{t_p,t_r}. Specifically, the prediction model is expressed as

$$\hat{N}_i\left(t_p, t_r\right) = \Theta^T_{t_p,t_r} \cdot X_i\left(t_r\right), \tag{3}$$

where $\Theta_{t_p,t_r} = \left(\theta_{t_p,1}, \ldots, \theta_{t_p,t_r}\right)^T$ and $X_i\left(t_r\right) = (x_i(1), \cdots, x_i(t_r))^T$. Here, $(x_i(t) = N_i\left(t\right) - N_i\left(t-1\right))$ is the view count at day t. Again, Θ_{t_p,t_r} should be learned from sampled videos with complete view count traces.

3) MRBF MODEL

The MRBF model is a generalization of the ML model [7] by involving patterns for view count traces. First of all, a number of videos will be selected from the training set as centers, denoted by set $\mathcal{C}$. Each sampled video represents a pattern of view count traces. Then, the similarity between each video and each center (i.e., each pattern) is calculated by Radial Basis Functions (RBFs) (e.g., Gaussian RBF). The similarity between any pair of video i and center video j with $j \in \mathcal{C}$ is $RBF_j\left(i\right)$. ω_j is the model weight of RBF feature for a center video j. The predictive model is

$$\hat{N}_i\left(t_p, t_r\right) = \underbrace{\Theta_{t_p,t_r} \cdot X\left(t_r\right)}_{\text{ML}} + \underbrace{\sum_{j\in\mathcal{C}} \omega_j \cdot RBF_j\left(i\right)}_{\text{RBF Features}}. \tag{4}$$

Here, both Θ_{t_p,t_r} and ω_j are parameters learned from dataset.

4) RPP-EX MODEL

RPP-ex model was originally proposed by Gao *et al.* [21] to predict popularity for microblog messages, which can be easily applied to predict video popularity. In this model, user view arrival for video i is assumed to obey a non-homogeneous Poisson process with rate $\lambda_i(t)$ at time t. The rate $\lambda_i(t)$ is expressed as

$$\lambda_i\left(t\right) = q \cdot \left[m + \sum_{j=0}^{N_i(t)} e^{-\alpha_i j}\right] \cdot f_i\left(t\right), \tag{5}$$

where m is a constant parameter, q is the video's inherent attractiveness, $N_i(t)$ is the cumulative view count until time t, $m + \sum_{j=0}^{N_i(t)} e^{-\alpha_i j}$ models a "richer-get-richer" effect, and $f_i(t)$ depicts the decay of user interests for video i with time.

5) ARIMA MODEL

ARIMA is a commonly used model in time series prediction [24], [25], [37], which combines the ideas of autoregressive models and moving average models. The prediction model is expressed as

$$x_i\left(t\right) = \mu(t) \sum_{j=1}^{p} \alpha_j x_i(t-j) + \sum_{k=1}^{q} \theta_k \mu(t-k), \tag{6}$$

where p is the number of autoregressive items, and q is the number of moving average items. $(x_i(t) = N_i\left(t\right) - N_i\left(t-1\right))$ is the view count at day t and μ is the white noise. α and θ are the parameters we want to learn. For an unstable time series, we first differentiate it and convert it into a stationary time series, and d is the number of differences required to make it a smooth time series.

D. THE PROPOSED PREDICTABILITY CLASSIFICATION METHOD

Our main contribution is the predictability classification method presented in Sec.IV. We designed a video classification methods so as to discriminate video predictability with Fourier transform. Videos with smooth trace are put into Group_S, the rest videos with fluctuating traces are put into Group_F. Then we use state-of-the-art prediction algorithms to prove the validity of our classification method. That is to verify that the two groups of video separated by our classification method have significant differences in prediction error. The average predicted error of videos in Group_S is denoted by $\delta_S(t_p, t_r)$, and the average predicted error of Group_F is represented by $\delta_F(t_p, t_r)$. Similarly, t_r is the reference time in the video popularity prediction, and t_p is the target time. Then the objective of our classification method is presented as follows:

$$\max \quad \delta_F(t_p, t_r) - \delta_S(t_p, t_r). \tag{7}$$

E. EVALUATION METRIC

We use the error metrics commonly used in time series prediction in many previous works(e.g., MRSE) [3], [5], [21] to evaluate prediction performance. Because of the great variability in popularity across different videos, we use the relative error MRSE instead of the absolute error. We use MBRE to measure the error in our paper. It adds a limit to the MRSE, that is, each individual error is bounded by 1, which can prevent some extreme values from affecting the overall error. MBRE is defined as:

$$MBRE = \frac{1}{m} \sum_{i=1}^{m} \min\left\{\left|\frac{\hat{N}_i\left(t_p, t_r\right) - N_i\left(t_p\right)}{N_i\left(t_p\right)}\right|, 1\right\}, \tag{8}$$

where m is the number of videos, t_r is the reference time in the video popularity prediction, and t_p is the target time, that is to use the previous t_r days of popularity to predict the popularity at day t_p. $\hat{N}_i\left(t_p, t_r\right)$ is the predicted view count, while $N_i\left(t_p\right)$ is the actual view count at day t_p.

IV. UNDERSTANDING VIDEO PREDICTABILITY VIA FFT

In this section, we first classify videos according to a simple criteria, which is further validated with experiments and case study. Then, a statistical study is conducted with the purpose to extract and exhibit features of videos in the group difficult for prediction.

The intuition to apply the Fourier transform is explained as follows. The Fourier transform can decompose a continuous time domain function into a sum of a series of *sin* and *cos*

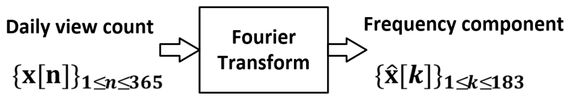


FIGURE 1. Input and output of Fourier transform.

functions, and get the frequency domain representation of the function [38]. A time domain function with more sharp jumps and drops will generate heavier high frequency components. We anticipate that a trace with many sharp and significant fluctuations will be uncertain for prediction. In contrast, a function with mild movement trends will be friendly for prediction. According to the feature of the Fourier transform, it is reasonable to capture sharp changes with high frequency components. The Discrete Fourier Transform (DFT) makes the Fourier transform applicable for discrete series. The time series of daily view counts is the input to Discrete Fourier transform (DFT). For the sequence of N points $\{x[n]\}_{1 \leq n \leq N}$, the discrete Fourier transform follows

$$\hat{x}[k] = \sum_{n=1}^{N} e^{-i\frac{2\pi}{N}nk} x[n] \tag{9}$$

where e is the base of the natural logarithm, i is the imaginary unit and $k = 1, \cdots, N$, represents the serial number of the frequency component. The input of the Fourier transform is the time series of the daily view count $\{x[n]\}_{1 \leq n \leq N}$, and the output is the frequency components in complex domain.

In our experiments, a 365-day daily view count sequence can produce 365 frequency components by conducting Fourier transform. However for a real sequence in time domain, the output of the discrete Fourier transform is Hermitian-symmetric. So in practical applications, for input sequences of length n, we only need to examine the first $\frac{n}{2} + 1$ frequency components after the transformation. That is, for the 365-day daily view count sequence, only the first 183 frequency components in frequency domain are used. The input and output of the Fourier transform in our paper are shown in Fig. 1.

We further use a specific sample to illustrate the input and output of DFT. Fig. 2(a) shows the daily view count for a video randomly selected from video data set. Fig. 2(b) is the magnitude spectrum of frequency components. The x-axis represents the index of the frequency component with total 183 components. The y-axis is the magnitude, i.e. the module of each frequency component. The energy of a frequency component is the square of its magnitude, hence the energy of high frequency components is the sum of energy contributed by components with high frequency.

A. CLASSIFICATION CRITERIA

The video classification criteria is determined by two thresholds: γ_1, the threshold to distinguish low frequency and high frequency, and γ_2, the threshold of high frequency energy ratio to classify videos. All videos with high frequency energy ratio less than γ_2 are put into *Group_S*, which is supposed to include videos with smooth traces. The rest videos will form *Group_F* with most videos with fluctuating traces.

The classification process based on these two thresholds is as follows. The daily view count trace of each video is a series of nonnegative integers with length $L = 365$. Each point in this series is the view count per day since the video was uploaded to Tencent Video, denoted by $x(t)$ ($x(t) = N(t) - N(t-1)$). By applying FFT on this series, we get a series of complex numbers in frequency domain. If the maximum frequency is f_m, frequency higher than $f_m \cdot \gamma_1$ will be regarded as high frequency. For a video, the high frequency energy is the summation of the square of each high frequency component. The energy ratio (denoted by τ) of high frequency energy over total energy can be computed correspondingly. Then, videos with high frequency energy ratio less than γ_2 are put into *Group_S*, the rest videos will be put into *Group_F*.

We select $\gamma_1 = 0.5$ as the threshold to divide high and low frequencies. Then calculate high frequency energy ratio τ for each video. The cumulative distribution function (CDF) curve of τ (i.e., high frequency energy ratio) is plotted in Fig. 3. Then we select $\gamma_2 = 0.4$ as the threshold to partition videos, which is marked by an arrow in Fig. 3. Because we just want to verify the difference in predictability between the two groups of videos, the thresholds are set heuristically. However, according to subsequent validation experimental results, this threshold works well. To facilitate understanding and further illustrate the effectiveness of threshold selection, the selection of the two thresholds and their impact on the classification results are discussed at the end of this section.

B. EXPERIMENTAL COMPARISON STUDY

To validate whether the video classification is effective, we implement aforementioned five video popularity prediction models to compare prediction performance between two video groups. The target is to validate that videos in Group_F is much harder for prediction.

We set the experiment as follows. t_r is varied from 10 to 364, while t_p is always set as $t_r + 1$. In other words, we use view count traces until day t_r to predict view counts of the next day.

Totally, we have 8955 videos. For RPP-ex model and ARIMA model, it is unnecessary to classify videos as training set and test set because they only utilizes a single video's historical view counts to make prediction. For the other three models, we need to partition all videos into five sets randomly for cross validation. Each time, four of them are consolidated as the training set to train parameters in SH model, ML model and MRBF model separately. The remaining one set is used as the testing set to compute MBRE. By rotating the testing set five times, we can compute prediction error for each video. Then, according to the group a video belongs to, we can calculate MBRE for each group.

The experiment results are presented in Fig. 4 with x-axis representing t_p and y-axis representing MBRE. Curves with circle marks are the results of Group_S, while curves

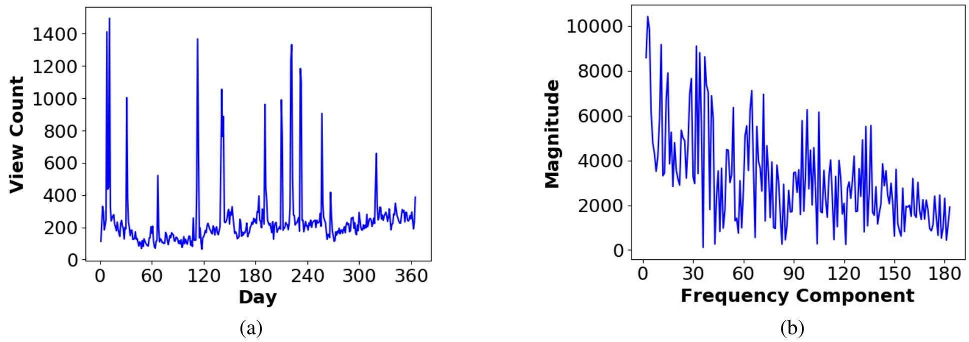


FIGURE 2. An example of Fourier transform on video dataset. (a) Daily view count trace. (b) Magnitude spectrum.

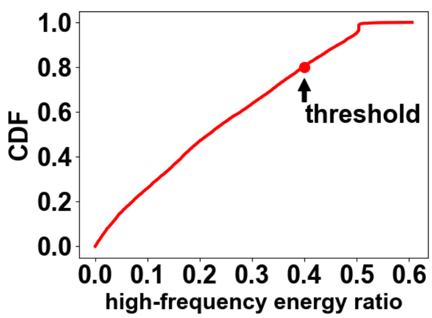


FIGURE 3. CDF of high-frequency energy ratio and the classification point.

with cross marks are the results of Group_F. Obviously, for all predictive models, the prediction performance of videos in Group_S is much better than that of videos in Group_F indicating the effectiveness of our video classification approach. Note that the performance of RPP-ex and ARIMA is very stable because they process each video independently without the necessity to learn parameters from a group of videos.

We can also observe that prediction errors gradually diminish with t_p because with the increase of t_p, cumulative view counts will become larger such that the extent of prediction errors becomes smaller. Nevertheless, this does not mean popularity prediction is a trivial problem by collecting more historical records. More often, we have to make the prediction with very limited historical view counts.

Considering the similarity between wavelet transform and Fourier transform, we also explore the feasibility of wavelet for predictability classification. Similar to the Fourier transform, the wavelet transform has high and low frequency information. The approximation coefficient obtained by the wavelet transform characterizes the low frequency part of the signal, and the detail coefficient characterizes the high frequency part of the signal. With this property, we also did a classification experiment. In our experiments, several commonly used wavelet basis functions were used, such as Haar, Daubechies, etc. In the wavelet classification experiment, we also divided videos into two groups. We implement ML model to compare their classification performance, and the experimental result is shown in Fig. 5. It can be seen that the classification performance of Fourier transform is slightly better. The difference between the prediction errors of the two groups of video separated by the Fourier transform is larger. In addition, the wavelet transform takes time in choosing appropriate wavelet basis function, so we use the Fourier transform in our work.

To facilitate a better understanding of classification results, we select four concrete videos to conduct a case study. Traces of these examples are presented in Figure 6 with x-axis representing day t and y-axis representing view counts.

Fig. 6(a) and Fig. 6(b) are two normal examples. Example 1 with low τ (high frequency energy ratio) is selected from Group_S, and example 2 with high τ is selected from Group_F. Apparently, the trace of example 1 is very stable. Its popularity can be predicted more accurately because its view count movement trends can be well captured by learning historical view counts. In contrast, in example 2, the view count trace keeps fluctuating severely with time resulting in the difficulty to predict its future popularity based on past view counts. The two examples typically show the view count trace patterns we intend to put to each group.

However, our criteria are imperfect, and there exist situations we cannot simply judge predictability based on τ, though such situations are rare. In our dataset, only 37 videos could not be classified correctly. Fig. 6(c) and Fig. 6(d) show two exceptional examples, which are not correctly discriminated by our approach. For example 3, there are two sharp and significant jumps. Consequently, predictive models cannot accurately predict its trace evolution, though τ is low because the trace curve is quite stable for most of the time. Example 4 is another typical exception with high τ caused by many small sharp peaks as we show in the figure. Nevertheless, the trace has a noticeable decreasing trend, and it is not difficult to predict its popularity with existing predictive models.

In summary, for most cases, our approach can classify them correctly. Some exceptional cases call for more sophisticated classification criteria, which will our future work.

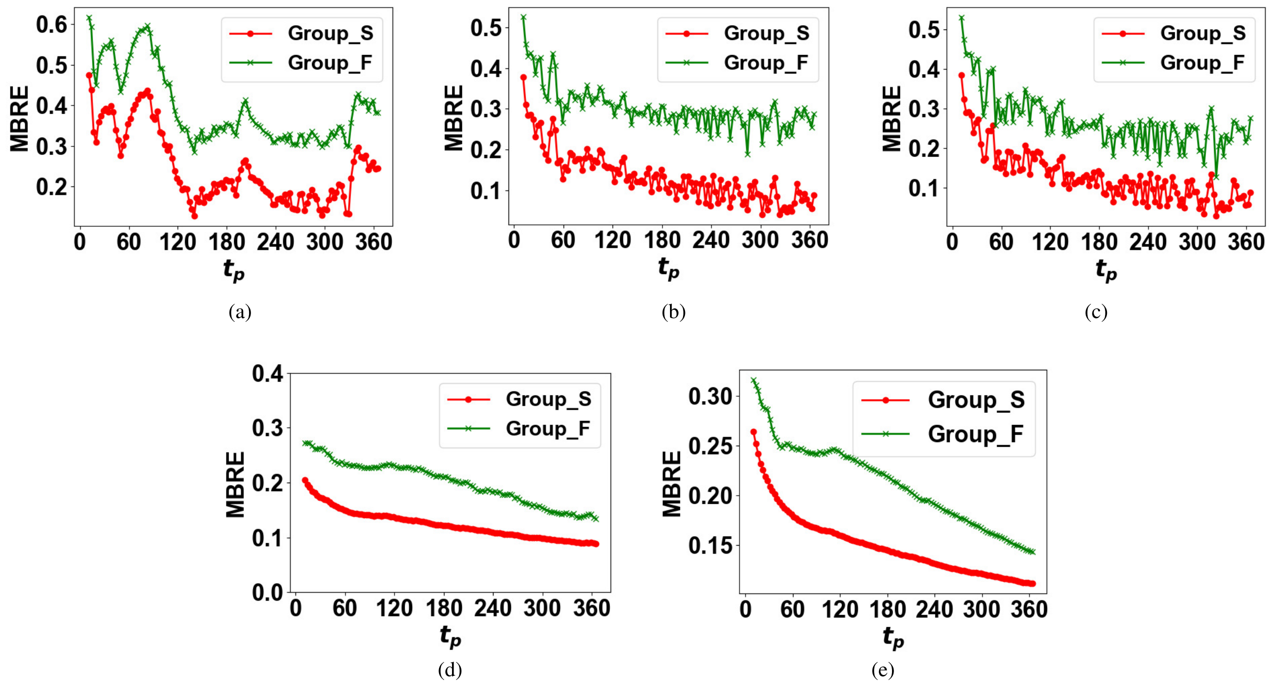


FIGURE 4. **The comparison of prediction performance between Group_S and Group_F for online videos by implementing five prediction models. (a) SH model. (b) ML model. (c) MRBF model. (d) RPP-ex model. (e) ARIMA model.**

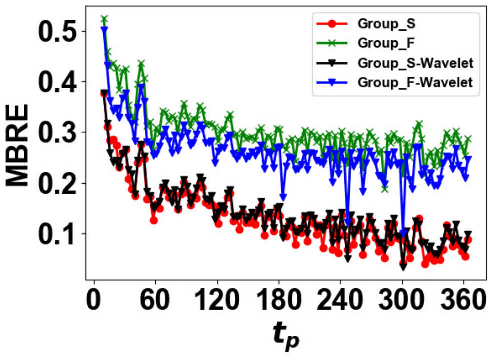


FIGURE 5. **Comparison of wavelet and Fourier transform classification results.**

C. STATISTICAL ANALYSIS

With two groups of videos at hand, we can compare them by conducting a statistical study to scrutinize videos in Group_F.

Briefly speaking, we compare Group_S and Group_F from three perspectives: popularity distribution, category distribution and fluctuation extent. The detailed statistical results are listed as below:

- We compare popularity distributions of two video groups by drawing their popularity CDF curves in Fig. 7(a). X-axis is the logarithm value of the total view count per video over 365 days. Apparently, Group_S contains more popular videos with much larger total view counts. Two jump steps appear on the Group_F curve indicating that both unpopular and popular videos take heavy portions in Group_F. It is easy to understand that view count traces of unpopularity videos will be more fluctuating and harder for prediction because their views may be mainly driven by occasional views. Popularity trends in the future can be very different from existing trends. However, more than 60% videos in Group_F are temporary popular. Their prediction inaccuracy cannot be simply explained by occasional views. To explore features of these videos, we exclude unpopular videos from Group_F if a video's total view count is less than 1000. The threshold to remove unpopular videos is also marked with a label in Fig. 7(a). The subsequent statistical study in this subsection only covers left videos.
- The category distribution is compared in Fig. 7(b). Black (White) histograms represent statistical results of videos in Group_S (Group_F). Movie and News are two categories with the highest percents. It is interesting to note that News category dominates Group_F in comparison to the other three categories.
- At last, we compare the peak height and the number of local peaks between two groups to observe their stability. For a video trace, the peak occurs at day t_m when the view count value reaches the maximum value of the whole year. To fairly compare videos of different popularity, we normalize peak value by the view count of a previous day. In other words, peak height is $\frac{x_i(t_m)}{x_i(t_m-1)}$ for video i. Here, $(x_i(t) = N_i(t) - N_i(t-1))$ is the view count at day t. We exclude special cases with $t_m = 1$ or $x_i(t_m - 1) = 0$ from our statistics. Similarly, for any day t if $\frac{x_i(t)}{x_i(t-1)} \geq 2$, a local peak is counted. We compare

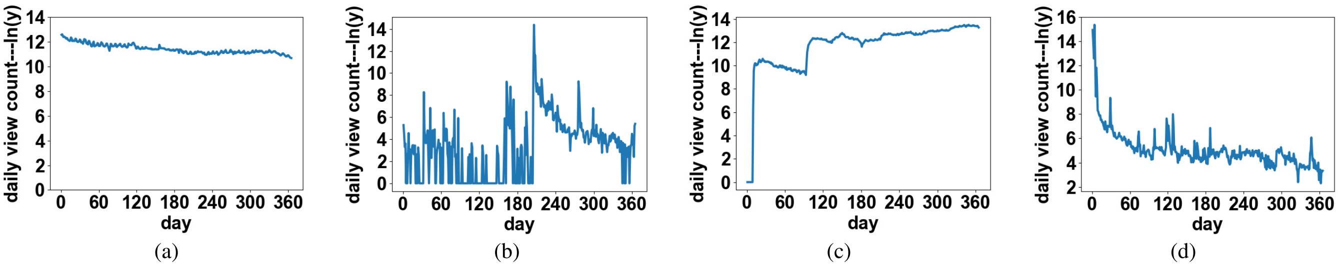


FIGURE 6. A case study of four video traces. (a) Normal example 1 with low τ. (b) Normal example 2 with high τ. (c) Exceptional example 3 with low τ. (d) Exceptional example 4 with high τ.

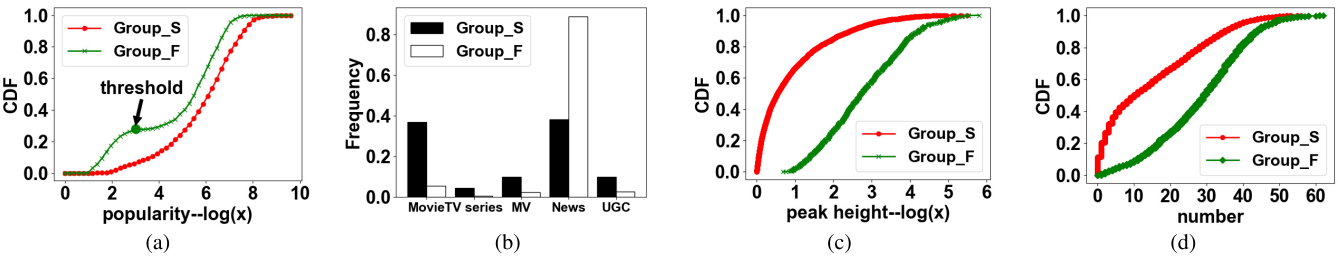


FIGURE 7. A statistical study of video groups by comparing them in popularity distribution, category distribution, peak height and local peak numbers. (a) Popularity distribution of videos. (b) Category distribution of videos. (c) Normalized peak height of videos. (d) The number of local peaks of videos.

peak height and the number of local peaks for two groups in Fig. 7(c) and Fig. 7(d) respectively by drawing their CDF curves. From the comparison results, we can easily draw the following conclusion. The volatility of video traces in Group_F is significantly keener and sharper with wider range than video traces in Group_S.

In light of above observations, we further explore the root of the challenge to predict video popularity as follows. We conjecture that unpopular videos are most probably viewed by users occasionally without any distinguishable movement trend in view count traces, hence it is difficult to develop accurate predictive models for them. Other than unpopular videos, we speculate that traces of popular videos could be distorted significantly by video providers, which may result in sudden jumps or drops on view count traces. For example, a news intensively recommended to users may attract massive views in a very short period. Inversely, a video removed from certain recommendation list may suffer a severe drop in view count. However, such events cannot be predicted merely through historical view counts, and therefore it is difficult to predict popularity based on traces.

D. SELECTION OF CLASSIFICATION THRESHOLD

Actually, the setting of threshold values has significant impact on classification performance, i.e., the group separation gap. We have designed a new set of experiments to illustrate the correlations between the thresholds and classification performance, based on which the optimal combination of threshold values can be determined. We introduce the detailed process of parameter selection in the following.

First, as both γ_1 and γ_2 are less than 1, we set their values from 0.1 to 0.9 in the experiment, increasing by 0.1 each time.

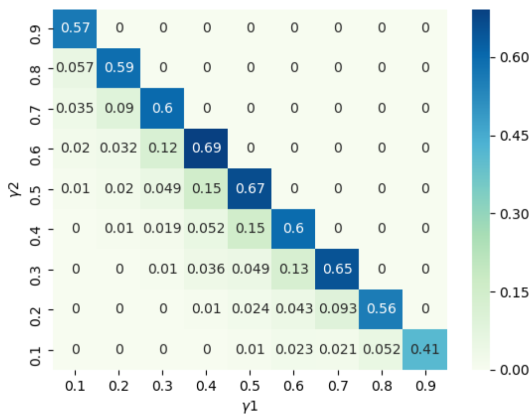


FIGURE 8. Classification results with different threshold combinations.

Then there are 9 choices for γ_1 and γ_2, forming 81 combinations. For each combination of thresholds, we divide Group_S and Group_F accordingly, and then use five prediction algorithms to perform prediction experiments, and calculate the average prediction errors of the two groups. Finally, we subtract the average prediction error of Group_S from the average prediction error of Group_F to obtain a gap value. These operations formulate a correlation map between the threshold values and classification performance.

Fig. 8 shows the heat map of the correlation between threshold combinations and prediction error gap. It could be found that some grids have a value of 0, which is because that the number of videos in a group is 0 after dividing them with these combinations. Namely, a threshold is set too small or too large, and we use 0 to indicate in the heat map. We can also see that the values on the diagonal are particularly large, that is, the classification is particularly good, but we did not select these combinations to show our effects. These combinations work well because the number of videos in Group_F is very small after using these combinations, and the number of Group_S is very large, so it is not representative. However, this can also reflect that these videos with more high frequency energy do have larger prediction errors, which also confirms the effectiveness of our use of Fourier transformation for predictability classification.

In order to better display the classification effect and the rationality of the classification, we have selected two thresholds, $\gamma_1 = 0.5$ and $\gamma_2 = 0.4$ in our experiments. Under these two thresholds, the number of videos in both groups has a certain scale, and a good classification effect is also achieved.

V. EXTENDED PREDICTIVE MODELS OF VIDEO POPULARITY

Inspired by findings we have obtained from the previous section, we propose a strategy to improve trace-based predictive models. In this section, the strategy principle is firstly illustrated before we apply it to existing predictive models. We extend RPP-ex model as the example for illustration. Prediction performance improvement will be evidenced by a new experiment using Group_F.

A. PRINCIPLE OF PREDICTION IMPROVEMENT

In accordance with the findings in the previous section, we come up with a strategy to improve trace-based predictive models for videos belonging to Group_F. The outcomes we have obtained manifest that the prediction of more fluctuating videos will be less accurate. We speculate that a significant view count jump could correspond to an event that can alter popularity evolution trends. For example, the popularity of a movie with mild popularity could surge once it is selected into the list of the top rated movies. In fact, such event can impact video traces at any time and could last a long period, whereas past view counts fail to contain this information. In view of that, we propose to remove out-of-date trace data in time once a significant peak occurs in that such a peak is possibly attributed to an unknown event. The view count data ahead of the peak could be useless and even harmful to predictive models. Future evolution trends should be learned with new traces.

B. REFINED RPP-EX MODEL

The strategy we have proposed is general and applicable to a broad range of predictive models based on trace data though the specific design may depend on modeling details, and we do not have space to cover all of them. In this work, we particularly demonstrate how to refine RPP-ex model using our strategy.

We stick with the same prediction problem by assuming that cumulative view count information before day t_r is known. The task is to predict view count at day t_p.

Originally, the input data is a series $N_i(1), \ldots, N_i(t_r)$ to RPP-ex model to predict popularity at day t_p. By incorporating our strategy, we first search t_h such that $x_i(t_h)$ is the maximum value of the daily view count series $x_i(1), \ldots, x_i(t_r)$. Here, $(x_i(t) = N_i(t) - N_i(t-1))$ is the view count at day t. Then, we remove the input data on or before day t_h by speculating that t_h is the day when the evolution trend has been altered. Subsequently, we recalculate the cumulative view count from 0 and obtain an updated series denoted by $N_i'(t_h+1), \ldots, N_i'(t_r)$ (i.e., $N_i'(t) = N_i(t) - N_i(t_h)$ for $t > t_h$). Then, we use the updated series as the input of the RPP-ex model to predict $N_i'(t_p)$. At last, we need to amend the predicted value by adding the removed view count back (i.e., $N_i(t_p) = N_i'(t_p) + N_i(t_h)$). Note that at least three input samples are required to solve RPP-ex model [21]. We insist on using original RPP-ex model if $t_r - t_h < 3$.

C. EXPERIMENT RESULTS

A thorough experiment is conducted to demonstrate that our simple strategy can indeed improve prediction accuracy for RPP-ex model.

Note that the refined strategy is particularly designed for Group_F with fluctuating traces. In total, we have three experimental curves representing RPP-ex with Group_S, RPP-ex with Group_F and refined RPP-ex with Group_F respectively, to show the gain we can achieve.

The experiment results are presented in Fig. 9. In Fig. 9(a), t_r varies from 10 to 364 and $t_p = t_r + 1$. In Fig. 9(b), $t_r = 10$ and t_p varies from 11 to 60. In Fig. 9(c), we fix $t_p = 60$ and vary t_r from 10 to 59. At last, Fig. 9(d) shows the performance of RPP-ex and our refined strategy on the complete data set with videos in both groups by varying t_p from 11 to 365.

Based on the experiment results, we have the following observations. In Fig. 9(a), Fig. 9(c) and Fig. 9(d), prediction errors decline with time since with more historical view counts we can always achieve better prediction performance. In Fig. 9(b), MBRE gradually increase with t_p because of the enlarging of $t_p - t_r$. In the first three experiments, the refined RPP-ex model has achieved much better prediction performance on Group_F in comparison with the original RPP-ex model. The refined performance is already very close to and occasionally even better than the performance of Group_S. In the last experiment with the the complete dataset, our strategy also achieves better performance.

In summary, our strategy can be extensively applied to enhance trace-based prediction algorithms. In a nutshell, our principle is to remove out-of-date dataset in time once an event that can significantly change view counts' future movement trends occurs, Otherwise, past view counts may hamper predictive models from digesting the latest changes in time, and hence result in inaccurate prediction results.

VI. A COMPARISON STUDY WITH MICROBLOG MESSAGES

At last, we explore to apply our approach to analyze traces generated by other similar web-based applications such as twitter. We use traces recording the number of times a message is reposted on microblog platforms to conduct this study and compare the result with that of online videos.

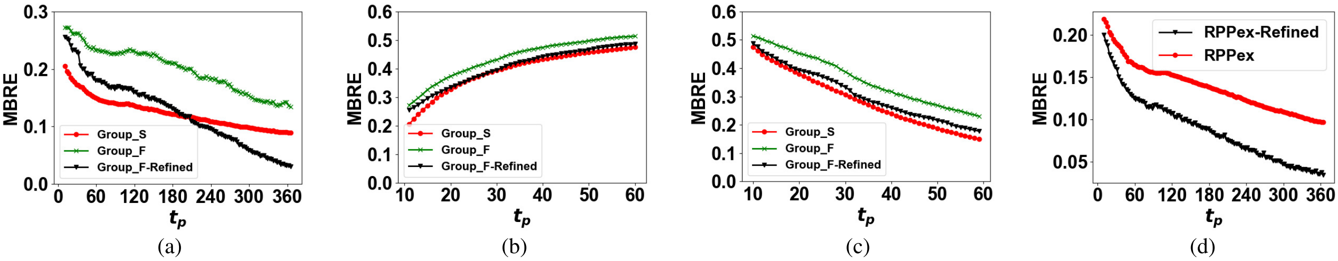


FIGURE 9. Comparing prediction performance between original RPP-ex model and refined RPP-ex model. (a) $t_p = t_r + 1$ and t_r varies from 10 to 364. (b) $t_r = 10$ and t_p varies from 11 to 60. (c) $t_p = 60$ and t_r varies from 10 to 59. (d) Performance on the entire original dataset.

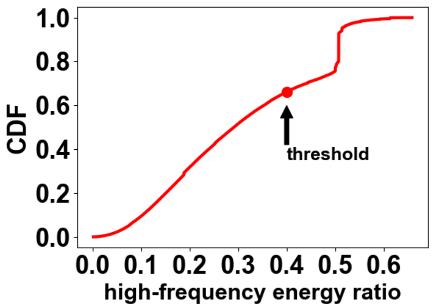


FIGURE 10. CDF of high-frequency energy ratio of microblog messages.

A. MICROBLOG DATASET

The dataset we use for this study is a public dataset that can be retrieved from [9]. The dataset is crawled from Sina Weibo, which analogous to twitter is one of the most reputed and popular microblog platforms in China [39]. The public dataset contains more than 1.7 million users (a selected subset of all users), and their connection information. This dataset has 0.3 million messages, which are posted and reposted by users from 30 Aug. 2012 to 27 Nov. 2012. For each message, the crawled data include its first post time, id, the number of reposts, and the time of each repost.

Note that the lifespan of a microblog message is much shorter than an online video. Typically, a video can keep active for one year or longer, whereas the spreading of a microblog message may cease a few days later. It is meaningless to predict repost times several days later. Due to this difference, we sample repost times every ten minutes and only consider the repost traces in the first day for microblog messages. Accordingly, we have 144 sampled values per trace because there are 1440 minutes per day.

B. CLASSIFICATION CRITERIA

With the same target, we intend to classify all microblog messages into two groups according to the high frequency energy ratio. First of all, we set $\gamma_1 = 0.5$ and use FFT to obtain the high frequency energy ratio of all messages with the same method. The CDF curve is plotted in Fig. 10.

Based on the CDF curve, we heuristically set $\gamma_2 = 0.4$ (of high frequency energy ratio) as the thresholds to classify all microblog messages into two groups. Finally, with a little bit abusing of notations, after removing the microblog that has never been reposted, there are 170212 microblog messages in Group_S and 86991 microblog messages in Group_F.

C. EXPERIMENT RESULTS

Again, we set the prediction problem as follows. With the repost traces of the first t_r time intervals, the objective is to predict the cumulative repost times at the end of time interval $t_p = t_r + 1$ with t_r varying from 5 to 143. The duration of each time interval is 10 minutes.

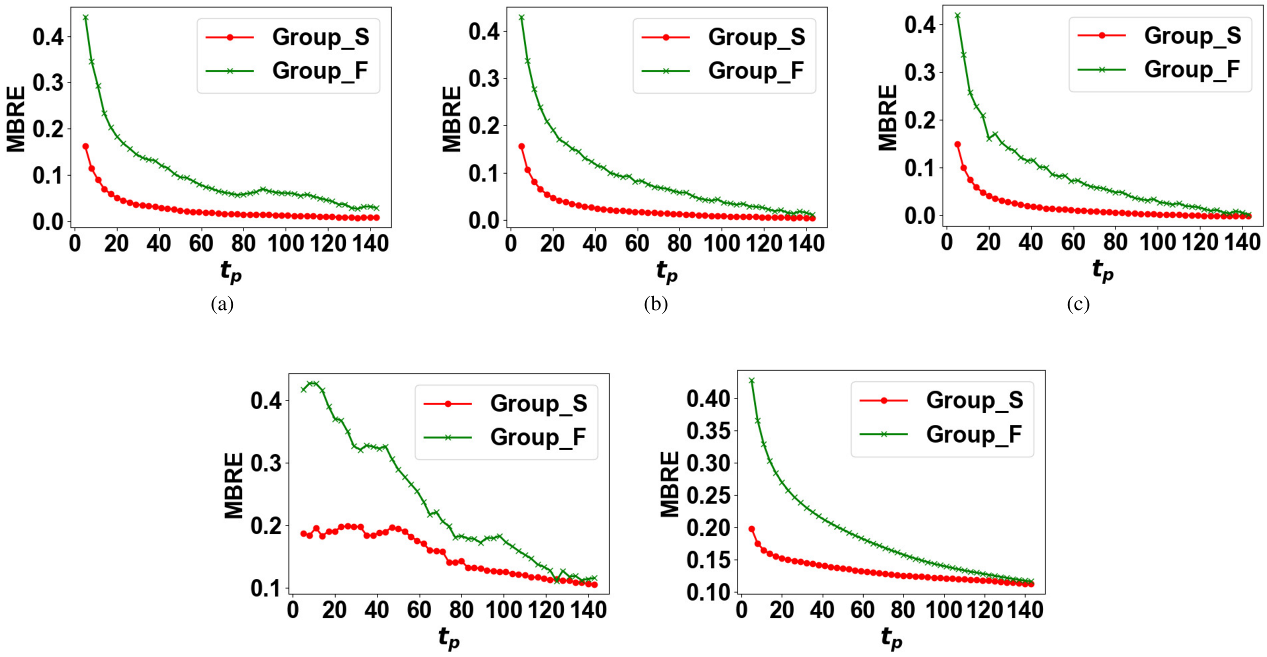
The five predictive models are executed with the microblog dataset. Similar to the experiments with video view count traces, we cut microblog messages into five sets for cross validation. MBRE is still adopted as the metric to evaluate prediction accuracy. Experiment results are shown in Fig. 11 with x-axis representing t_p and y-axis representing MBRE.

Interestingly, the results in Fig. 11 seem to have a similar pattern with the results in Fig. 4. Messages out of Group_S are friendly for prediction with more accurate prediction results in comparison with messages out of Group_F in all five experiments. This comparison study indicates the value of our approach that can be generally applied to discriminate traces difficult for prediction.

D. DISCUSSION

The difference of trace evolution between online videos and microblog messages can be unveiled by further study. In Fig. 12, we plot the CDF curves for the popularity distributions of microblog messages out of Group_S and Group_F separately. Different from the result of online videos, all microblog messages in Group_F are unpopular messages with very few repost times. This is a very nice property because the impact of unpopular messages is usually much less essential than popular messages. In contrast, only about 30% videos in Group_F can be asserted to be unpopular as shown in Fig. 7(a). This difference also implies that the prediction difficulty of microblog messages is mainly attributed to the occasionally repost behavior and low repost times.

In addition, the comparison study reveals the complication to predict online video popularity. We try to explore the reason from video recommendation perspective. For modern online video systems, video recommendation is indispensable and seminal to aid users in locating videos of their interests. Otherwise, users will be overwhelmed by the explosively growing video population. However, video recommendation could considerably change view count traces'



(a)

(b)

(c)

(d)

(e)

FIGURE 11. The comparison of prediction performance between Group_S and Group_F for microblog messages by implementing five prediction models. (a) SH model. (b) ML model. (c) MRBF model. (d) RPP-ex model. (e) ARIMA model.

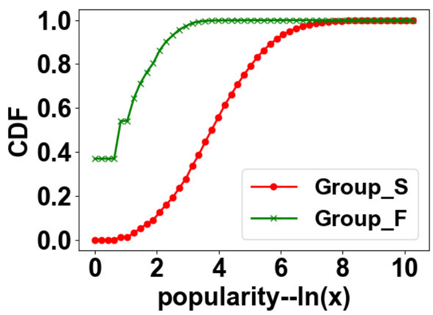


FIGURE 12. Popularity distribution of microblogs.

evolution trends resulting in sharp turns and fluctuations. Unfortunately, a trace-based predictive model relies on historical view counts to infer future popularity by learning the underlying popularity evolution trends. It is almost impossible to incorporate unexpected factors (e.g., recommendation) into consideration by merely considering historical traces. Therefore, we conjecture that video providers' recommendation activities are one of the major reasons leading to inaccurate prediction. Contrarily, there is no such counterpart on microblog platforms who can make global recommendation for messages. Some super users may influence more other users, however it is still negligible in comparison with the large scale of an entire microblog system. The trace of a microblog message is likely a summation of repost behavior contributed by a vast number of users in a distributed manner, and thus microblog messages' traces are more stable and smooth.

Above arguments also give us insights on how to improve prediction accuracy. This problem can be discussed from two perspectives: 1) Online video providers with the full knowledge of their recommendation strategies should quantify the impact of their recommendation on view count traces and develop a model to incorporate this factor into prediction algorithms; 2) Third parties without recommendation information can infer recommendation operations from sharp fluctuations of view count traces (e.g., a sudden jump) so as to develop models that can adapt with such fluctuations.

In the end, we need to emphasize that recommendation activities do not necessarily lead to trace fluctuations. According to our observations, prediction of these most popular videos is rather accurate. These videos usually have been fully recommended to users prior to their births. In most cases, their traces gradually drop with time until they lose user interests eventually. The main trouble is caused by recommendation strategies that try to surprise users because their purpose is help users discover videos that are not well recommended yet [40]–[42].

At the same time, our method can help recommendation systems to catch videos that are going to be popular, which means that these videos with rising popularity can be exposed to users in time. In other words, by predicting future popularity of each video accurately, one can decide which video will attract more views in the future, such that recommendation strategies can be adjusted accordingly [1], [7], [31].

VII. CONCLUSION

Video popularity prediction is a challenging research topic. To our best knowledge, the idea of analyzing view count traces from the frequency domain has not been explored in previous works. By using the Fourier transform, We successfully classify all videos into two groups according to the extent of their trace fluctuations. A statistical study further reveals the characteristics of those videos that are difficult for prediction. Accordingly, a prediction refinement strategy is proposed by removing out of date trace records in time so as to adapt with the latest situations. RPP-ex model is refined as an example to show the effectiveness of our strategy. At last, a comparison study using repost traces of microblog messages is conducted. Surprisingly, only extremely unpopular microblog messages are difficult for prediction. Through the comparison, we discuss trace fluctuations possibly caused by recommendation strategies deployed by online video providers to surprise users.

REFERENCES

[1] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "Analyzing the video popularity characteristics of large-scale user generated content systems," *IEEE/ACM Trans. Netw.*, vol. 17, no. 5, pp. 1357–1370, Oct. 2009.

[2] J. Chen, X. Song, L. Nie, X. Wang, H. Zhang, and T.-S. Chua, "Micro tells macro: Predicting the popularity of micro-videos via a transductive model," in *Proc. 24th ACM Int. Conf. Multimedia*, Oct. 2016, pp. 898–907.

[3] H. Li, X. Ma, F. Wang, J. Liu, and K. Xu, "On popularity prediction of videos shared in online social networks," in *Proc. 22nd ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2013, pp. 169–178.

[4] S. D. Roy, T. Mei, W. Zeng, and S. Li, "Towards cross-domain learning for social video popularity prediction," *IEEE Trans. Multimedia*, vol. 15, no. 6, pp. 1255–1267, Oct. 2013.

[5] H. Shen, D. Wang, C. Song, and A.-L. Barabási, "Modeling and predicting popularity dynamics via reinforced poisson processes," in *Proc. 38th AAAI Conf. Artif. Intell.*, vol. 14, Jul. 2014, pp. 291–297.

[6] G. Szabo and B. A. Huberman, "Predicting the popularity of online content," *Commun. ACM*, vol. 53, no. 8, pp. 80–88, 2010.

[7] H. Pinto, J. M. Almeida, and M. A. Gonçalves, "Using early view patterns to predict the popularity of youtube videos," in *Proc. 6th ACM Int. Conf. Web Search Data Mining*, Feb. 2013, pp. 365–374.

[8] (2018). *Tencent Video*. [Online]. Available: https://v.qq.com/

[9] (2018). *Microblog Dataset*. [Online]. Available: https://www.aminer.cn/data-sna

[10] F. Figueiredo, F. Benevenuto, and J. M. Almeida, "The tube over time: Characterizing popularity growth of youtube videos," in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, Feb. 2011, pp. 745–754.

[11] Y. Zhou, J. Wu, T. H. Chan, S.-W. Ho, D.-M. Chiu, and D. Wu, "Interpreting video recommendation mechanisms by mining view count traces," *IEEE Trans. Multimedia*, vol. 20, no. 8, pp. 2153–2165, Aug. 2018.

[12] J. Wu, Y. Zhou, D. M. Chiu, and Z. Zhu, "Modeling dynamics of online video popularity," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1882–1895, Sep. 2016.

[13] Y. Borghol, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti, "The untold story of the clones: Content-agnostic factors that impact youtube video popularity," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2012, pp. 1186–1194.

[14] Y. Zhou, X. Gu, D. Wu, M. Chen, T. H. Chan, and S. W. Ho, "Statistical study of view preferences for online videos with cross-platform information," *IEEE Trans. Multimedia*, vol. 20, no. 6, pp. 1512–1524, Jun. 2018.

[15] W. Huang, Y. Zhou, X. Xie, D. Wu, M. Chen, and E. Ngai, "Buffer state is enough: Simplifying the design of QoE-aware HTTP adaptive video streaming," *IEEE Trans. Broadcast.*, vol. 64, no. 2, pp. 590–601, Jun. 2018.

[16] Y. Zheng, D. Wu, Y. Ke, C. Yang, M. Chen, and G. Zhang, "Online cloud transcoding and distribution for Crowdsourced live game video streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 8, pp. 1777–1789, Aug. 2017.

[17] J. He, D. Wu, X. Xie, M. Chen, Y. Li, and G. Zhang, "Efficient upstream bandwidth multiplexing for cloud video recording services," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 10, pp. 1893–1906, Oct. 2016.

[18] H. Tian, D. Wu, J. He, Y. Xu, and M. Chen, "On achieving cost-effective adaptive cloud gaming in geo-distributed data centers," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 12, pp. 2064–2077, Dec. 2015.

[19] Z. Xue, D. Wu, J. He, X. Hei, and Y. Liu, "Playing high-end video games in the cloud: A measurement study," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 12, pp. 2013–2025, Dec. 2015.

[20] D. Wu, J. Huang, J. He, M. Chen, and G. Zhang, "Toward cost-effective mobile video streaming via smart cache with adaptive thresholding," *IEEE Trans. Broadcast.*, vol. 61, no. 4, pp. 639–650, Dec. 2015.

[21] S. Gao, J. Ma, and Z. Chen, "Modeling and predicting retweeting dynamics on microblogging platforms," in *Proc. 8th ACM Int. Conf. Web Search Data Mining*, Feb. 2015, pp. 107–116.

[22] G. Gürsun, M. Crovella, and I. Matta, "Describing and forecasting video access patterns," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 16–20.

[23] Z. Tan, Y. Wang, Y. Zhang, and J. Zhou, "A novel time series approach for predicting the long-term popularity of online videos," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 436–445, Jun. 2016.

[24] D. Niu, Z. Liu, B. Li, and S. Zhao, "Demand forecast and performance prediction in peer-assisted on-demand streaming systems," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 421–425.

[25] D. Niu, B. Li, and S. Zhao, "Understanding demand volatility in large vod systems," in *Proc. 21st Int. Workshop Netw. Operating Syst. Support Digit. Audio Video*, Jun. 2011, pp. 39–44.

[26] K. Lerman and T. Hogg, "Using a model of social dynamics to predict popularity of news," in *Proc. 19th Int. Conf. World Wide Web*, Apr. 2010, pp. 621–630.

[27] B. Wu, W.-H. Cheng, Y. Zhang, and T. Mei, "Time matters: Multi-scale temporalization of social media popularity," in *Proc. 24th ACM Int. Conf. Multimedia*,Oct. 2016, pp. 1336–1344.

[28] H. Yu, L. Xie, and S. Sanner, "Twitter-driven youtube views: Beyond individual influencers," in *Proc. 22nd ACM Int. Conf. Multimedia*, Nov. 2014, pp. 869–872.

[29] M. Mazloom, R. Rietveld, S. Rudinac, M. Worring, and W. van Dolen, "Multimodal popularity prediction of brand-related social media posts," in *Proc. 24th ACM Int. Conf. Multimedia*, Oct. 2016, pp. 197–201.

[30] C. Ma, Z. Yan, and C. W. Chen, "LARM: A lifetime aware regression model for predicting youtube video popularity," in *Proc. ACM Conf. Inf. Knowl. Manage.*, Nov. 2017, pp. 467–476.

[31] W. Ding, Y. Shang, L. Guo, X. Hu, R. Yan, and T. He, "Video popularity prediction by sentiment propagation via implicit network," in *Proc. 24th ACM Int. on Conf. Inf. Knowl. Manage.*, Oct. 2015, pp. 1621–1630.

[32] J. Xu, M. van der Schaar, J. Liu, and H. Li, "Timely video popularity forecasting based on social networks," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr./May 2015, pp. 2308–2316.

[33] P. Jing, Y. Su, L. Nie, X. Bai, J. Liu, and M. Wang, "Low-rank multi-view embedding learning for micro-video popularity prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 8, pp. 1519–1532, Aug. 2018.

[34] D. Rafiei and A. O. Mendelzon, "Querying time series data based on similarity," *IEEE Trans. Knowl. Data Eng.*, vol. 12, no. 5, pp. 675–693, Sep./Oct. 2000.

[35] K.-P. Chan and A. W.-C. Fu, "Efficient time series matching by wavelets," in *Proc. 15th Int. Conf. Data Eng.*, Mar. 1999, pp. 126–133.

[36] I. Popivanov and R. J. Miller, "Similarity search over time-series data using wavelets," in *Proc. 18th Int. Conf. Data Eng.*, Feb./Mar. 2002, pp. 212–221.

[37] C. Liu, S. C. H. Hoi, P. Zhao, and J. Sun, "Online ARIMA algorithms for time series prediction," in *Proc. 30th AAAI Conf. Artif. Intell.*, Feb. 2016, pp. 1867–1873.

[38] R. N. Bracewell, *The Fourier Transform and Its Applications*. New York, NY, USA: McGraw-Hill, 1986.

[39] (2018). *Sina Weibo*. [Online]. Available: https://weibo.com/

[40] Y. C. Zhang, D. Ó. Séaghdha, D. Quercia, and T. Jambor, "Auralist: Introducing serendipity into music recommendation," in *Proc. 5th ACM Int. Conf. Web Search Data Mining*, Feb. 2012, pp. 13–22.

[41] M. de Gemmis, P. Lops, G. Semeraro, and C. Musto, "An investigation on the serendipity problem in recommender systems," *Inf. Process. Manage.*, vol. 51, no. 5, pp. 695–717, Sep. 2015.

[42] N. Hurley and M. Zhang, "Novelty and diversity in top-n recommendation—Analysis and evaluation," *ACM Trans. Internet Technol.*, vol. 10, no. 4, Mar. 2011, Art. no. 14.

Authors' photographs and biographies not available at the time of publication.