# A First View on Mobile Video Popularity as Time Series

Aining Wang
anwang13@fudan.edu.cn

Chen Zhang
zhang_chen12@fudan.edu.cn

Yuedong Xu *
ydxu@fudan.edu.cn

Department of Electronic Engineering,
Fudan University, Shanghai, China, 200433

## ABSTRACT

With the rise of mobile video streaming service, video popularity has drawn a great deal of attention in both academia and streaming industry. In this paper, we present a novel view of video popularity as time series. We collect more than two billion view records from one of the largest mobile streaming content providers in China. Using Discrete Fourier transform, we decompose popularity series of videos into two parts, the seasonal and residual component series. We find that the traffic dynamics of videos is governed by a small number of components in the frequency spectrum, and that traffic patterns within and across days differ from type to type. We further investigate the short-term and long-term stability of video popularity, and find that the stability metrics are highly correlated with total video requests.

## CCS Concepts

**•Mathematics of computing → Time series analysis; Computation of transforms; •Networks → Social media networks;**

## Keywords

Video Popularity Dynamics, Discrete Time Series, Time Series Decomposition, User Behavior, Popularity Stability

## 1. INTRODUCTION

The popular online VoD (Video-on-Demand) service attracts a lot of users. The major video service providers offer fancy mobile apps, providing users with easy access to online videos. With a simple touch on the screen, video users leave footprints on servers. These footprints include not only timestamps and IP addresses, but also users' viewing behavior. The huge volume of data brings many benefits to video service providers, ISPs, as well as advertisers, and largely enriches research on video popularity and user behavior, which is helpful in designing, configuring and managing video content distribution systems.



*HotPOST 16, July 04-08, 2016, Paderborn, Germany*



DOI: http://dx.doi.org/10.1145/2944789.2944872

Dated back to 2006, the authors of [1] modeled user behavior pattern based on a large-scale VoD systems deployed by China Telecom. They also made an early attempt in understanding video popularity. Authors of [2] thoroughly studied the popularity life-cycle of videos, the intrinsic statistical properties of requests and their relationship with video age, and the level of content aliasing or of illegal content in the system. Both [1] and [2] only took the number of user request into account. However, they provided fundamental perspectives in understanding VoD systems and user behavior. Other video-related metrics, such as #comments, #favorites, #ratings and average rating, were first taken into consideration in [3]. The authors of [3] also studied users' behavior pattern of higher-level interaction, i.e., leaving a comment, adding a video to one's favorite set, rating the video, and sharing a video to OSN (Online Social Network). Social network data were latter taken into analysis. In [4], researchers merged data from YouTube and Twitter to investigate the relation between social network and video popularity. They obtained an effective cross-network predictor of two kinds of Twitter-driven YouTube views, namely, JUMP (sudden increase) and EARLY (increase shortly after upload). Authors of [5] pointed out that popularity dynamics of OSN-driven views is nonlinear, therefore linear models cannot predict video popularity well. They applied an epidemic model based on the dataset of Renren, the largest OSN in China at that time, and achieved surprisingly accurate prediction results.

In most of the previous works, researchers only focused on the popularity variation between days. Various regression models and learning methods were applied, and amazing outcomes were archived. However, these methods had problems depicting the details in the popularity evolution and user behavior patterns. In [6], the authors present a qualitative description on the access patterns and the daily and hourly change in user interest. This is a step towards VoD service user behavior study on an hourly basis.

As we look into our dataset collected from one of the largest online VoD service providers in China, we find that the variation of video popularity follows certain short-term and long-term patterns. It seems to us that video popularity dynamics resembles signals in the signal processing perspective. We regard the data as more than merely numbers, but also the amplitude of the popularity fluctuation. Therefore, we tried signal processing methods on our dataset and look deeply into the fluctuation of the popularity curves.

To the best of our knowledge, our work makes the first attempt towards understanding the lifecycle of video popularity as discrete time series and apply signal processing methods on it. In this paper, we seek to look into the evolution of video popularity dynamics and underlying user behavior patterns.

The main contributions of this work is as follows:

- We decompose video popularity data into **Seasonal Component Series** and **Residual Component Series** using frequency domain methods. By doing so, we manage to capture the major activities and the trends of video popularity dynamics within and across days.

- We investigate the popularity curve in frequency domain. The frequency spectrum reveals details about popularity evolution. Concretely, we observe periods of 8, 6, and 4.8 hours in the fluctuation of video popularity, which is non-intuitive and can hardly be obtained via traditional methods.

- We compared the frequency spectra of different types of videos and show the various user behavior patterns within and across days.

- We redefine two metrics originally used to evaluate series of time interval [7] and apply them to video popularity series. The redefined metrics evaluate stability in the traffic of individual video within and across days respectively. It is found that the two stability metrics are strong correlated with the total request of videos. This discovery may lead to a new understanding towards the concept of popularity.

The popularity dynamics of videos with different types or viewcounts evolve in very different ways. These differences have large impact on the efficiency of content distribution. We believe that our findings will benefit the design of content replication systems. Concretely, we propose a guideline for replication priority in the means of video type and viewcount.

The rest of the paper is organized as follows. In Section 2, we demonstrate our dataset. Section 3 introduces Digital Signal Processing methods used in our work. In Section 4, we qualitatively analyze the popularity dynamics in time and frequency domain, whereas in Section 5, we analyze the long-term and short-term stability using the redefined metrics. We conclude in Section 6.

# 2. DATASET

## 2.1 Overview

We collect data from one of the largest VoD service providers in China from September 1 to 14, 2014. The dataset involves around 1.5 million videos and more than 2.5 billion requests started from mobile clients. The features of our data include:

$<$*Timestamp, VideoID, VideoType, UserID, Duration*$>$

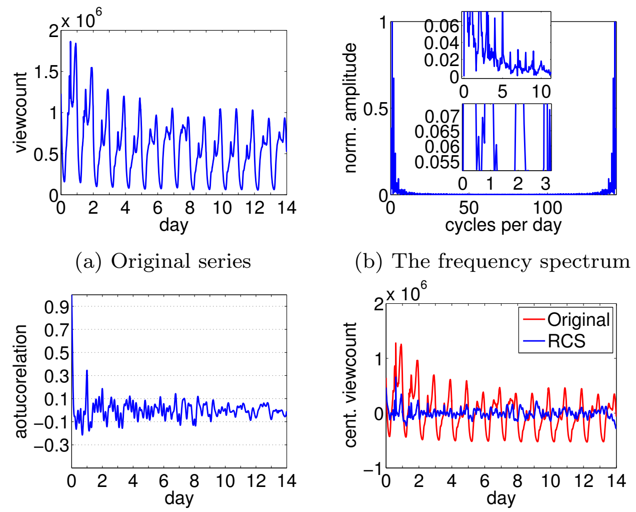
*Timestamp* marks the end of a session. By subtracting the duration of the session, we get the time when the request arrived. *VideoType* identifies the type of video, i.e., TV, music video, news, entertainment, etc. *Duration* stands for the viewing time of the video session.

## 2.2 Data Pre-Processing

In this work, videos are not treated as independent clips. We are interested in the similarities and differences between groups of videos. By summing over all video in the same category, the trend of popularity evolution is easy to obtain. To focus our study on the mainstream videos, we choose only those that fall into the six categories: **news**, **variety show**, **music video**, **entertainment**, **animation**, and **TV**. As shown in Table 1, videos of these six types make up a considerate proportion of either amount or viewcount. We further filter out the invalid requests, such as those with viewing time of 0 and repeated requests within a short time. Since we are looking into the lifecycle of videos, we ignore videos that did not generate any traffic on Day 1, so that all videos in the cleansed datatset are guaranteed to have been uploaded no later than Day 1. With the above pre-processing, we end up with around half a million videos (nearly a third of the original volume).

| Type | % of videos | % of viewcount |
|---|---|---|
| News | 28.7 | 7.8 |
| Variety Show | 12.6 | 15.7 |
| Music Video | 9.3 | 1.9 |
| Entertainment | 8.3 | 3.2 |
| Animation | 4.9 | 14.8 |
| TV | 4.3 | 35.1 |
| **Total** | **68.1** | **78.5** |

Table 1: Proportion of 6 major types of video in the dataset




(a) Original series

(b) The frequency spectrum

(c) *RCS*'s autocorrelation

(d) *RCS* v.s. original

Figure 1: An overview of the cleansed dataset. Figure 1a shows the overall viewcount curve. As is illustrated by Figure 1b, major components in the spectrum are mainly concealed in low frequency parts. Figure 1c shows the autocorelation of the less important components in the popularity dynamics. In Figure 1d, the original series is shifted to achieve an average of zero for better visualization.

The original *Timestamp* is recorded to the second. Since we are mainly interested in the general trend of popularity dynamics within and across days, we group the timestamps of all video into bins of 10-minute intervals based on the reduced dataset, so that the popularity of each video is represented with a time series with 2016 elements. Figure 1a demonstrates the traffic of all videos in the dataset.

# 3. METHODOLOGY

### 3.1 Discrete Fourier Transform

Discrete Fourier Transform (DFT) converts a finite sequence of equally spaced samples into the coefficients of a finite combination of complex sinusoidal signals. It decomposes a time series into a combination of sinusoidal waves of certain amplitude, frequency, and phase. DFT of a discrete time series is defined as

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j\frac{2\pi kn}{N}}, k = 0, 1, 2, ..., N-1 \tag{1}$$

where $x_n$ is the $n^{th}$ sample in the time series, and $X_k$ is the complex coefficient of the $k^{th}$ frequency component. The variations of $|X_k|$ (amplitude) and $\angle X_k$ (phase) in frequency domain make up the frequency spectrum of $x_n$. Our discussion is limited to the amplitude spectrum.

### 3.2 An Overview of Amplitude Spectrum

Let $V = [1, m]$ be the set of all videos, and $T = [t_1, t_2, ..., t_n]$, $n = 2016$, be the set of timestamps (10-minute time interval as defined in Section 2.2). For each type-$i$ video $v(i, j) \in V$, we denote the original popularity series as

$$O_{i,j}(T) = \{o_{i,j}(t_1), o_{i,j}(t_2), ..., o_{i,j}(t_n)\}.$$

We first map our time series into the frequency domain by applying DFT to convert $O_{i,j}(T)$for all $v(i, j) \in V$. Figure 1b shows the corresponding amplitude spectrum of the time series in Figure 1a. Significant frequency components can be observed at lower frequencies such as $0.07, 0.14, 0.35$ and $0.71$ cycles per day, which corresponds to periods of $14, 7, 3.5$ and $1.4$ days respectively. Components at relatively higher central frequencies such as $1, 2, 3, 4$ and $5$ cycles per day correspond to periods of $24, 12, 8, 6$ and $4.8$ hours respectively. Some of the periods, such as those of $8, 6$ and $4.8$ hours, are almost impossible to be discovered in time domain.

### 3.3 Extraction of Major Frequency Components

Using frequency domain method to extract components from a time series is feasible because DFT decomposes a series into a finite combination of sinusoidal waves. For most meaningful signals that behave largely different from white noise, there will be only a small number of frequency components with significant amplitude, while the rest remain very close to zero. Empirically, the major frequency components of the spectrum largely shape the original signal, and the less dominant frequency components form the detail.

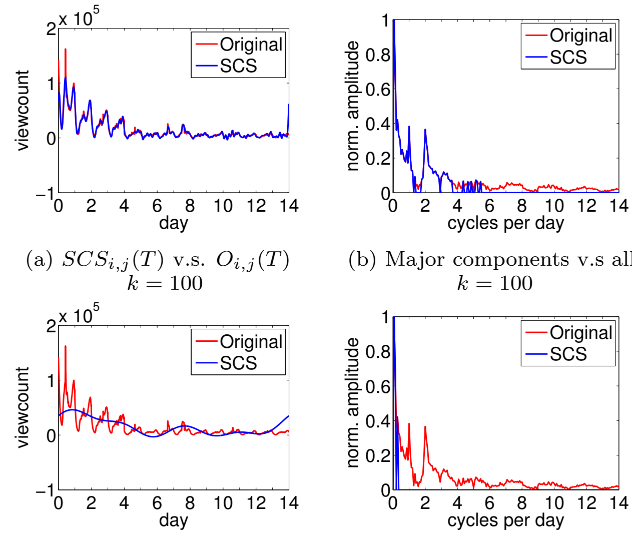
We hereby decompose the original series into two parts: **Seasonal Component Series** ($SCS$), which comprises of the major frequency components of the original popularity series, and **Residual Component Series** ($RCS$), which is the rest. $SCS$ can incorporate regular patterns occurring on any time scale. $RCS$, on the other hand, can represent irregular traffic. [8]

To extract the $SCS$, we apply the similar method as that in [8]. For each video $v(i, j) \in V$, we decompose the original time series $O_{i,j}(T)$ into seasonal and residual components through the following steps:

1. The $k$ frequency components of $DFT[O_{i,j}(T)]$ with the largest amplitudes are selected.

2. $SCS_{i,j}(T)$ is generated using inverse DFT of the $k$ components selected in Step 1, where $SCS_{i,j}(T) = \{scs_{i,j}(t) | t \in T\}$

3. $RCS_{i,j}(T)$ is obtained by subtracting $SCS_{i,j}(T)$ from $O_{i,j}(T)$: $RCS_{i,j}(T) = \{o_{i,j}(t) - scs_{i,j}(t) | t \in T\}$




(a) $SCS_{i,j}(T)$ v.s. $O_{i,j}(T)$ $k = 100$

(b) Major components v.s all $k = 100$

(c) $SCS_{i,j}(T)$ v.s. $O_{i,j}(T)$ $k = 7$

(d) Major components v.s all $k = 7$

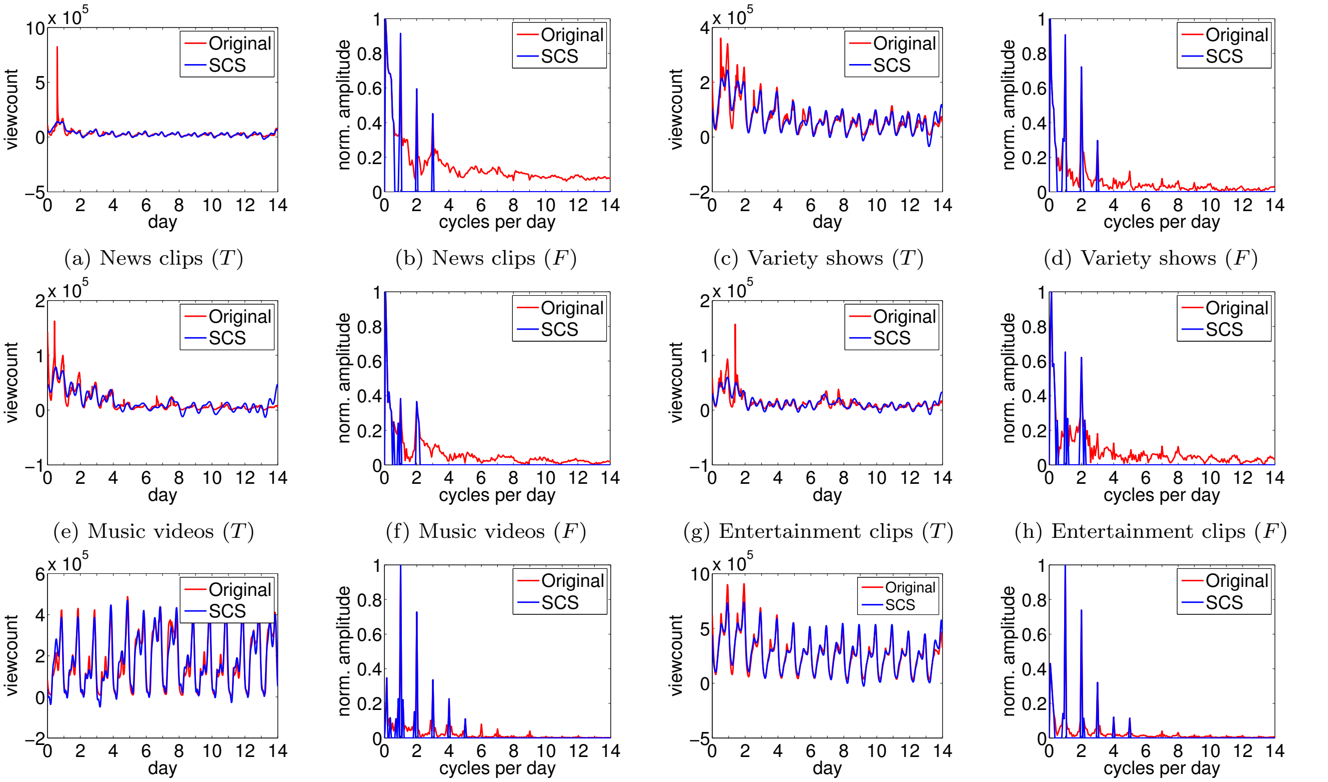
Figure 2: $O_{i,j}(T)$ vs $SCS_{i,j}(T)$ for variety shows and corresponding amplitude spectra using $k = 100$ and $k = 7$. Obviously, choice of $k$ determines the performance of extraction.

The choice of $k$ determines the performance of the decomposition. The larger $k$ is, the closer $SCS$ is to the original series. However, larger $k$ brings more components with relatively small amplitude, causing inefficiency in tracking the trend, as is shown in Figure 2a and Figure 2b. To the other end, the smaller $k$ is, the more frequency components the procedure will filter. An extremely small $k$ will even take away important components. As an illustration, Figure 2c and Figure 2d suggest that when $k$ is unreasonably small (in the case of variety shows, $k = 5$), $SCS$ fails to capture some important fluctuations in the original series. A good $k$ should retain most of the regular activities of the original series in $SCS$, while leaving the $RSC$ with little information (See Figure 1d). We set $k$ to be 25 by manually finding the smallest value such that the average autocorrelation of $RCS$ quickly drops below 0.1 (See Figure 1c).

## 4. POPULARITY DYNAMICS IN TIME AND FREQUENCY DOMAINS

### 4.1 Time Domain Analysis

Figure 3 shows the results of $SCS$ extraction. Aside from the extreme pulse-like cases, in each case, the $SCS$s fit the original series rather well. Common features in the $SCS$ of each type of videos can be easily generalized. The $SCS$ of each type has two peaks each day. The first traffic peak occurs near 1:00 PM, when viewers are enjoying their lunch break. The second appears around 10:30 PM, when viewers are just about to get prepared for bed. The position of the peaks are a little different from that of web traffic [1, 3], since nowadays not many people stick in front of computer screens during breaks and before bed time. Moreover, the latter peak tends to be much stronger than the former, suggesting




(a) News clips ($T$) (b) News clips ($F$) (c) Variety shows ($T$) (d) Variety shows ($F$)

(e) Music videos ($T$) (f) Music videos ($F$) (g) Entertainment clips ($T$) (h) Entertainment clips ($F$)

(i) Animation clips ($T$) (j) Animation clips ($F$) (k) TV ($T$) (l) TV ($F$)

Figure 3: *SCS* and corresponding frequency spectra using $k = 25$, compared with those of the original. Zero-frequency components are removed. Captions below each subfigure marks the video type. ($T$) and ($F$) marks **Time** and **Frequency** domain respectively.

that more videos are watched before bed time than during lunch breaks.

Because of the frequency filtering mechanism, *SCS* fails in following pulse-like behavior in the original series. For example, the *SCS* of news clips (Figure 3a) does not capture the traffic explosion in the afternoon of Day 1. This pulse-like burst of request brings much disturbance to the frequency spectrum, causing poor performance of *SCS* in approximating the original series. A feasible explanation is that the traffic of news clips explodes on the occurrence of unpredictable breaking news. Another explanation is that news clips stay appealing only for a short period after being uploaded, since viewers are only interested in news that occurred very recently. The latter explanation also works in the cases of variety shows (Figure 3c), music videos (Figure 3e), and entertainment videos (Figure 3g).

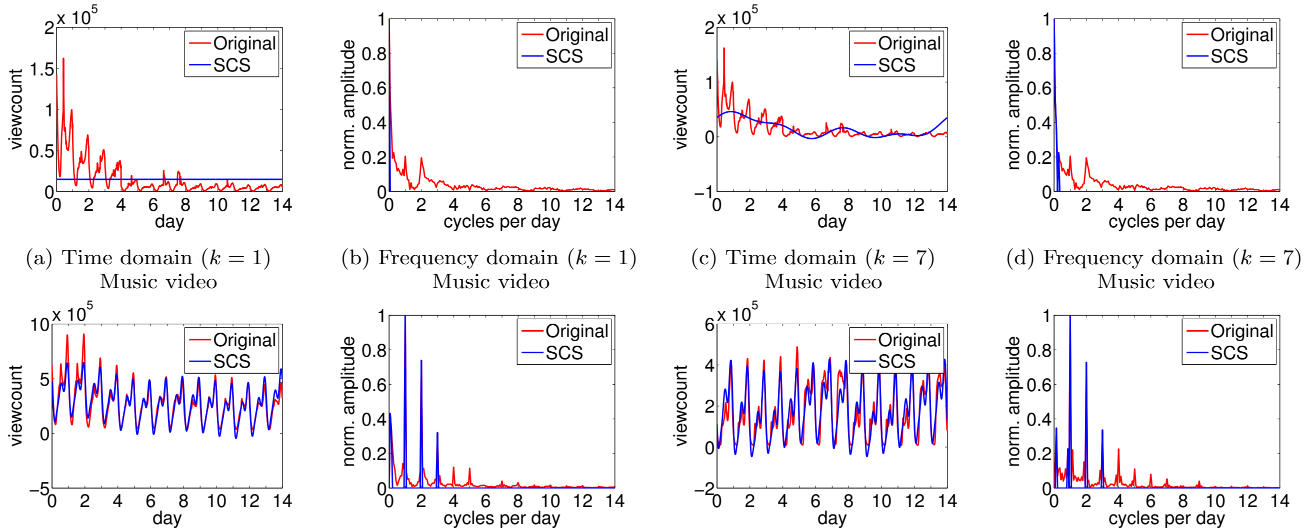
## 4.2 Frequency Domain Analysis

Mathematically, the zero-frequency component represents the average of the whole original series (see Figure 4a and Figure 4b). In some cases, the large value of zero-frequency component deteriorates the visualization effect. In most cases, we set the zero-frequency component to zero for better visualization. The value of $k$ in this subsection may be changed for the purpose of analyzing.

The component at frequency $f$ stands for a $\frac{1}{f}$-day, or a $\frac{24}{f}$-hour period in the traffic dynamics. For every type of videos, major components around central frequencies of $1, 2$ and $3$ cycles per day are observed, which corresponds to the periods of $24, 12$ and $8$ hours respectively. Significant components between 0 and 1 cycle per day are also found. Due to DFT, we are able to capture the periods like 14, 7 and 3.5 days as well as 8, 6 and 4.8 hours. These periods cannot be explicitly revealed from time domain, but can show user behavior patterns caused by different work schedule, lifestyle, content type and etc, that have not been depicted by previous methods.

Strong components at lower frequencies (between 0 and 1 cycle per day) indicate the evolution of video popularity between days. Significant components in this interval indicate severe traffic fluctuations on the daily basis. It is noted that news clips (Figure 3b), variety shows (Figure 3d), music videos (Figure 3f) and entertainment clips (Figure 3h) have significant components near 0 in their spectra. In the case of music videos (Figure 4d), the components gathering around 0 are actually the strongest in the spectrum, implying drastic fluctuation in popularity between days (Figure 4c). Actually, the request drops drastically after the day 1. The phenomenon also appears in the spectrum of news clips, whose popularity is all about timeliness. In the spectra of TV (Figure 4f) and animations (Figure 4h), however, there are only weak components within the $(0, 1)$ interval, suggesting that their popularity are relatively stable on the daily basis (Figure 4e and Figure 4g).

Major Components at higher frequencies depict popularity variation pattern within each day. Most videos have relatively higher components at 1 and 2 cycles per day, with the one at 2 slightly smaller than that at 1. This manifests




(a) Time domain ($k = 1$) Music video
(b) Frequency domain ($k = 1$) Music video
(c) Time domain ($k = 7$) Music video
(d) Frequency domain ($k = 7$) Music video
(e) Time domain ($k = 11$) TV
(f) Frequency domain ($k = 11$) TV
(g) Time domain ($k = 11$) Animation
(h) Frequency domain ($k = 11$) Animation

Figure 4: Frequency domain components and corresponding time domain popularity. $k$ is manually selected for analysis.

that the 24-hour period is much stronger than the 12-hour one. The existence of 24 and 12-hour periods are quite intuitive. However, 6, 4.8, and especially 8-hour periods are also illustrated by the spectra(all, especially Figure 3j and Figure 3l). These non-intuitive periods can hardly be observed in time domain. They indicate user behavior patterns that are have not been previously observed. These patterns may be resulted from different working schedule or lifestyle. This observation is worthy of further investigation.

The distribution of significant components across the spectrum is also a good indicator of popularity dynamics and user behavior patterns. Empirically, deviation from central frequencies are caused by multiplicative noise, which indicates small variance in user behavior patterns. Components concentrated within small margin from central frequencies guarantee regular the user behavior pattern, therefore the request is predictable. Components spreading over a wide range across the spectrum indicates poor predictability. In the case of TV and animation, the energy of their spectrum is quite concentrated (Figure 4f and Figure 4h). Their request evolves in a stable and regular way. It is foreseeable that their popularity are far more predictable then that of the other types, whose energy spoils over the spectra.

The discoveries above enlighten us that when replicating video contents that have been living for a long time, it is unnecessary to replicate videos like news clips, while is it still necessary to replicate videos like TV.

# 5. POPULARITY STABILITY

## 5.1 Grouping Criteria

In this section we focus on videos with different popularity. Videos are grouped by their viewcounts. The video traffic follows a power-law distribution, as is described in Table 2. In order to keep the video popularity variance within a group at a reasonable level, we manually classify the videos into five groups by viewcounts as Table 2 shows.

| Viewcount | % of videos |
|---|---|
| $\leq 199$ | 72.7 |
| $200 - 499$ | 8.7 |
| $500 - 999$ | 5.1 |
| $1000 - 4999$ | 8.5 |
| $\geq 5000$ | 5.0 |

Table 2: Distribution of videos based on popularity

## 5.2 Redefinition of Metrics

In previous work on human mobility [7], metrics such as **burstiness** and **memory** are introduced to evaluate the temporal distribution of human activities. However, in this work, we treat video traffic directly as time series. We therefore redefine the two following metrics:

$$B(i,j) = \frac{\sigma_{i,j}/\mu_{i,j} - 1}{\sigma_{i,j}/\mu_{i,j} + 1} \tag{2}$$

$$M(i,j) = \frac{1}{n-k} \sum_{r=1}^{n-k} \frac{(o_{i,j}(t_r) - \mu_{i,j}^{(1)})(o_{i,j}(t_{r+k}) - \mu_{i,j}^{(2)})}{\sigma_{i,j}^{(1)} \sigma_{i,j}^{(2)}} \tag{3}$$

where $\mu_{i,j}$ and $\sigma_{i,j}$ are the average and the standard deviation of $O_{i,j}(T)$ respectively, and $\mu_{i,j}^{(\theta)}$ and $\sigma_{i,j}^{(\theta)}$ are the average and the standard deviation of the series consisting of the heading ($\theta = 1$) or the tailing ($\theta = 2$) $n - k$ elements in $O_{i,j}(T)$ respectively.

The range of $B(i,j)$ is $[-1, 1]$. It increases monotonously as $\sigma_{i,j}/\mu_{i,j}$ grows, which is the normalized variance in popularity from the average traffic. When directly applied to time series, $B$ correlates negatively with the long-term stability of video popularity. Larger value of $B$ suggests larger deviation in video request from the average level, therefore less long-term stability of popularity.

$M(i,j)$ is the Pearson Correlation Coefficient between the time series excluding the last $k$ values and that excluding the first $k$. The range is $[-1, 1]$ by definition. $M$ will have a larger value only when the intervals with a lot of traffic are followed by the intervals with considerate request. It measures the stability of popularity within $k$ time slots. In other words, $M$ correlates positively with the short-term stability of video popularity. Here, we set $k$ to be 1.

### 5.3 Popularity Stability Analysis

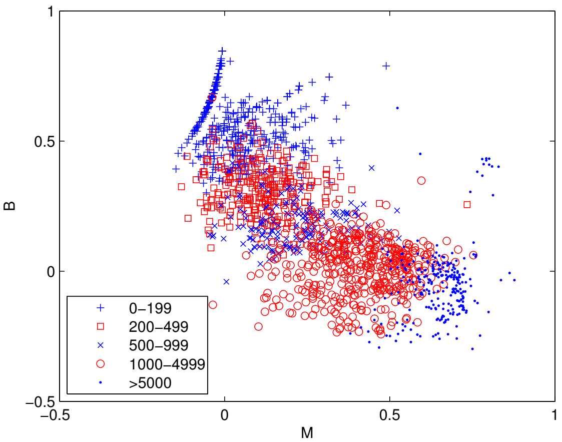



Figure 5: $M$ v.s. $B$ for a random sample of all videos

In Figure 5, we randomly select a subset of the $(B, M)$ pairs of all videos in our dataset. Videos in the same group are mapped to $(B, M)$ pairs in neighboring regions on the plain, showing cluster-like features.

Generally, popular videos tend to have bigger $M$ values and $B$ values around zero (mapped to the lower right corner in Figure 5), indicating a good stability both short-term and long-term. This shows that most of the popular videos stay attractive for a long time. Unpopular videos tend to have smaller $M$ values and bigger $B$ values (mapped to the upper left corner), suggesting significant instability both short-term and long-term. This implies that unpopular videos failed to gain stable traffic. The traffic of these videos remained dormant for most of the time, with several sporadic requests. This discovery intuitively suggests higher priority for popular videos when caching.

We can actually see a borderline to the upper left corner in Figure 5. These $(B, M)$ pairs marks videos with the most unstable popularity both long-term and short-term. Most of such videos have the lowest total viewcount. However, a few videos with relatively considerate request are also spotted on the curve. Manual check shows that at some mysterious times, a flood of requests arrived at these relatively popular ones. There is a small cluster of popular videos are found near $(0.4, 0.8)$, suggesting good short-term stability but poor stability over the long period. Judging from video type and traffic pattern, these are mostly news clips covering breakout events. These videos remained heated for a short period of time, but eventually faded out just like other news clips. The combination of $B$ and $M$ provide us with a novel view on the definition of popularity.

The discoveries above enlighten us that it is unnecessary to replicate unpopular videos, while is it of great necessity to replicate popular videos except for news clips.

## 6. CONCLUSIONS AND FUTURE WORK

We take the first step towards understanding video popularity data as discrete time signals. Popularity series of major types of videos is decomposed and qualitative analysis is conducted on the Seasonal Component Series as well as major frequency components of popularity series. The frequency domain methods does well in following the popularity dynamics and capturing non-obvious patterns, such as periods of 8,6 and 4.8 hours. We then redefined metrics to evaluate the long-term and short-term stability of popularity series. In most cases, the combination of the two metrics are highly correlated with total traffic of a certain video. Our findings provide reasonable guideline for the design of content replication systems.

We have shown that DFT is a powerful tool in time series analysis. However, DFT evaluates the distribution of frequency components of the whole time series, paying not attention to the distribution of amplitude in time. In our future work, we plan to exploit wavelet or short-term Fourier transform to get a better understanding of user behavior and to predict the video popularity. Other features in our dataset, such as viewing time, will be added into analysis. We are also working on the popularity dynamics of list of videos, such as episodes of the same TV or variety shows.

## 7. ACKNOWLEDGMENTS


This work is supported in part by Natural Science Foundation of China (No. 61402114), and Shanghai Pujiang Program (No. 14PJ1401400).


## 8. REFERENCES


[1] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng. Understanding user behavior in large-scale video-on-demand systems. In *Proc. of ACM EuroSys'06*, pp. 333–344, Leuven, Belgium, 2006.

[2] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proc. of ACM IMC'07*, pp. 1–14. San Diego, CA, USA, 2007.

[3] G. Chatzopoulou, C. Sheng, and M. Faloutsos. A first step towards understanding popularity in youtube. In *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pp. 1–6. San Diego, CA, USA, 2010.

[4] H. Yu, L. Xie, and S. Sanner. Twitter-driven youtube views: Beyond individual influencers. In *Proc. of ACM MM'14*, pp. 869–872. Orlando, FL, USA, 2014.

[5] H. Li, X. Ma, F. Wang, J. Liu, and K. Xu. On popularity prediction of videos shared in online social networks. In *Proc. of ACM CIKM'13*, pp. 169–178. San Francisco, CA, USA, 2013.

[6] H. Abrahamsson and M. Nordmark. Program popularity and viewer behaviour in a large tv-on-demand system. In *Proc. of ACM IMC'12*, pp. 199–210. Boston, MA, USA, 2012.

[7] K.-I. Goh and A.-L. Barabási. Burstiness and memory in complex systems. *EPL*, 81(4):48002, 2008.

[8] B. Cici, M. Gjoka, A. Markopoulou, and C. T. Butts On the decomposition of cell phone activity patterns and their connection with urban ecology. In *Proc. ACM MobiHoc'15*, pp. 317–326. Hangzhou, China, 2015.