Physica A 453 (2016) 236–241

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Characterizing popularity dynamics of online videos

Zhuo-Ming Ren [a,*], Yu-Qiang Shi [b], Hao Liao [c,a]

[a] Department of Physics, University of Fribourg, Chemin du Musée 3, CH-1700, Fribourg, Switzerland

[b] School of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, PR China

[c] Guangdong Province Key Laboratory of Popular High Performance Computers, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, PR China

HIGHLIGHTS

- A temporal analysis of the popularity dynamics in two online video-provided websites.
- Dynamics of the online video popularity can be characterized by the burst behaviors.
- The burst behaviors typically occur in the early life span of videos.
- Lately the online video popularity restricts to the classic preferential mechanism.

ARTICLE INFO

Article history:
Received 9 August 2015
Received in revised form 9 January 2016
Available online 24 February 2016

Keywords:
Popularity dynamic
Online video
Burst behavior

ABSTRACT

Online popularity has a major impact on videos, music, news and other contexts in online systems. Characterizing online popularity dynamics is nature to explain the observed properties in terms of the already acquired popularity of each individual. In this paper, we provide a quantitative, large scale, temporal analysis of the popularity dynamics in two online video-provided websites, namely MovieLens and Netflix. The two collected data sets contain over 100 million records and even span a decade. We characterize that the popularity dynamics of online videos evolve over time, and find that the dynamics of the online video popularity can be characterized by the burst behaviors, typically occurring in the early life span of a video, and later restricting to the classic preferential popularity increase mechanism.

1. Introduction

The online individual behaviors for videos, music, news and other contexts are implicated in the online popularity whose breadth surpasses individual awareness [1–3]. The online behaviors can predict the popularity trend, and then the online services such as Netflix, Youtube, Facebook would benefit greatly from an in-depth understanding of the popularity trend to optimize strategies for update and recommendation [4,5]. For example, in Netflix, if the evolution of the video popularity is predictable in advance, the Netflix service can replicate more copies of videos with increasing popularity and reduce replication of videos losing popularity to make space for the hot videos. Meanwhile, the online behavior information produces a big amount of time stamped data, making it possible to study the dynamics of the online popularity and how it evolves over time on a global scale [2,6]. From a theoretical point of view, the massive amount of available data from these online services provides an unprecedented opportunity to understand the online popularity dynamics. Hence, characterizing

* Corresponding author.
E-mail addresses: zhuomingren@gmail.com (Z.-M. Ren), hao.liao@unifr.ch (H. Liao).

http://dx.doi.org/10.1016/j.physa.2016.02.019
0378-4371/

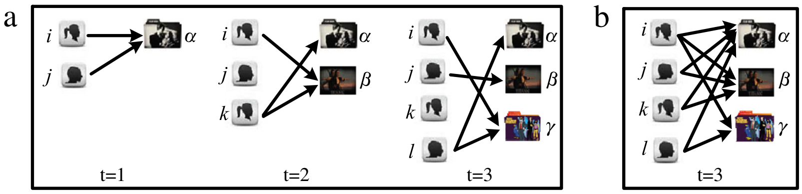

Fig. 1. An example of the relationship between user and video which represents a bipartite network with the temporal information. (a) The temporal bipartite networks. The new user and the new video enter the bipartite network as the time varying. At the time point $t = 2$ and $t = 3$, a new video and a new user are added to the network respectively. The new video is more popular than the old ones, and is selected by the two users, but the old video is selected by one user. (b) shows that historical snapshot until the time $t = 3$. The older video α has been selected four times. The old one β has been selected three times. The new one γ has been selected two times. The example indicates that the older the video is, the more accumulative but the less temporal popular it will be.

the online popularity dynamics is necessary for explaining the observed features in terms of the already acquired online popularity of each individual and predicting the popularity trend. The behaviors of the online popularity dynamics have been widely studied in the literature focused on videos [7–9], music [10], news [11], and other online social collective dynamics [12,13]. Cha et al. [14] found that a high linear correlation existed between the number of video views on early days and later days in YouTube. Borghol et al. [15] performed an empirical study to measure the popularity of videos, and argued a strong linear rich-get-richer behavior with the number of previous views as the most important factor [16]. Shen et al. [17] employed the reinforcement Poisson mechanism documenting the well-known rich-get-richer phenomenon to model the popularity dynamics. Compared to a stronger presence of the rich-get-richer phenomenon, Vasconcelos et al. [18] observed a lower correlation between the early and late popularities in Foursquare. Chen et al. [19] considered the lifetime perspective of video popularity, and found that the relative popularity of a video was not only dependent on its age, but also the type of the video. Specifically, the online video of popularity is by no means restricted to rich-get-richer behaviors. It can depend on exogenous attributes of the online contexts which make popularity dynamics suitable for applications in many different contexts in the online systems [20]. Furthermore, Ratkiewicz et al. [21] examined the popularity of Wikipedia topics and Web pages and presented a model that combines the classic preferential popularity increase mechanism with the occurrence of random popularity shifted capturing the influence of exogenous factors on online popularity. There also have some works based on the exogenous attributes to result in changes to modeling the popularity dynamics in different online contexts such as user-generated videos [22], the citation of scientific release [23,24], and the activity of scientists [25]. While, few works are centralized when the online contexts emerge suddenly popularly in their life span. In this paper, we conduct an in-depth study on the online popularity dynamics in two online video-provided websites, namely MovieLens [26] and Netflix [27], who both invite users to rate videos. The two big time stamped data sets contain over 100 million rating records and even span a decade, which provide an opportunity to understand the dynamics of online popularity and how they evolve over time on a global scale. We characterize that how video rating behaviors of the online popularity dynamics evolve over time, and find that the dynamics of rating popularity are characterized by the burst behaviors (i.e. far exceed preferential popularity increase). Typically, the presences of rating burst occur in the early life span of a video, and later restrict to the classic preferential popularity increase mechanism.

2. Model and method

Here, we use bipartite networks with temporal information to represent the online rating systems which include the set of users (denoted by $U(t)$), the set of videos (denoted by $O(t)$) and the records between users and videos (denoted by $L(t)$). A link in the bipartite network connecting user i and video α represents a historical record $l_{i\alpha}$ ($\in L$). If a video has been chosen by more users than others, we could see this video is more popular than others. Thus, the popularity of the video $k_\alpha(t)$ is defined as the number of rating records which the video α was received at time t. We give a simple example in Fig. 1 to show how to construct a temporal bipartite network. The new video gets more temporal popularity than the old videos (at time point $t = 3$, the older video α has 1 records, the old video β has 1 records, and the new video γ has 2 records) as displayed in Fig. 1(a). While as shown in Fig. 1(b), the older videos get more accumulative popularity than the newer objects' (until time $t = 3$, the older video α has 4 records, the old video β has 3 records, and the new γ has 2 records).

Preferential attachment [28] focuses on the fact that the rich gets richer, which indicates that the popular videos get more popularity. While, as the time varying, there always emerge new popular videos, and some old popular videos perform less popularly than the new popular videos. The fitness model [29] describes that the ability between the new node and the old node to compete for popularity. Numerous examples [21,23,24] convincingly indicate that in real systems the new one is easy to get more popularity in the early life span. For example, the online rating systems provide services or products (i.e., Netflix, MovieLens for videos, Amazon for books/other products, alibaba, Ebay e-business platform for goods). The users could prefer to pay more attention at the appearance of new videos, new news or new goods in the online rating systems, then the new products immediately get a lot of attention and become popular. The burst is considered as accumulations of a large number of rapidly occurring events during short time intervals which is observed in many real systems such as human dynamics, citation dynamics and other online collective dynamics [30]. To measure the burst of production, Goh et al. [31]