Neurocomputing 205 (2016) 529–541

Contents lists available at ScienceDirect

Neurocomputing

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

Boosting video popularity through keyword suggestion and recommendation systems

Renjie Zhou [a,b], Samamon Khemmarat [c], Lixin Gao [c], Jian Wan [a,b], Jilin Zhang [a,b,*], Yuyu Yin [a,b], Jun Yu [a,b]

[a] School of Computer Science and Technology, Hangzhou Dianzi University, Zhejiang, China
[b] Key Laboratory of Complex Systems Modeling and Simulation, Ministry of Education, China
[c] Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, 01003 MA, USA

ARTICLE INFO

Article history:
Received 26 July 2015
Received in revised form
7 April 2016
Accepted 2 May 2016
Communicated by Tao Mei
Available online 20 May 2016

Keywords:
Video popularity
View propagation
Keyword suggestion
Recommendation system

ABSTRACT

YouTube offers a great opportunity for people to entertain, advertise, gain popularity, and generate revenue. How to increase views for a video has become the key question for anyone who wish to be famous or gain more revenue. Recognizing that a recommendation system is a major view source for videos, our goal in this paper is to increase video views in YouTube by leveraging on the recommendation system. We first perform measurements to understand how views are propagated among videos through recommendation links and identify factors that influence the recommendation produced by the system. Our measurement results show that similarity in video meta-data is a crucial ingredient in connecting videos. We then propose a keyword suggestion method for a video with the aim to raise video views through the recommendation system. The keyword suggestion method utilizes video clusters on a referrer video graph to obtain relevant keywords and ranks keywords based on both their relevance and their potential to attract video views. The effectiveness of the keyword suggestion method is demonstrated through a case study, showing that using the keywords suggested by our method leads to a larger number of video views and higher average watching time per video playback compared to initial user-given keywords.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Social media websites have been experiencing explosive developments in the recent years. According to the Alexa statistics, among the top ten most popular websites around the world, five are social media websites. Content creators, advertisers and anyone who wish to gain fame and generate revenue through social media websites are interested in the popularity of their contents. Most of popular social media websites deploy search engines to assist users in finding desired items from a huge collection of items. When it comes to YouTube, the search engine retrieves a video via its textual description. However, the textual information of a video is usually incomplete, inconsistent or even unavailable, because it is a time-consuming task for content uploaders to describe their videos appropriately, completely and attractively. Consequently, the search engine becomes inefficient and thus the visibility of poorly described videos decreases.

Recommendation systems are currently broadly utilized as a part of online social networking and e-commerce websites. These recommendation systems exploit the clues from the history of users' behavior and content metadata to predict the items that will capture the interests of users. For instance, the collaborative filtering algorithm measures similarity between items by the times of co-access, and thus it is able to recommend an item to users even without textual data of item itself. Therefore, recommendation systems play a critically important role in helping users to discover interesting contents. For example, it has been shown in our previous work [1] that approximately thirty percent of the views on YouTube are contributed by the YouTube recommendation system.

We focus on a common scenario where an item is shown with a recommendation list containing related items. An example of this scenario is the related video recommendation on YouTube, where a list of related videos is shown to the user on the right of video playing page. While a user usually starts the viewing process from search engines, homepage, or external links embedded on other websites, s/he probably continues watching videos through clicking on one of the related videos. Thus, a single video view is likely to be followed by a series of video viewing on the site under the impact of the recommendation system. In other words, a view received at a video may propagate to its related videos. The goal of the paper is to study this propagation process systemically, which

* Corresponding author.
E-mail address: jilin.zhang@hdu.edu.cn (J. Zhang).

http://dx.doi.org/10.1016/j.neucom.2016.05.002
0925-2312/© 2016 Elsevier B.V. All rights reserved.

would allow one to better utilize the recommendation system to help users find videos of their interests and meanwhile boost the popularity of videos.

With the goal in mind, we propose a model which captures the mechanism underlying the propagation of views through a recommendation system, allowing us to investigate how a video's popularity impacts popularity of other videos via these recommendation links [2]. Quantifying the influence between videos is the basis of identifying the most influential videos in terms of contributed views. By examining influential videos, we are able to identify the characteristics of videos that are potentially beneficial to a video. Taking advantage of this knowledge, we investigate the possibility of improving metadata of a video to induce recommendation linkage with potentially advantageous videos. Finally, we propose a keyword suggestion algorithm, which will help uploaders compose the textual description of a video and also increase the popularity of his/her videos. The contributions of the paper are summarized as follows.

- A Markov-based view propagation model is proposed to estimate the number of views one video contributes another video. Based on the model, we investigate the roles of factors like distance between videos, the number of injected views in impacting the view propagation. The results show that approximately two thirds of propagated views are from the direct top referrers and nearly 15% of propagated views are from the indirect top referrers. Furthermore, we find that the number of injected views of a referrer video (means the number of views coming from sources except recommendation system) is the most important factor in determining the number of views propagated from it in general. These results imply that it is the most beneficial to attach a video directly to referrer videos with a large number of views.
- It is found that the similarity of meta-data between videos plays a significant role in establishing the connection between videos. Approximately 65% of the videos contain the video with the highest title/tag similarity in their related video lists. This result further suggests that it is possible to use the title/tag similarity in the attempt to create recommendation linkage between videos.
- A keyword suggestion algorithm is proposed for suggesting keywords that are relevant to a topic given by a user and have high potential to attract views to a video. We demonstrate that videos with similar topics tend to form a cluster in the referrer video graph induced by the recommendation system and exploit this property to obtain keywords relevant to a topic. Then, a metric is proposed for ranking keywords based on both their relevance and potential to attract views. Our case study experiment demonstrates effectiveness of the keyword suggestion algorithm in helping users discover videos of their interests and increasing video views as well.

The paper is organized as follows. A view propagation model for estimating influence between videos is proposed in Section 2. The dataset and our method for analyzing the influence between YouTube videos is described in Sections 3 and 4, respectively. Section 5 investigates the roles of factors in impacting view propagation. Section 6 examines how the titles, tags and categories of videos affect the formation of recommendation linkage between videos. In Section 7, we describe our keyword suggestion algorithm and demonstrate its effectiveness. Related work is discussed in Section 8 and the paper is concluded in Section 9.

2. View propagation model

In this section, we describe a view propagation model which captures the mechanism underlying the propagation of views through recommendation links between videos. The model allows us to derive influence between each pair of videos even though there are no direct recommendation links between them. This model is a basis for our investigation of the influence of one video's popularity to other videos' popularity through the recommendation system in the later sections.

Consider a common scenario where an item is shown with a recommendation list containing recommended items. For example, a video page on YouTube shows a list of videos related to the video. A user normally reaches the first item via sources except the recommendation lists, then s/he may continue to view another item in the recommendation list of the first item and continue on. Our interest in this paper is to gain insights into how one item's popularity affects any other item's popularity via these recommendation links. Usually, it is not difficult to obtain the click through rate from item i to any item j in i's recommendation list. However, the click through rate is not sufficient for satisfying our need as it only captures the influence between items that are directly connected by recommendation links. Therefore, we propose a view propagation model which captures the mechanism underlying the propagation of views through recommendation links, allowing us to derive influence between items, even though there are no direct links between them.

Let us now introduce several terminologies. As mentioned before, it is common that an item i is shown along with a recommendation list. A user viewing item i may continue to view any item j in is' recommendation list. Here, item i is named as a *referrer item* of item j. The view item j received from its referrer item i through the recommendation list is called *a propagated view*. A view an item received from sources except the recommendation list is defined as an *injected view*.

Let P be the transition probability matrix of the view propagation process. It is a square matrix of order N, where N equals the number of items in a system, and each entry $P(i,j)$ represents the click through rate from i to j. Where, $0 \le P(i,j) \le 1$, and $P(i,j) = 0$ when j is not in is' recommendation list. Let F be the $N \times N$ influence coefficient matrix, where each entry $F(i,j)$ represents the coefficient of influence from i to j, meaning the probability that a user who views item i will watch item j by clicking on recommendation lists one by one. It is necessary to mention that matrix P and F are different, since $P(i,j)$ is the probability of directly clicking to watch item j on is' recommendation list after viewing item i.

The the view propagation process can be described as follows. It starts with an injected view of item i, then it will either propagate a view to one of the items in the recommendation list of i or stop propagating, in accordance with the transition probability matrix P. The propagation continues iteratively until stop propagating. Given the number of injected views and the transition probability matrix P, we can obtain the number of propagated views between items by iterating the view propagation. Let V_I be the row vector of injected view count, V_P be the row vector of propagated view count, and $V_P^{\langle k \rangle}$ be the row vector of propagated view count in kth iteration. Then, we can derive $V_P^{\langle 1 \rangle} = V_I P$ and $V_P^{\langle k \rangle} = V_P^{\langle k-1 \rangle} P$. The total propagated view count of a video equals the summation of propagated view count from every iteration. Thus, we have

$$V_P = \lim_{n \to \infty} \sum_{k=1}^{n} V_P^{\langle k \rangle} = \lim_{n \to \infty} \sum_{k=1}^{n} V_I P^k. \qquad (1)$$

The series converges [3] if $\|P\|_\infty < 1$. This condition is equivalent to that, the sum of the click through rates originated from each item i is less than one. The condition is usually satisfied in reality because not all users will continue to view items in the recommendation

Download English Version:

https://daneshyari.com/en/article/405724

Download Persian Version:

https://daneshyari.com/article/405724

Daneshyari.com