How is Attention Allocated? Data-driven Studies of Popularity and Engagement in Online Videos

Siqi Wu
Australian National University & Data 61, CSIRO, Australia
Canberra, Australia
siqi.wu@anu.edu.au

ABSTRACT

The share of videos on Internet traffic has been growing, e.g., people are now spending a billion hours watching YouTube videos every day. Therefore, understanding how videos capture attention on a global scale is also of growing importance for both research and practice. In online platforms, people can interact with videos in different ways – there are behaviors of active participation (watching, commenting, and sharing) and that of passive consumption (viewing). In this paper, we take a data-driven approach to studying how human attention is allocated in online videos with respect to both active and passive behaviors. We first investigate the active interaction behaviors by proposing a novel metric to represent the aggregate user engagement on YouTube videos. We show this metric is correlated with video quality, stable over lifetime, and predictable before video's upload. Next, we extend the line of work on modelling video view counts by disentangling the effects of two dominant traffic sources – related videos and YouTube search. Findings from this work can help content producers to create engaging videos and hosting platforms to optimize advertising strategies, recommender systems, and many more applications.

KEYWORDS

YouTube; engagement; popularity; empirical measurement

ACM Reference Format:

Siqi Wu. 2019. How is Attention Allocated? Data-driven Studies of Popularity and Engagement in Online Videos. In *The Twelfth ACM International Conference on Web Search and Data Mining (WSDM '19), February 11–15, 2019, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3289600.3291599

1 INTRODUCTION

Multimedia is central to modern online platforms in which people can view, discuss and share such content. In this paper, our focus is on the largest online video hosting site – YouTube, which currently attracts a billion hours of watch time every day [10]. There has been a rich body of literature on studying online videos. However, while current research has extensively measured and modelled video popularity, engagement measures, sometimes referred as "post-clicking behaviors" [5], remain largely unexplored in academia despite gaining growing attention in industry. Hence, we consider online video attention (behaviors) as two classes: popularity (viewing) and engagement (watching, commenting, and sharing). Intuitively, our notions of popularity and engagement describe the decision to *click* on a video and the decision to *interact* after clicking.

Modelling the mechanism of attention allocation from observational data is challenging due to multiple confounding factors. The data is often presented asynchronously and heterogeneously. In practice, the YouTube eco-system reacts to *exogenous* stimulus such as Google search [1]. The signal then amplifies and steers through its own recommender systems, creating an *endogenous* response [4]. Previous studies have focused on analyzing the semantic structure within each individual time series data and revealed the strong correlation between early and future popularity [7]. However, the analysis on the effects of two dominant sources of views – related videos and YouTube search [11] – is missing in current research.

Specifically, our work endeavors to answer two research questions in the context of online videos:

RQ1: How to measure the video engagement? Can it be predicted?
RQ2: How to model the attention flow on YouTube eco-system?

We address the first question by presenting the first large-scale measurement on video-level aggregate engagement [9]. In contrast to prior work that requires an auxiliary toolkit [6], our data collection strategy raises the data volume by several orders of magnitude. We study a set of metrics including time and completion rate of a video being watched, and further propose a new metric, *relative engagement*, which is calibrated against video length and closely correlated with recognized notions of quality. We also find that engagement measures are stable over time and predictable even before video's upload. The result is significant as it separates the concerns for modelling engagement and popularity – the latter is known to be unstable and driven by external promotions [8].

The second question is to model and predict video popularity. We curate a dataset that records the traces of related videos from recommender systems, search volume inside YouTube and mentions on Twitter. We intend to exploit this dataset to answer the question of how attention flows from one video to another.

A better understanding of how online videos attract human attention would provide us with a set of useful tools to improve user experience in online platforms. For content producers, the observations can help them choose engaging topics to create videos, or adjust advertising strategies for higher rewards. For hosting platforms, it can help prioritize quality content in recommender systems. The impacts of this work go beyond quantitative observations. New datasets, research software and real-world applications are developed to facilitate future studies upon online videos.

WSDM '19, February 11–15, 2019, Melbourne, VIC, Australia

ACM ISBN 978-1-4503-5940-5/19/02.
https://doi.org/10.1145/3289600.3291599

2 RELATED WORK

Measuring engagement in online content. Many researchers have analyzed engagement behaviors towards web content. For example, the line of work that measures web page reading patterns often exploits auxiliary toolkit such as mouse-tracking browser [6]. In search engine and recommender systems, dwell time, which is conceptually close to video watch time, has been widely used [3]. All the above works focus on engagement with an individual user. However, user-level data is often unavailable to content producers on YouTube. Our work measures engagement at an aggregate level, as complementary to individual engagement study.

Modelling popularity in online content. For online videos, the most studied attributes is the popularity dynamics, defined as the number of times they are viewed. A number of models have been proposed to describe it, such as endogenous relaxations [4], epidemic contagion [1], and external stimulation [8]. On social media sites like Twitter and Facebook, information diffusion patterns are extensively studied. Many works leverage different sets of features and methods to predict future cascade, with network structure and past success being the two most notable factors [2].

3 DATASETS

The initial challenge is the limitation of available data. To empower our data-driven analysis, a new software "`YouTube-insight`" for crawling integrated YouTube data is developed and released [1]. We have curated two new datasets for answering **RQ1** and **RQ2**, respectively. The first is 5M Tweeted Videos dataset, which is not only one of the largest public YouTube datasets, but also one of the few datasets that contain information about user watching behavior. The second is Vevo Music Graph dataset, featuring licensed English songs on YouTube. It contains video metadata, popularity time series, snapshots of video recommendation network, associated YouTube search and Twitter mentions over several years. This new attention dataset bridges the information gap across multiple web services and allows us to study more complex patterns towards understanding human attention in online videos.

4 RESEARCH OVERVIEW

4.1 Measuring and predicting engagement

Recent studies show that the quality of a digital item is linked to the audience's decision to continue watching after first opening it [5]. Therefore, the average amount of time that the audience spend on watching a video should be indicative of video quality. However, we also observe that duration is an important covariate on video watching patterns. Intuitively, longer videos are less likely to be fully watched compared to shorter videos due to the limited human attention span. Here we define watch percentage as the average ratio of video being watched, bounded between 0 and 1.

To calibrate the effect of video duration, we first construct a 2-dimensional tool – *engagement map* – that captures the non-linear relationship between video duration and watch percentage. Based on it, we propose a new metric *relative engagement* as the watch percentage rank percentile among videos of similar lengths. We show that relative engagement closely correlates to external notions of video quality. Relative engagement values are high for content deemed of good quality by the experts. Furthermore, we find that the temporal dynamics of relative engagement appears to be stable over time. This implies that early watch pattern is a strong predictor for future engagement, which prompts us to examine whether engagement can be predicted *before* upload. We then setup regression tasks to predict video engagement. We find engagement metrics predictable in a cold-start setup, having most of its variance explained by video context, content topics and channel information at coefficient of determination R^2 of 0.77. For details, we refer to our previous work [9].

[1] Available at https://github.com/avalanchesiqi/youtube-insight

4.2 Modelling and predicting popularity

This section provides a brief description for our future work. In online platform, products are rarely independent with each other. On YouTube, users switch videos by clicking the recommendation list on the right-hand side bar, which contributes as one of the main sources of video views [11]. The original video and associated recommended videos naturally form a directed network and human attention flows through the linkages in such network. The aim of this work is to provide an interpretable manner for predicting video popularity and quantitatively measure the effects of YouTube recommender systems on steering human attention.

5 SUMMARY

The focus of my ongoing PhD research is to study collective human behaviors in online videos, specifically, how is attention allocated on the largest video hosting site YouTube. We have measured a set of aggregate engagement metrics and proposed a new metric *relative engagement*, which appears to be stable over the video lifetime, closely correlate with video quality and help the task of predicting engagement in a cold start setup. Our future work studies human attention in a more complex social system by incorporating the structure of recommendation network and internal search volume. Broadly, this work makes contributions to the data mining and computational social science community.

REFERENCES

[1] Christian Bauckhage, Fabian Hadiji, and Kristian Kersting. 2015. How Viral Are Viral Videos?. In *ICWSM*.

[2] Justin Cheng, Lada Adamic, P Alex Dow, Jon Kleinberg, and Jure Leskovec. 2014. Can cascades be predicted?. In *WWW*.

[3] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *RecSys*.

[4] Riley Crane and Didier Sornette. 2008. Robust dynamic classes revealed by measuring the response function of a social system. *PNAS*.

[5] Coco Krumme, Manuel Cebrian, Galen Pickard, and Sandy Pentland. 2012. Quantifying social influence in an online cultural market. *PloS one* (2012).

[6] Dmitry Lagun and Mounia Lalmas. 2016. Understanding user attention and engagement in online news reading. In *WSDM*.

[7] Henrique Pinto, Jussara M Almeida, and Marcos A Gonçalves. 2013. Using early view patterns to predict the popularity of youtube videos. In *WSDM*.

[8] Marian-Andrei Rizoiu, Lexing Xie, Scott Sanner, Manuel Cebrian, Honglin Yu, and Pascal Van Hentenryck. 2017. Expecting to be hip: Hawkes intensity processes for social media popularity. In *WWW*.

[9] Siqi Wu, Marian-Andrei Rizoiu, and Lexing Xie. 2018. Beyond views: Measuring and predicting engagement in online videos. In *ICWSM*.

[10] YouTube. 2017. You know what's cool? A billion hours. https://youtube.googleblog.com/2017/02/you-know-whats-cool-billion-hours.html.

[11] Renjie Zhou, Samamon Khemmarat, and Lixin Gao. 2010. The impact of YouTube recommendation system on video views. In *IMC*.