The Tube Over Time: Characterizing Popularity Growth of YouTube Videos

Flavio Figueiredo **Fabricio Benevenuto** **Jussara Almeida**

Universidade Federal de Minas Gerais, Belo Horizonte – MG - Brazil

{flaviov,fabricio,jussara}@dcc.ufmg.br

Abstract

In this work, we characterize the growth patterns of video popularity on YouTube. Using newly provided data by the application, we analyze how the popularity of individual videos evolves since the video's upload. Also with this data, it was also possible to characterize the types of the referrers (external links) that most often attracted users to each video. We compared popularity evolution and referrals for three different datasets: videos which were promoted, copyrighted and selected according to a random procedure.

Evolution of Popularity in YouTube

Problem:

- How does video popularity evolve?
- How do users find these videos?

Basic Statistics:

- YouTube search is the second most queried search engine (2009)
- 2 Billion views per day (2010)

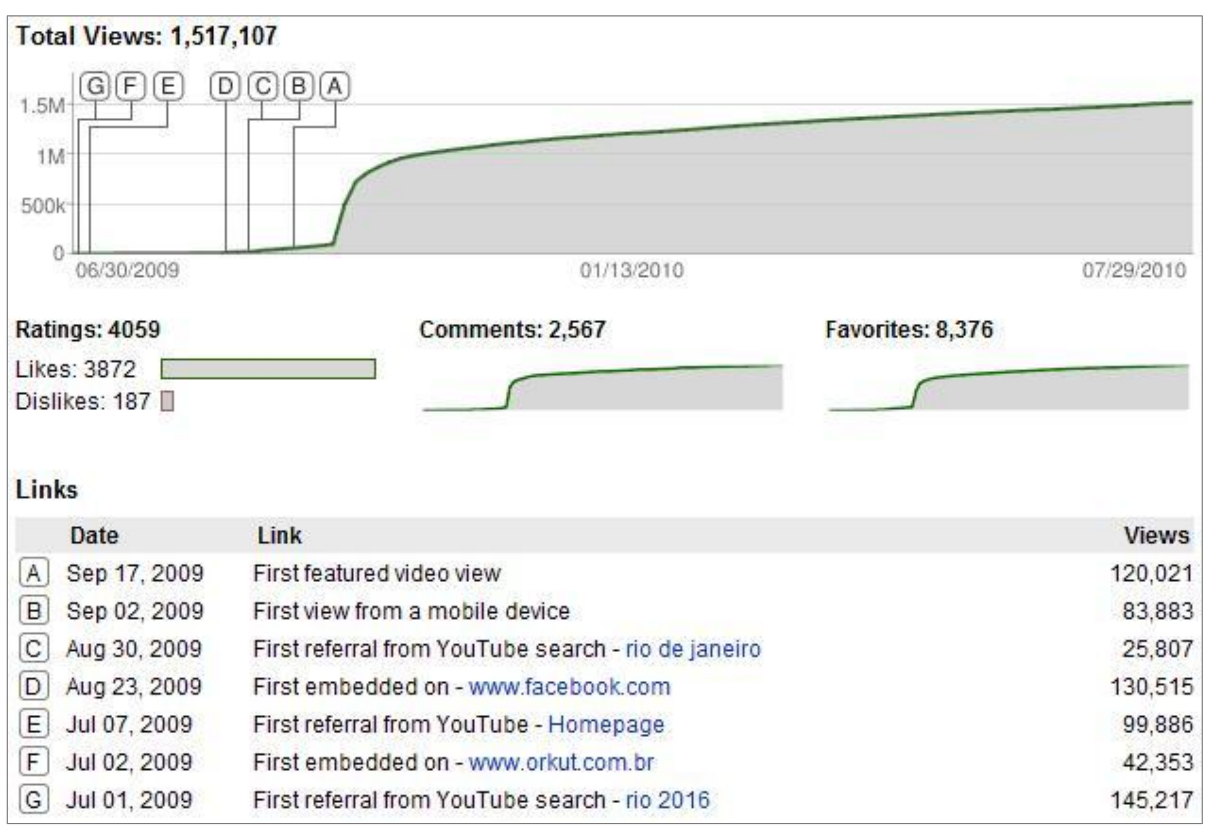
For whom is this important?

- Content creator and video partners
- Internet providers, caching services, CDNs
- Online marketers

Data Collection

We collected the cumulative growth in popularity and top ten referrers for each video

- **Top:** Videos which appeared on top lists (promoted)
- **YouTomb:** Copyrighted videos
- **Random:** Selected based on random queries



Statistics provided by YouTube

Table 1: Crawled datasets.

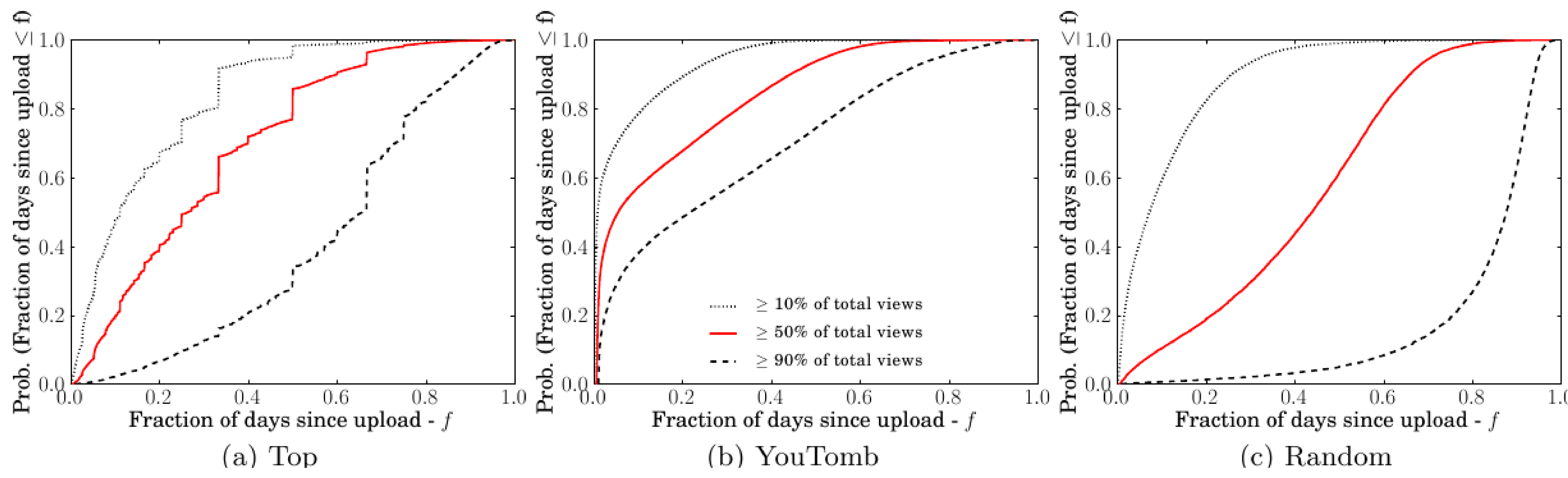
Video Datasets	Top	YouTomb	Random
# Videos	17,127	73,257	18,095
Avg # of views/videos	843,001	279,486	126,056
Average age (# days)	136	627	493

How Early Does a Video Get Popular?

Copyrighted videos (YouTomb) are generally consumed faster than Top and Random ones. For half of the videos:

- On YouTomb they take at most 21% of lifetime to reach 90% of final popularity
- On Top this value is 65%
- On Random it is 87%

This work was funded by the Brazilian National Institute of Science and Technology for Web Research (MCT/CNPq/INCT Grant 573871/2008-6), and by individual grants from CAPES, CNPq and FAPEMIG.



Cumulative distribution of the fraction of time until a video reaches at least 10%, 50% and 90% of its total views

Temporal Dynamics Model

We classified videos based on the model by Crane and Sornett [1], which takes the number of views on the peak day/week:

$$\text{Viral} \longrightarrow views_{peak} \leq t \quad (1)$$

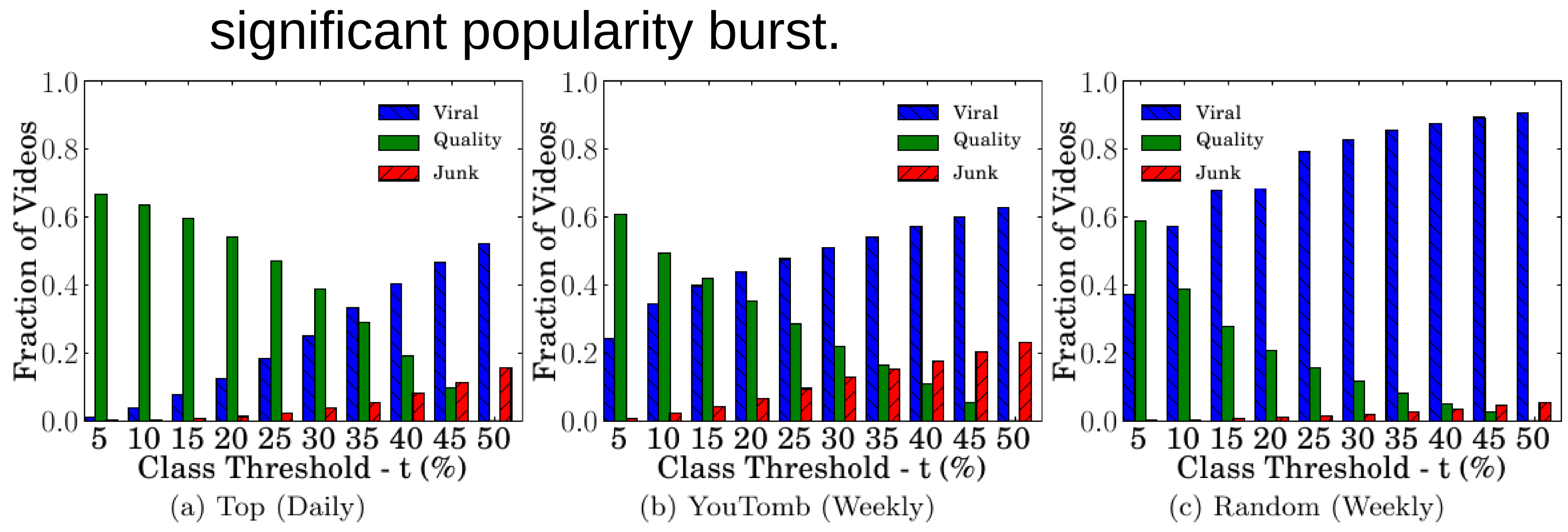
$$\text{Quality} \longrightarrow t < views_{peak} \leq (1-t) \quad (2)$$

$$\text{Junk} \longrightarrow views_{peak} > (1-t) \quad (3)$$

Results

Considering a 20% threshold model:

- Most Random and YouTomb videos are viral
- For Top videos, most are of quality, exhibiting a significant popularity burst.



Fraction of videos in each class

Referrer Importance

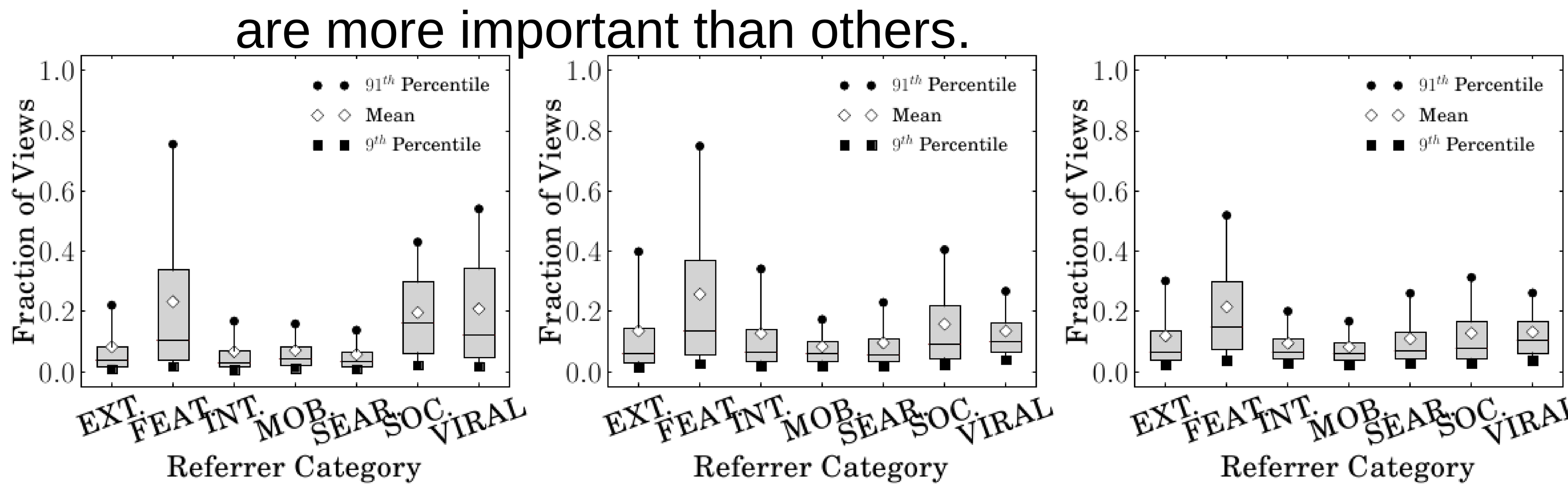
How do users reach videos on different datasets?

Considering datasets as whole (table omitted):

- Browsing is responsible for **29%**, **36%** and **18%** of views for Top, YouTomb and Random datasets
- Search for **20%**, **35%** and **37%** of views for Top, YouTomb and Random datasets

How important are referrers for each video?

- When present, Featured, Social and Viral referrers are more important than others.



Fraction of views each referrer was responsible for per video

Conclusions

- Copyrighted videos tend to get most views earlier
- Copyrighted & random videos exhibit viral propagation
- Search & browsing are important referrers
- Social &featured referrers are important when present

References

[1] R. Crane and D. Sornette. Robust dynamic classes revealed by measuring the response function of a social system. Proceedings of the National Academy of Sciences (PNAS), 105(41):15649-15653.