Viral Video Style: A Closer Look at Viral Videos on YouTube

Lu Jiang, Yajie Miao, Yi Yang, Zhenzhong Lan,
Alexander G. Hauptmann

School of Computer Science, Carnegie Mellon University

Carnegie Mellon University

Carnegie Mellon University

Outline

- Introduction
- CMU Viral Video Dataset
- Statistical Characteristics
- Peak Day Prediction
- Conclusions

Carnegie Mellon University

Outline

- Introduction
- CMU Viral Video Dataset
- Statistical Characteristics
- Peak Day Prediction
- Conclusions

What is a viral video?

- A viral video is a video that becomes popular through the process of (most often) Internet **sharing** through **social media**.

WIKIPEDIA
The Free Encyclopedia

Gangnam Style

What is a viral video?

- A viral video is a video that becomes popular through the process of (most often) Internet sharing through social media.

WIKIPEDIA
The Free Encyclopedia

Charlie bit my finger

Gangnam Style

Missing pilot MH370

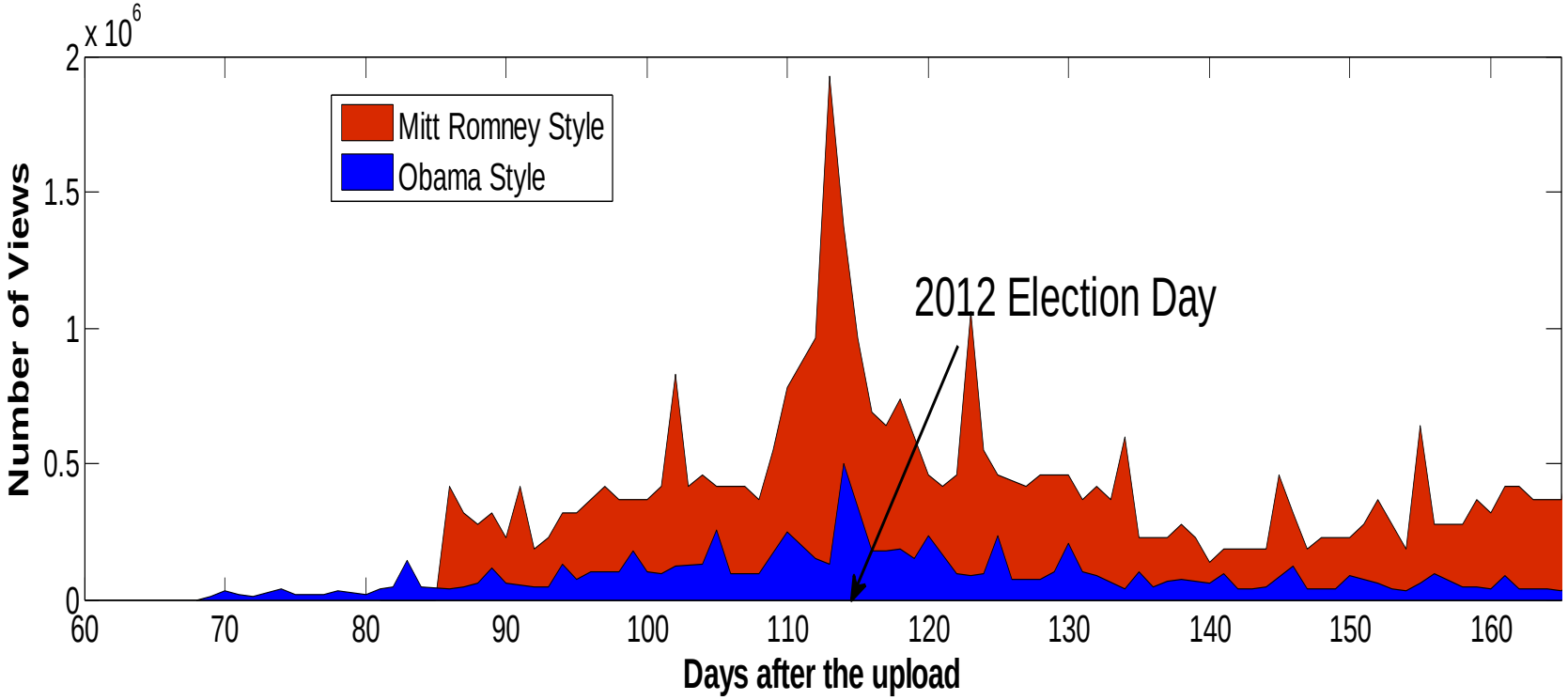
Social Validity

- Viral videos have been having a **profound impact** on many aspects of society.
- Politics:
  - Pro-Obama video “Yes we can” went viral (10 million views) in 2008 US presidential election [Broxton 2013].

Social Validity

- Viral videos have been having a profound impact on many aspects of society.
- Politics:
  - Pro-Obama video “Yes we can” went viral (10 million views) in 2008 US presidential election [Broxton 2013].
  - We found that Obama Style and Mitt Romney Style went viral (30 million views in the month of Election Day), and **peaked on Election Day**.

Social Validity



Social Validity cont.

- Viral videos have been having a profound impact on many aspects of society.
- **Financial marketing**:
  - Old Spice's campaign went viral and improved the brand's popularity among young customers[West et al 2011].

Social Validity cont.

- Viral videos have been having a profound impact on many aspects of society.
- Financial marketing:
  - Old Spice's campaign went viral and improved the brand's popularity among young customers[West et al 2011].
  - Psy's commercial deals has amounted to 4.6 million dollars from Gangnam Style.

Motivation

- Existing studies are conducted on:
  - Small set (tens of videos) → biased observations?
  - Large-scale Google set → confidential!
- A relatively large and public dataset on viral videos would be conducive.
- Solution: CMU Viral Video Dataset.

Beware the content are hilarious!

Carnegie Mellon University

Outline

- Introduction
- CMU Viral Video Dataset
- Statistical Characteristics
- Peak Day Prediction
- Conclusions

CMU Viral Video Dataset

- By far the **largest** public viral video dataset. Time's list was the largest dataset (50 videos).
- Videos are **manually selected by experts**
  - Editors from Time Magazine, YouTube and the viral video review episodes.
- Statistics: 446 viral videos, 294 quality 19,260 background videos.

10+ million subscribers!

Carnegie Mellon University

CMU Viral Video Dataset Cont.

- For a video, it includes:
  - Thumbnail
  - Video and user metadata

CMU Viral Video Dataset Cont.

- For a video, it includes:
  - Thumbnail
  - Video and user metadata
  - Insight data: historical views, likes, dislikes, etc.

Date	Day	View
12/18/2012	0	1296022
12/19/2012	1	15552276
12/20/2012	2	8025373
12/21/2012	3	3987763
12/22/2012	4	2990822
12/23/2012	5	2392658
12/24/2012	6	2093575

CMU Viral Video Dataset Cont.

- For a video, it includes:
  - Thumbnail
  - Video and user metadata
  - Insight data: historical views, likes, dislikes, etc.
  - Social data: #Inlinks, daily tweeter mentions (pending)

[Number of In-links]

624000

[#Tweets mentioned]

Date	#Tweets with Video ID
12/18/2012	0
12/18/2012	79
12/19/2012	14
12/20/2012	3
12/21/2012	7

CMU Viral Video Dataset Cont.

(a) ID: rYEDA3JcQqw (b) ID:mBRUkdQa6Is

(c) ID: _OBlgSz8sSM (d) ID: Y13oB-IhE7I

– Near duplicate videos (automatic detection + manual inspection).

Carnegie Mellon University

Outline

- Introduction
- CMU Viral Video Dataset
- Statistical Characteristics
- Peak Day Prediction
- Conclusions

Carnegie Mellon University

Statistical Characteristics

Table 3: Basic statistics about viral videos.

Statistics	Viral	Quality	Background
View Count Median	3,079,011	55,455,364	7,528
Title length	5.0±0.1	5.4±0.1	7.0±0.1
Duration(s)	138.6±16.0	248±3.9	252±24.6
Average Rate	4.69±0.03	4.75±0.03	4.04±0.08
Rater/View	0.54±.03%	0.38±.01%	0.87 ±.07%
Cor(inlinks, view)	0.54	0.25	0.28
Days-to-peak Median	24	63	30
Lifespan Median	7	166	10

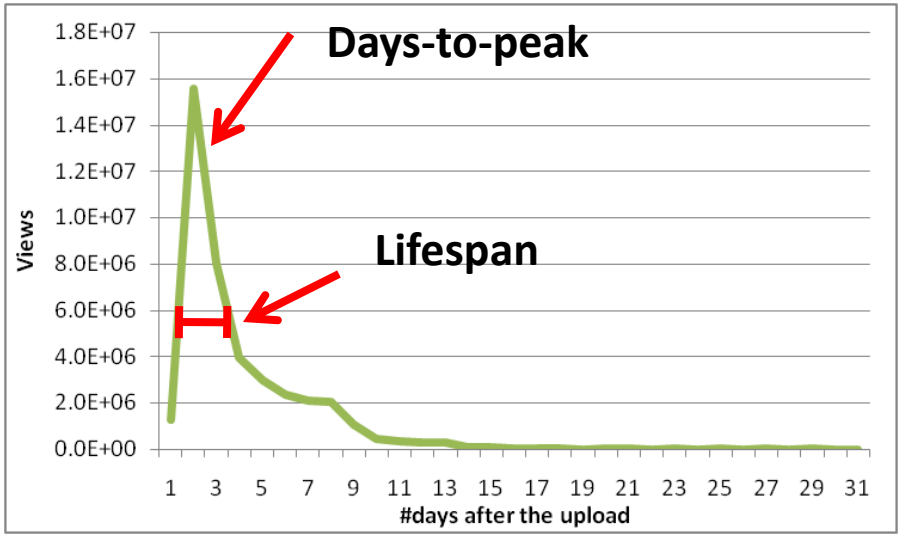
- Observations agree with existing studies including the study on Google's dataset
  - Short title, Short duration.
- Less biased.

Observation I

- *The days-to-peak and the lifespan of viral videos decrease over time.*

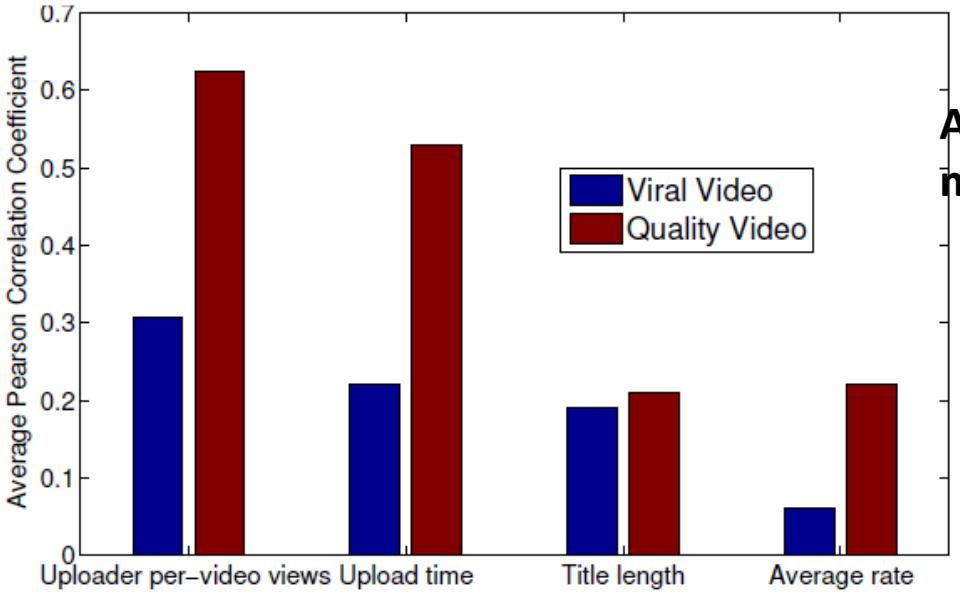
Table 4: Evolution of viral videos.

Statistics	2010	2011	2012
Days-to-peak Median	24	14	9
Lifespan Median	8	6	3



Observation II

- *The popularity of the uploader is a more substantial factor that affects the popularity than the upload time.*
- Upload time is believed to be the most important factor in background videos[Borghol et al. 2012], coined as First-mover advantage.



An example: official music videos

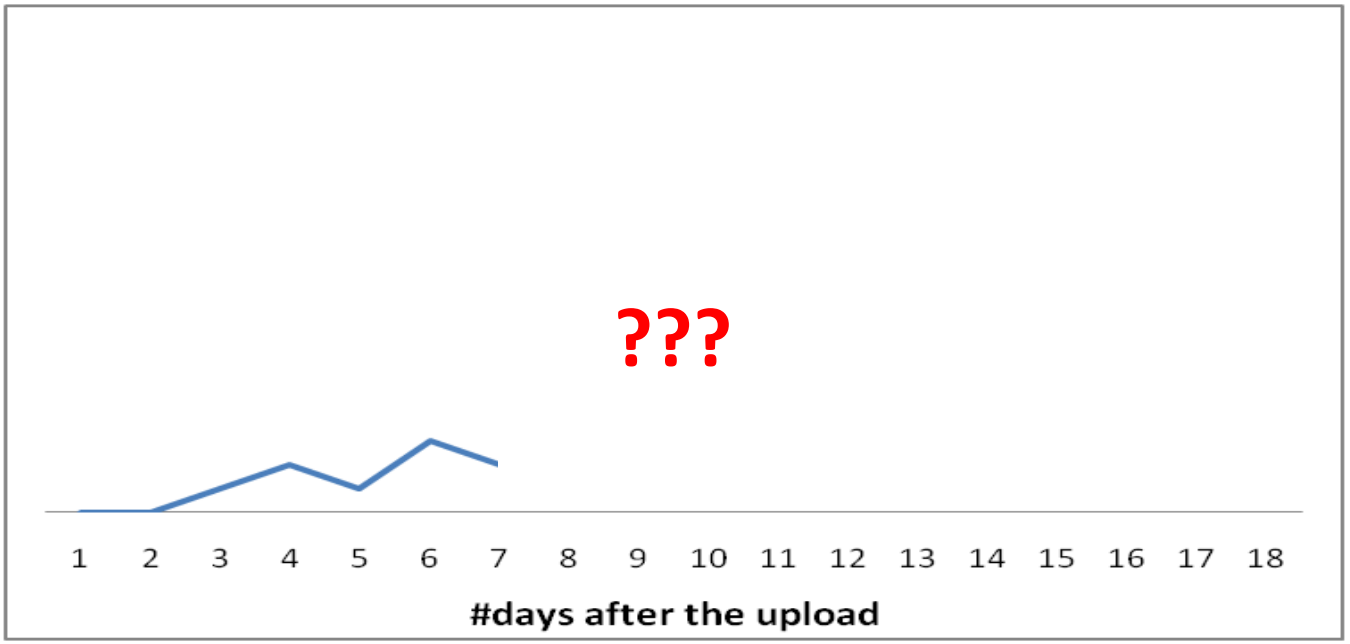
Carnegie Mellon University

Outline

- Introduction
- CMU Viral Video Dataset
- Statistical Characteristics
- Peak Day Prediction
- Conclusions

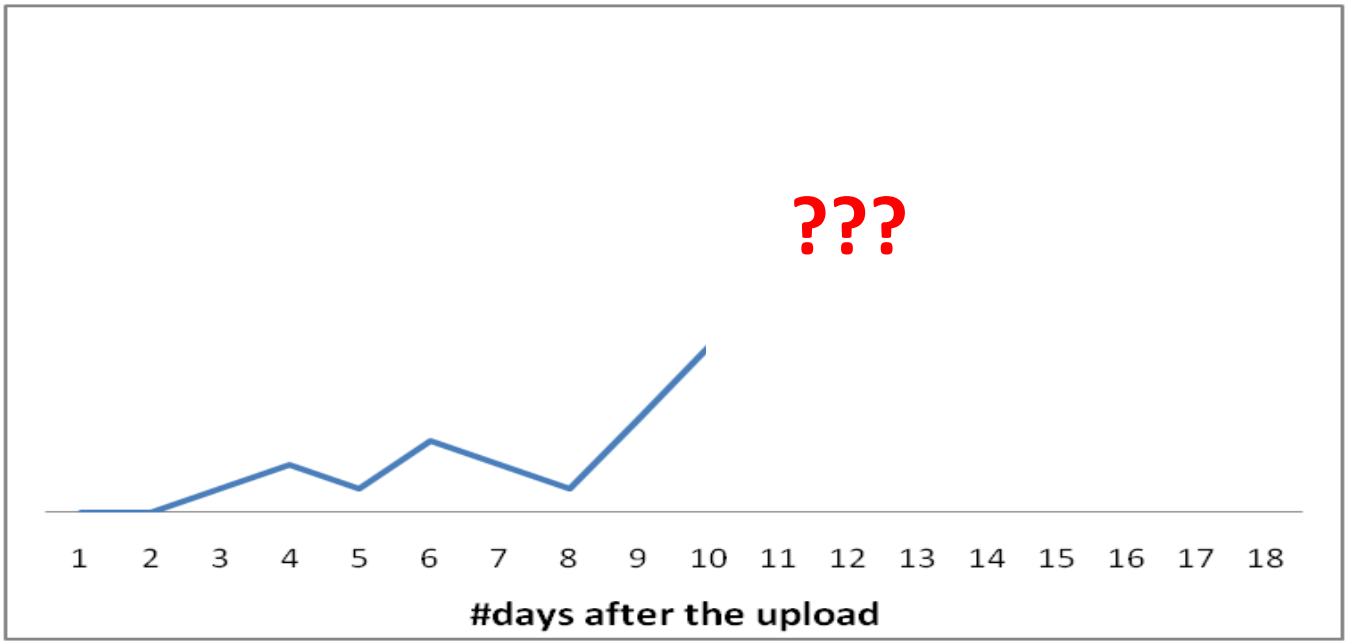
Peak Day Prediction

- Forecast when a video can get its peak view based on its historical view pattern.



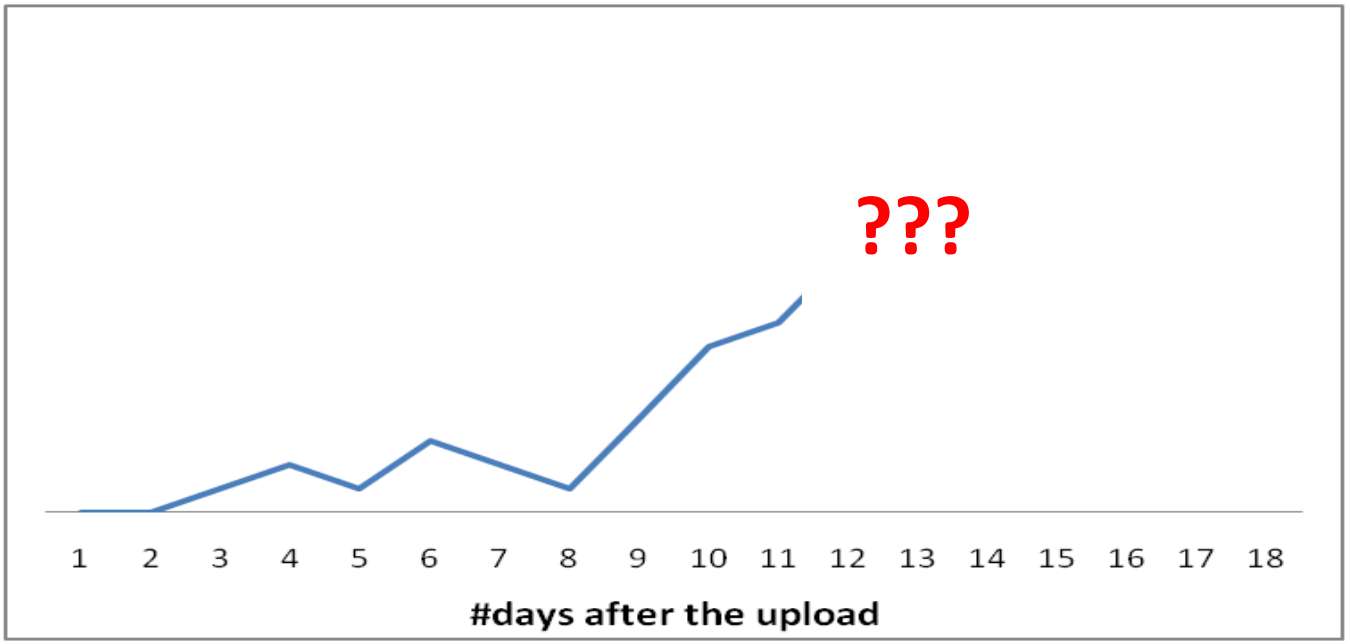
Peak Day Prediction

- Forecast when a video can get its peak view based on its historical daily view pattern.



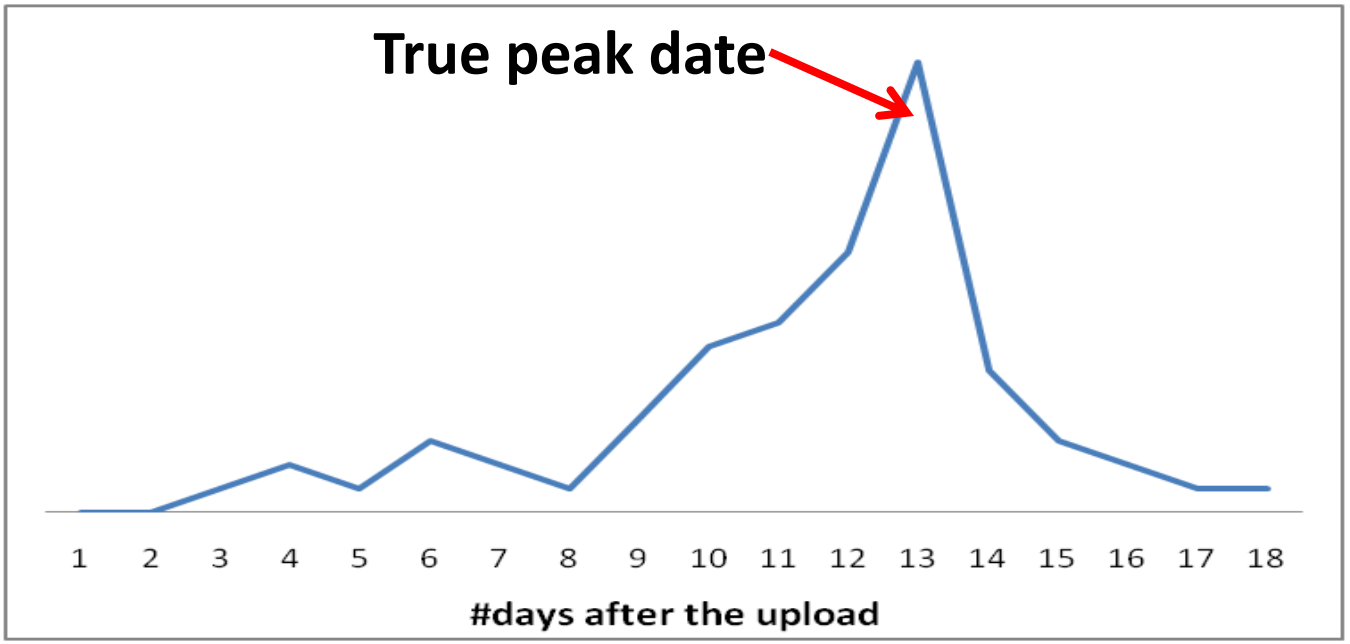
Peak Day Prediction

- Forecast the date a video can get its peak view based on its daily view pattern.



Peak Day Prediction

- Forecast when a video can get its peak view based on its historical daily view pattern.



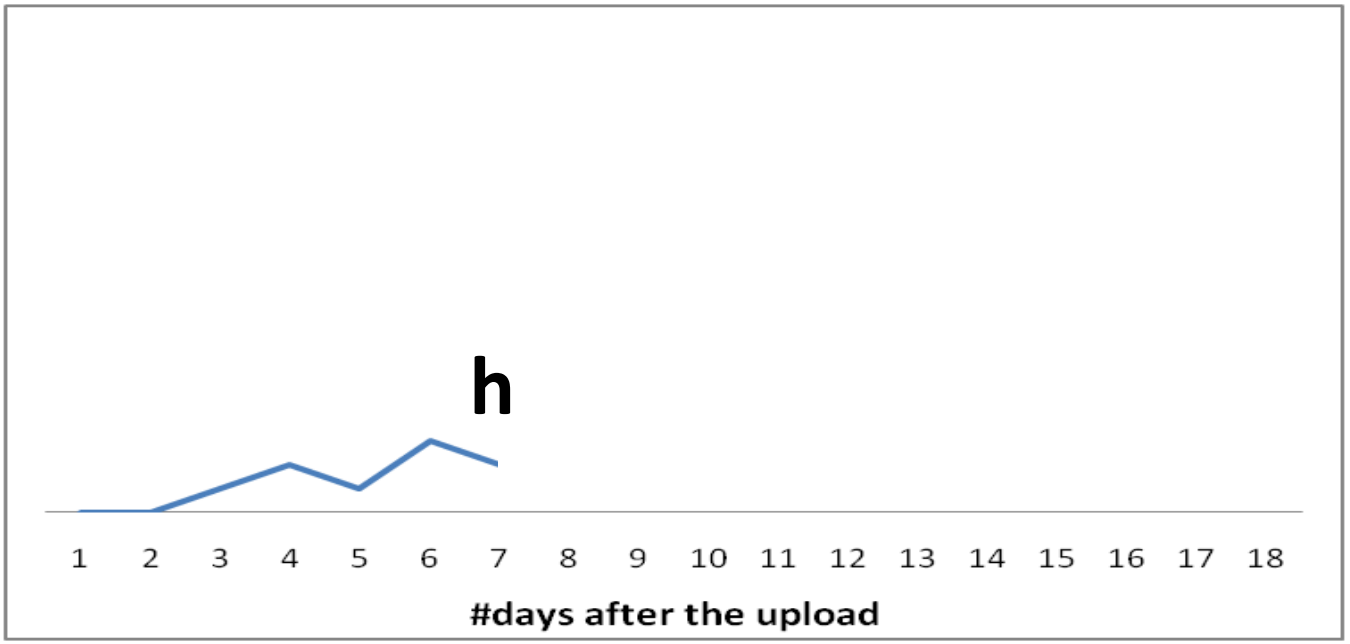
Carnegie Mellon University

Peak Day Prediction

- Forecast the date a video can get its peak view based on its daily view pattern.
- This application is **significant** in supporting and driving the design of various services:
  - Advertising agencies: determine timing and estimate cost
  - YouTube: recommendation
  - Companies/Politicians: respond viral campaigns

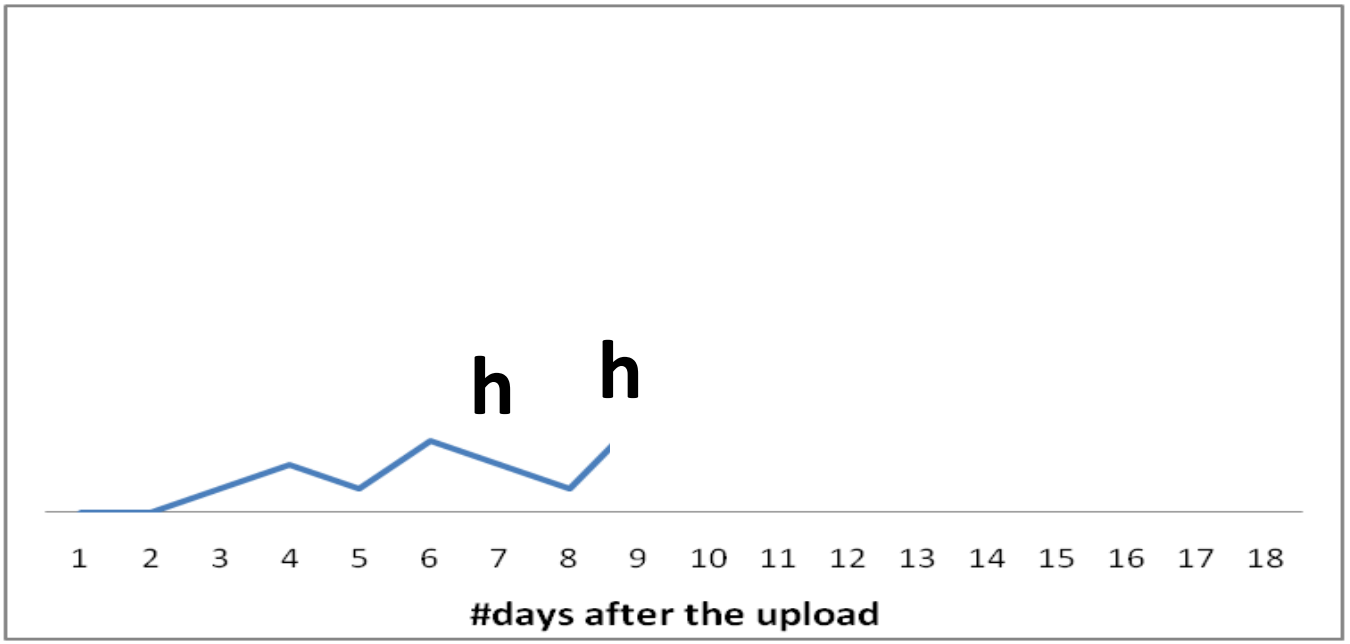
HMM Model

- Model daily views using HMM model:
- Two types of states:
  - hibernating → less views
  - active → more views



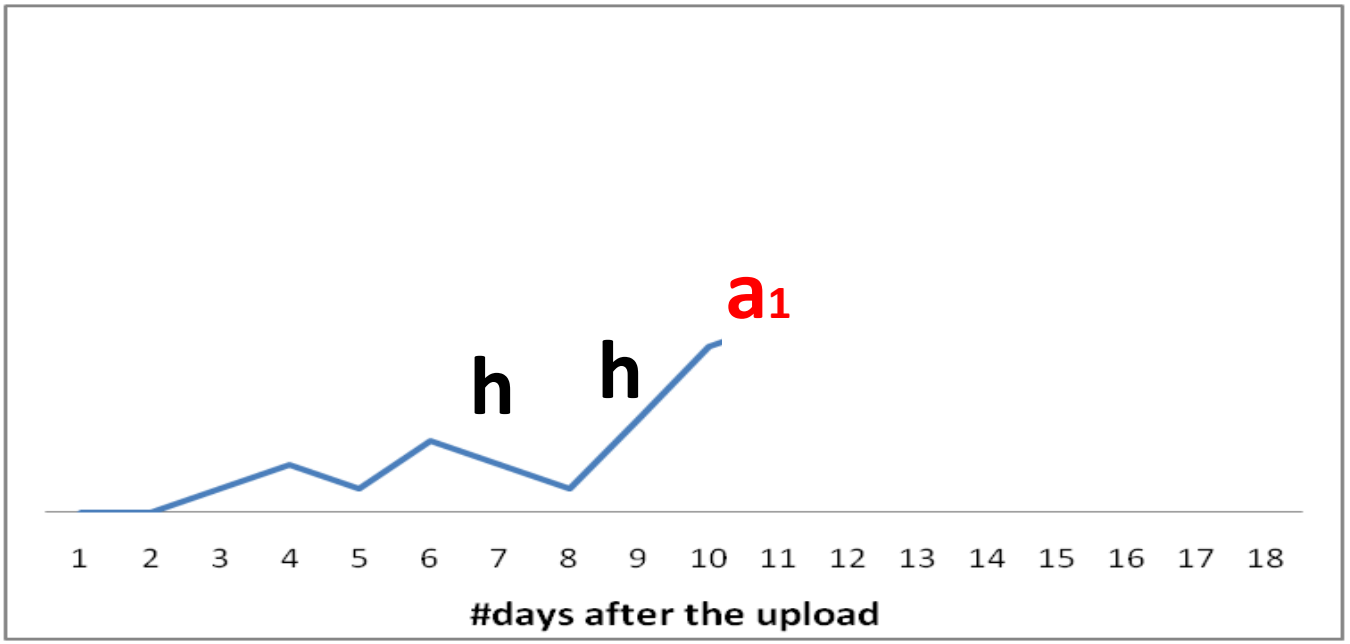
HMM Model

- Model daily views using HMM model:
- Two types of states:
  - hibernating → less views
  - active → more views



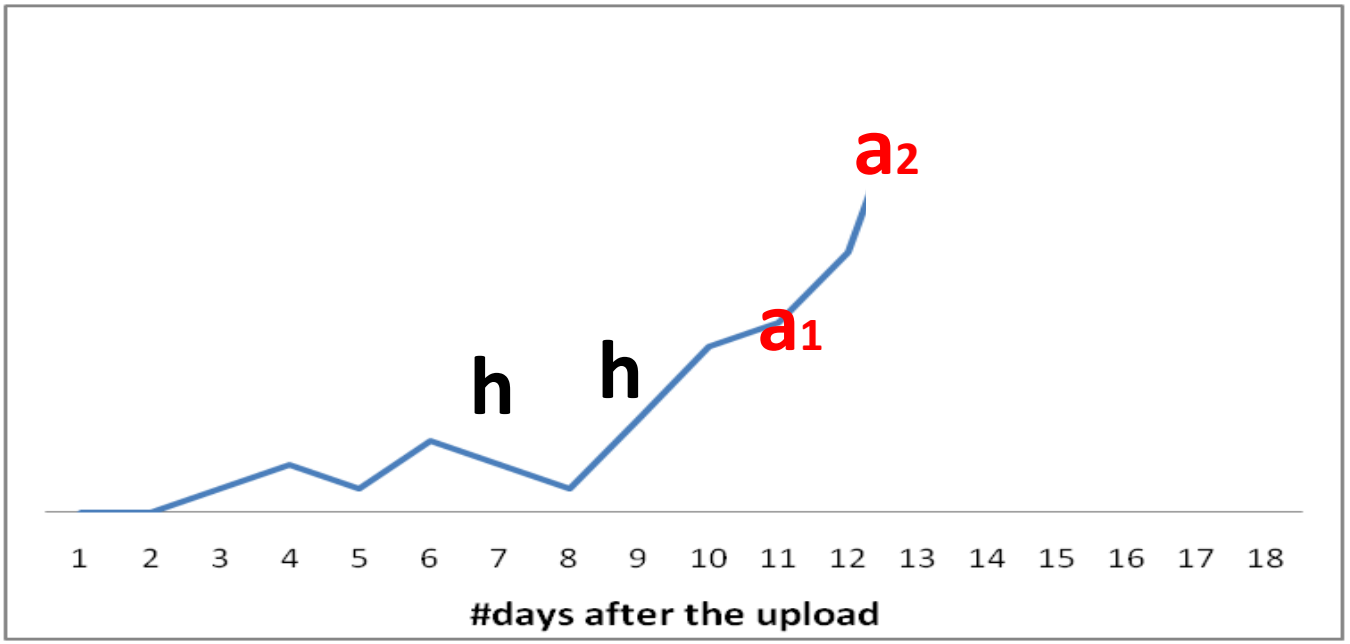
HMM Model

- Model daily views using HMM model:
- Two types of states:
  - hibernating → less views
  - active → more views



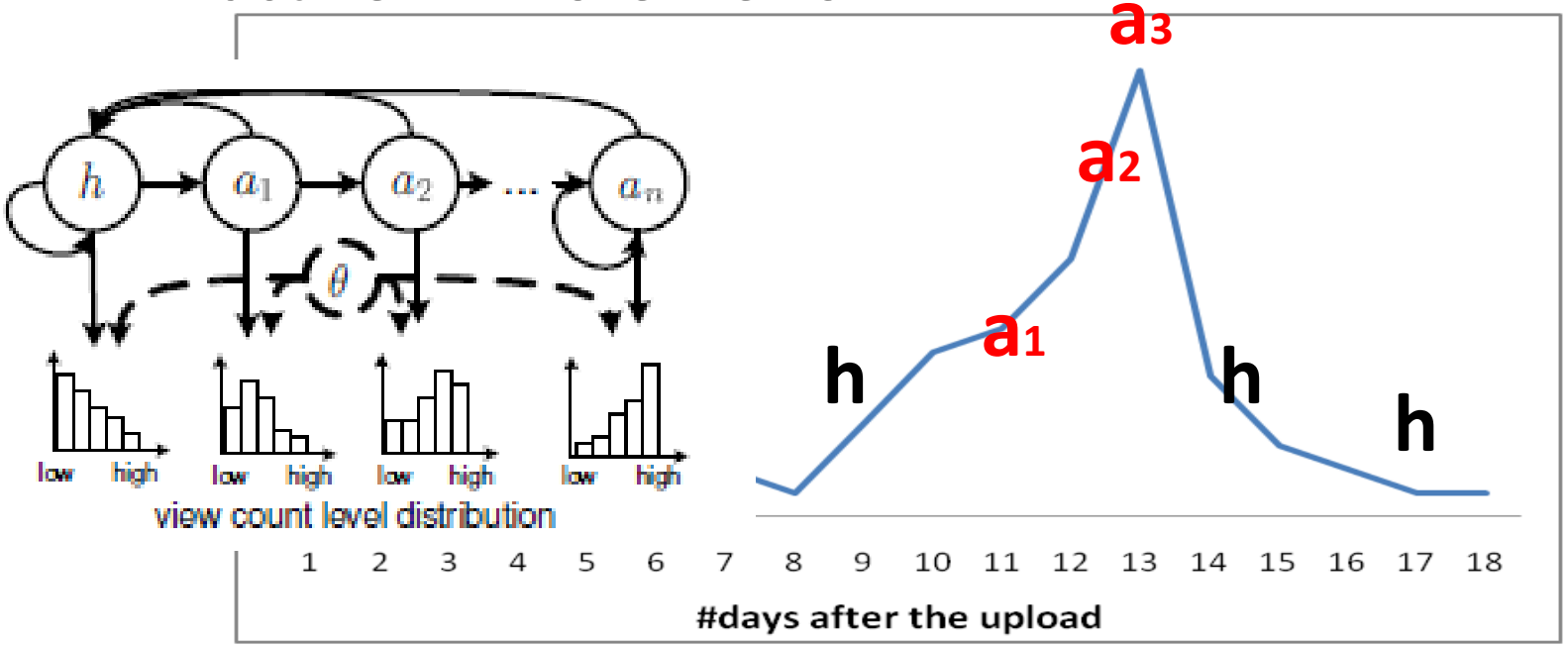
HMM Model

- Model daily views using HMM model:
- Two types of states:
  - hibernating → less views
  - active → more views



HMM Model

- Model daily views using HMM model:
- Two types of states:
  - hibernating → less views
  - active → more views

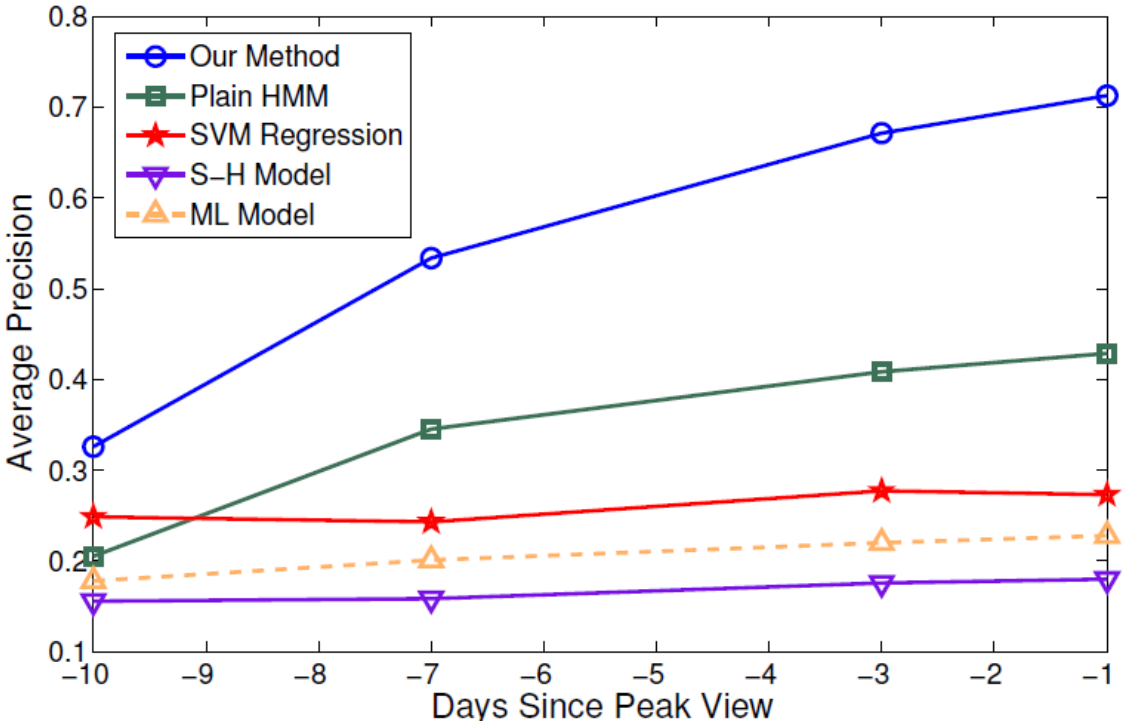


Carnegie Mellon University

HMM Model

- Model daily views using HMM model:
- Two types of states:
  - hibernating → less views
  - active → more views
- Novel modifications:
  - Incorporate metadata in the prediction. Other work only use the pure view count [Pinto et al. 2013].
  - Smooth transition probability by a Gaussian prior.

Experimental Results



Considering **metadata** in peak day prediction is instrumental.

Result cont.

- Early warning system for viral videos.
- Detect viral videos and forecast their peak dates.

Reference Date	Viral Videos: Golden Eagle Snatches Kid	Evolution of Dance	The Sneezing Baby Panda	Friday - Rebecca Black	HOW TO PLAY: P!nk
7 days before the true peak day	peak in 7 days	peak in 9 days	peak in 6 days	peak in 9 days	**Not** a viral video
3 days before the true peak day	peak in 3 days	peak in 3 days	peak in 2 days	peak in 6 days	**Not** a viral video
1 day before the true peak day	It will peak tomorrow.	It will peak tomorrow.	It will peak tomorrow.	peak in 2 days	**Not** a viral video

Carnegie Mellon University

Outline

- Introduction
- CMU Viral Video Dataset
- Statistical Characteristics
- Peak Day Prediction
- Conclusions

Carnegie Mellon University

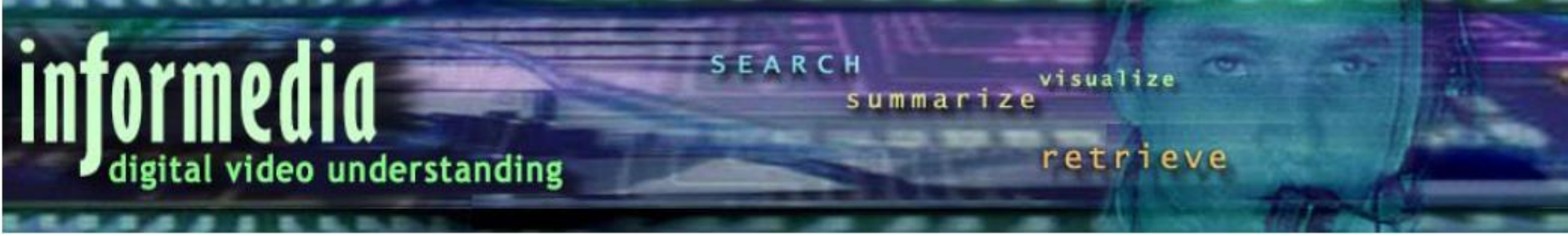
Conclusions

- A few messages to **take away from this talk**:
  - CMU Viral Video Dataset is by far the largest open dataset for viral videos study.
  - This paper discovers several interesting characteristics about viral videos.
  - This paper proposes a novel method to forecast the peak day for viral videos. The preliminary results look promising.

Carnegie Mellon University

References

- T. West. Going viral: Factors that lead videos to become internet phenomena. Elon Journal of Undergraduate Research, pages 76–84, 2011.
- T. Broxton, Y. Interian, J. Vaver, and M. Wattenhofer. Catching a viral video. Journal of Intelligent Information Systems, 40(2):241–259, 2013.
- Y. Borghol, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti. The untold story of the clones: content-agnostic factors that impact youtube video popularity. In SIGKDD, pages 1186–1194, 2012.
- H. Pinto, J. M. Almeida, and M. A. Gon¸calves. Using early view patterns to predict the popularity of youtube videos. In WSDM, pages 365–374, 2013.


informedia
digital video understanding
SEARCH
summarize
visualize
retrieve

THANK YOU.
Q&A?