

Socially Motivated Multimedia Topic Timeline Summarization

Mathilde Sahuguet
EURECOM
Sophia Antipolis, France
mathilde.sahuguet@eurecom.fr

Benoit Huet
EURECOM
Sophia Antipolis, France
benoit.huet@eurecom.fr

ABSTRACT

As the amount of social media shared on the Internet grows increasingly, it becomes possible to explore a topic with a novel, people based viewpoint. Contrasting with traditional man-made topic summarization which provide the personal view of its author, we want to focus on public reaction to events. To this end, we propose an approach to automatically generate a timeline of popular events related to a given topic. Time segments of interest are extracted from Google Trends results using a simple statistical approach. Each event, relevant to the specified topic, is illustrated on a timeline by videos mined from social media sharing platforms that gives context to the events and offers an overview of what has caught people's attention. We report the results provided by our approach for automatically illustrating the popular moments of four celebrities.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Retrieval models

Keywords

Video-based highlights, Timeline, Search Trends, Web Mining, Video Search, Social Media, Summarization

1. INTRODUCTION

Every day, millions of new documents are published on the Internet. This amounts to a huge mass of available information; it is not straightforward to retrieve useful content or to have an overview of a topic. The data is out there, but a question still remains: how to make sense of it? This highlights a need for summarization techniques that enable an automatic collection of the most relevant and informative data, in order to present the topic as a whole to the user in a way (s)he will easily understand.

In parallel, millions of search queries are issued each day over the Internet. Search logs contain a lot of information.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SAM'13, October 21, 2013, Barcelona, Spain.

Copyright 2013 ACM 978-1-4503-2394-9/13/10 ...\$15.00.

<http://dx.doi.org/10.1145/2509916.2509925>.

In particular, it is possible to exploit search behavior to discover hot topics: what people are searching for often reflects a general context. Indeed, popular term queries can be shown in the light of the current happenings and events. Google Flu Trends¹ builds accurate predictions of influenza activity based on certain search term popularity in diverse geographical areas. Hence, search behavior gives an insight in topic-related events. Similarly, uploading behavior reveals current interests and shows the actuality as seen by people. They share what they deem interesting in a specific context, so web videos of hot topics are massively uploaded to the Internet. On major storing platforms such as YouTube, the set of daily uploaded videos represents a collective repository of knowledge regarding the current context.

We want to leverage from both social media sharing and search trends as a source of knowledge to identify important events. We define an event as an occurrence of abnormal activity on a limited time segment, that captured a lot of interest and triggered massive web search. For a celebrity, an event could be a public event (concert), a personal event (wedding) or even a viral video. We aim at building a time oriented visual summary of events, using videos to illustrate events along a timeline. Indeed, videos capture information in a rich and effective manner, allowing viewers to quickly grasp the whole semantic content with limited effort. The popularity of social media sharing platforms provides access to a massive amount of multimedia documents of varying genre and quality. Timelines enable to represent information on a linear axis, which makes it easier to follow the evolution of an event or to distinguish between different events while providing a global view. We aim to exploit wisdom of the crowd by querying Google Trends² and YouTube, by mining retrieved data and combining them to discover events.

In this paper, we address the problem of automatic timeline generation by mining search behavior and video information. We use Google Trends data to extract time segments which captured attention, and perform a focused query on social media platforms to retrieve a set of candidate videos that we further process to ensure relevance. We test our framework on the creation of four celebrities summaries.

2. RELATED WORK

Organizing a news summary in a structured fashion has been researched over the past years. A structured output enables the visual aspect of the summary to be an integral

¹<http://www.google.org/flutrends/>

²<http://www.google.com/trends/>

part of the understanding process: timeline summarization displays events along a linear axis that is a key in interpreting the chain of events. While metro maps [6] enable to link different stories together, here we focus on summarization of a single story or chain of events.

The term "timeline generation" includes two different tasks regarding topic summarization. The first task is to illustrate the evolution of a topic over time and capture the diverse relative events, from a query-oriented collection of documents. The second one is that of extracting topics and spot their evolution from a pool of documents of all sort (e.g., a collection of news articles from a defined period of time).

[9] studies the construction of multiepisode video summaries in the form of a table of images. The authors of [3] propose to generate a timeline summary of a given topic using sentences extracted from a collection of documents. In [10], the authors extend the timeline representation to images, in order to build a visual summary of events along time. Textual content and images are jointly ranked. A main difference with our work is that the document collection used is a static dataset of archive articles. Events and their dates are extracted from the collection with no prior focus on a particular time segment, while we incorporate mining from the crowd to our framework.

On the other hand, research on topic detection can also make use of timeline analysis. Chen et al.[1] use the time dimension to extract hot topics from a collection of documents. They aim at summarizing what has happened during a period of time based on words and sentences of the articles.

Christiansen [4] takes a different approach to model the evolution of a topic over time. The paper analyses the frequency of occurrence over time of a search term in Google Trends, and of the same term as tag in Diigo (a social bookmarking website). Time series analysis of the term frequency lead them to segment topic behavior in different time intervals and to the definition of "topic signatures".

In [2], the authors describe their framework, based on different cues (tags, keyframes and hot search queries) to enhance topic detection in videos. Contrarily to our work, search queries are used as a cue to refine the detected topics, whereas we use them as an input for event detection.

The closest work to our is that of [8]. They extract hot times related to a query from Google context information (including Google Trends) by comparing term frequency data to a fixed threshold. Then, they pair news articles and videos from events on those extracted time segments. Our work proposes a new methodology to mine hot times, using an adaptive burst detection technique. Our approach to summarization is also different in the sense that it is biased towards user generated search terms and video uploading behavior. Hence, we focus on events as conveyed by people's minds, contrarily to news created by journalists.

3. FRAMEWORK

Our framework (Figure 1) is composed of the following steps: first, we query Google Trends with the given query term in order to have an overview of its popularity through time. We identify time segments of interest and then query social media platforms on those segments in order to get a pool of videos for each segment. Illustrating events on the timeline implies making a choice on which videos to display. Sections 3.2.1 and 3.2.2 describe the two approaches devel-

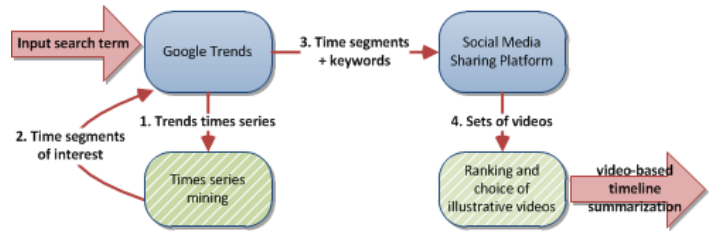


Figure 1: Proposed framework

oped to solve this issue. The first video of each set will be the representative video in the timeline summarization.

3.1 Time segment extraction

The first challenge to our approach is to define time segments corresponding to times when people's interest was high with respect to the topic of interest.

Google Trends enables to retrieve the time series representing the popularity of search query terms on a weekly basis. This is why the week is the base time unit in our work. Given a term, it returns the time series representing the likelihood of a random user to search for this term. This data has been normalized and scaled on a 0 to 100 axis. In this paper, we don't make use of geographical information and we will refer to the trends values as *popularity values*.

We assume that trends reveal two kinds of time segments, and so divide related extracted events in two types:

- *Bursts* show a sudden increase in the number of search, which indicates that an unexpected event happened and gave rise to a massive search behavior. Thus, bursts are time segments restricted to a week, which is the smallest time unit considered. For each week, we compute the *burst value* as the difference between its popularity value and that of the previous week; a week is a burst week when its burst value is positive (i.e., the given keyword draws more queries during this week than during the previous week). We order them by decreasing burst values. They constitute the main focus of this work. For example, the query 'Mickael Jackson' reveals a sudden increase in popularity between the 21st and 27th of June 2009, which can be correlated with the announcement of his death on the 26th.
- *Long-term events* are those events that lasts over several time units. Those events give rise to an abnormal interest; the values of the query for such an event stays high (compared to the average value) over several time units. Thus, time-segments representing long-term events spread over several weeks. For example, the popularity values for the query "Obama" shows abnormal high values between August and December 2008, during the time of the American presidential election. This aspect will be studied in future work.

Therefore, we obtain an ordered list of burst weeks. Behind this term lie both bursts that reveal the happening of an event, and those that only reflect a slow variation in popularity due to the general context. Hence, we need a threshold to separate them given their burst value. We designed an adaptive threshold that takes into account the distribution of the bursts. For each list of burst week, we compute the

mean value and standard deviation, and define the threshold as the sum of those values. Value above the threshold are those which considerably differ from the mean.

The output of this operation is a set of week dates that we want to link to some event in the real world. Event discovery will be made after querying a video sharing platform, but we want to orient our queries towards what people were searching during those times. This is why we perform an additional query to Google Trends: for each burst, we perform a query focused on the month (the finest grain for Google Trends queries) of the event. It returns a list of rising search terms associated with this query. At the end of the process, we obtain a set of time segments and associated terms that reveal the motivation underlying the queries. The next step is then to query online social sharing platforms in order to give context to those events.

3.2 Video focused search

We query the YouTube API on the relevant time intervals and their associated terms. For each time segment, we obtain a set of videos that were uploaded during the queried week and are supposed to be related to the event at stake. Users of such storing platforms are aware that in some cases, retrieved documents may not fit the query perfectly. Therefore it is necessary to prioritize the most representative videos, as a base for the choice of the video that will illustrate the timeline.

We perform those actions based on the semantics of the user-generated text that surrounds each video (title and description). First, we discard none English-language content using [7], so it is possible to compare textual features on their semantic meaning. This process may lead to losing some interesting content; nevertheless, we argue that this also lead to remove some unrelated or low quality content: some descriptions are just a combination of keywords that make no sense but that have been put together to raise a lot of views; also, poor English description may be correlated with poor content of the video. We extract textual features in order to use natural language processing techniques for the analysis. We use the vector space model with TF-IDF weighting and cosine similarity distance to represent videos.

Next, we order the datasets; the first item of each set is chosen as the representative video that appears in the timeline.

3.2.1 Baseline ranking by simple distance averaging

A first baseline approach is to compute a ranking of the videos using the average similarity to all other videos. Hence, the medoid is chosen as the most representative of the dataset.

3.2.2 DBSCAN clustering

In a second time we used DBSCAN [5] clustering to order the video data. DBSCAN is a density based clustering algorithm which also enables to remove outliers in the dataset. As we assume that the dataset is composed of videos with similar descriptions and a minority of noisy data with low similarity to other items, we argue that density can be a way to separate those items. DBSCAN clustering was also chosen for its ability to automatically generate the number of clusters and to remove noise.

Parameters have been set up empirically following [5] so they generate a low number of clusters and remove about 25% of noise on average on the tested data.

	# burst	TP	FP	FN	DE	ER1	ER2
O.P.	1	1	0	8	0	0	0
B.K.	9	7	1	21	1	1	1
						0.5	0.5
						1	0.5
						0.5	0.5
						1	1
						0.5	0.5
M.Z.	6	2	4	25	0	1	0
B	6	2	2	7	2	1	1
						1	1
						1	1
						1	0.5

Table 1: We report the number of bursts along with the number of true positives (TP), false positives (FP), false negatives (FN) and discovered events (DE) for each topic. Results of events relatedness for both approaches (ER1 and ER2) are then listed

This step generates several curated clusters (in practice, it is generally 1 or 2); again, we take the medoid of each set as representative data and rank the items by average similarity of each cluster.

4. EXPERIMENTS

Our goal is to summarize what captured people’s attention regarding a certain topic. We will focus on the person scenario, where we aim to generate a biography of popular moments of a celebrity, but this work could apply to many different concepts. For evaluation, we built timelines for the following persons: Oscar Pistorius (O.P), Beyonce Knowles (B.K), Mark Zuckerberg (M.Z.) and Batman (B.). The timeline is drawn from January 2004 (start date for Google Trends data) to present.

4.1 Popular event extraction

First, we evaluate our event extraction framework. For each query, we want to compare the extracted time segment or burst weeks to a manually created ground truth. This ground truth was constructed based on expert biographies³ and Wikipedia data, although for the "Batman" query the motivation was different: as it is a fictional character, we created ground truth by listing movies and video game releases that are the most generally popular associated events.

For each person, we manually compare top words from the extracted videos on the given time segments with description of the events in our ground truth to reveal matches or misses. Table 1 displays the results in term of : true positive events (TP), false positive events (FP), false negative events (FN) and discovered events (DE) which are events not described in the ground truth but we could find trace of on the Web.

As the timeline is based on popular moments which do not exactly match official biographies, evaluation of such results is complicated. On the one hand, it does not return all highlights of a biography, but only unforeseen events that

³<http://www.biography.com/>

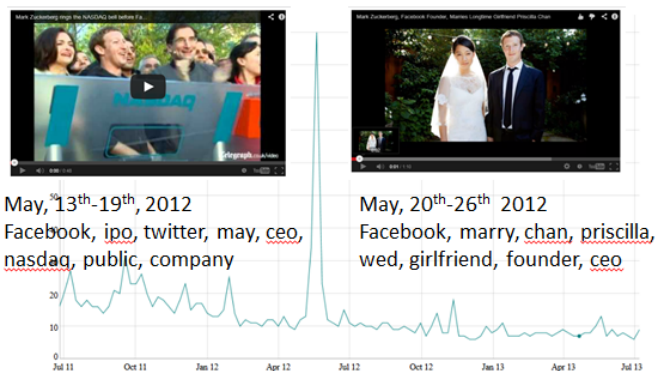


Figure 2: Extracted events for Mark Zuckerberg (M.Z.) from July 2011 to July 2013

caught public attention. On the other hand, it may reveal events that are not part of a classic biography but that could be linked to actual events that were discussed a lot. For example, Beyonce falling during a live show in Orlando was not part of any descriptive biography, but we could discover this happening with our system.

Also, true negative is hard to assess : how can one classify an event as worth appearing on the timeline ? If all happenings of a lifetime are displayed, we are losing the point in the summarization. We generated the ground truth by exhaustively taking every date and event mentioned in the expert biography and Wikipedia page, with no consideration of the importance of the event. Hence, false negatives are not very representative of the capacity of the algorithm to capture "important" moments. For example, Mark Zuckerberg became a public figure around 2007 and the events captured by our framework are no earlier than 2010 (see figure 2 for an extract of timeline summarization).

Also, a dissimilarity of granularity between our framework (week unit) and Google Trends (month unit) made it hard to extract focused search term when several events happened during the same month. While our algorithm has selected the week from the 9th to 15th of January 2011 as a peak week, the top words did not reveal a unified event; nevertheless external knowledge lead us to correlate the peak in the search to rumors of Facebook shutting down.

4.2 Video summarization

The second part of the evaluation relates to the choice of illustrative video. The baseline illustrates each event with one video while the second approach returns one or more videos depending on the number of clusters found. This evaluation focuses on time segments granted as true positives (TP) or discovered events (DE).

We evaluate if the video is related to the event at stake on a scale from 0 to 1 : for each video, the value of *event relatedness* is set to 0 if the video has no link with the event, 0.5 if it partly matches, or illustrates the event among other things, and 1 if it is related to the event. If we have several illustrative videos (second experiment), the average is given. Note that the quality of the information is not evaluated. The results reported in the second part of table 1 (ER1 for the first approach, ER2 for the second) show that both approaches perform reasonably well on the majority of

events. Some low scores can be explained by the fact that we rely on textual content to represent videos, although it may differ from actual content. Performing content-based analysis will be the subject of future work. Also, it can be noted that DBSCAN clustering does not perform as well as the baseline given this metric. In fact, creating clusters enables to separate different events that are reported during the same week; only one of them may match the event that is chosen as the most representative of the burst. Further work can lead to choose the cluster of interest among them.

5. CONCLUSION

In this paper, we tackle the issue of topic summarization from people's point of view, making use of social media data and search trends. Hence, we focus on events that were revealed by their existence on both people's interest and a social media platform such as YouTube. We design a framework to automatically build timelines of events regarding a topic, focusing on celebrities as they raise a lot of interest. We are able to discover both events that are part of general highlights, summarized on experts biographies, but also to point out events that were forgotten, but which at the time created much interest among people.

Our approach is based on textual features which are user generated; in order to have an insight on the actual video content, future work will perform video content analysis based on visual and audio information. We will also attempt to discover long-term events whose atomic unit will be more than a week by time-series mining. Last, using geographical information can be an useful cue for event summarization.

6. ACKNOWLEDGMENTS

This work was supported by the European Commission under contracts FP7-287911 LinkedTV and FP7-318101 MediaMixer.

7. REFERENCES

- [1] K.-Y. Chen, L. Luesukprasert, and S.-C. Chou. Hot Topic Extraction Based on Timeline Analysis and Multidimensional Sentence Modeling. *Knowledge and Data Engineering, IEEE Trans. on*, 19(8):1016–1025, 2007.
- [2] T. Chen, C. Liu, and Q. Huang. An effective multi-clue fusion approach for web video topic detection. *MM '12*, pages 781–784, New York, NY, USA, 2012. ACM.
- [3] H. L. Chieu and Y. K. Lee. Query based event extraction along a timeline. In *ACM SIGIR*, pages 425–432, New York, NY, USA, 2004. ACM.
- [4] L. Christiansen, T. Schimoler, R. Burke, and B. Mobasher. Modeling topic trends on the social web using temporal signatures. *WIDM '12*, pages 3–10, New York, NY, USA, 2012. ACM.
- [5] M. Ester, H. Peter Kriegel, J. S., and X. Xu. A density based algorithm for discovering clusters in large spatial databases with noise. pages 226–231. AAAI Press, 1996.
- [6] D. Shahaf, C. Guestrin, and E. Horvitz. Trains of thought: generating information maps. *WWW '12*, pages 899–908, New York, NY, USA, 2012. ACM.
- [7] N. Shuyo. Language Detection Library for Java, 2010.

- [8] S. Tan, H.-K. Tan, and C.-W. Ngo. Topical summarization of web videos by visual-text time-dependent alignment. In *ACM Multimedia*, pages 1095–1098, 2010.
- [9] I. Yahiaoui, B. Merialdo, and B. Huet. Comparison of multi-episode video summarization algorithms. *EURASIP Journal on applied signal processing*, 01 2003.
- [10] R. Yan, X. Wan, M. Lapata, W. X. Zhao, P.-J. Cheng, and X. Li. Visualizing timelines: evolutionary summarization via iterative reinforcement between text and image streams. In *CIKM*, pages 275–284, 2012.