# Understanding Real-World Events via Multimedia Summaries Based on Social Indicators

Mauricio Quezada and Barbara Poblete

PRISMA Research Group
Department of Computer Science, University of Chile, Chile
{mquezada,bpoblete}@dcc.uchile.cl

**Abstract.** We present a novel methodology for creating multimedia summaries of real-world events through social media information. Summaries are generated using selected multimedia data disseminated through Twitter. The proposed summarization technique takes into account social indicators of relevance, which are used to select a set of representative multimedia objects for summarizing the event from a social perspective. In addition, our approach incorporates different news angles by extracting topics within each event.

**Keywords:** Social Networks, Collaborative Summaries, Multimedia.

## 1 Introduction

The Web and in particular social networks are characterized for generating and publishing huge amounts of information on a daily basis. In particular, important real-world events (such as recent Boston bombing attacks, for example) overflow social media platforms with millions of messages. Moreover, many of these user-generated messages not only include textual data, but also hyperlinks to external media documents, images, and videos. It is understandable that under this type of scenario the volume of data becomes overwhelming for any human to analyze. Therefore, users searching for fresh information about news events must settle with just browsing a few messages or waiting for traditional news sources to report the information.

One of the most challenging social media platforms in terms of data volume and user adoption, is Twitter[1]. With over 400 million messages (tweets) posted every day[2], its users regularly use Twitter to share all sorts of multimedia. Mostly, multimedia content embedded (as hyperlinks) in tweets consists of text documents (on-line news articles, blog posts, Web pages), photographs (from services like Instagram[3], Flickr[4]), videos (from YouTube[5], Vimeo[6]), or even audio (SoundCloud[7] for instance). Depending on

[1] http://www.twitter.com/
[2] https://blog.twitter.com/2013/celebrating-twitter7
[3] http://www.instagr.am/
[4] http://www.flickr.com/
[5] http://www.youtube.com/
[6] http://www.vimeo.com/
[7] http://www.soundcloud.com/

P. Antunes et al. (Eds.): CRIWG 2013, LNCS 8224, pp. 18–25, 2013.

the nature of the real-world event, the most significant data can be in the form of text, images or video.

To address the multimedia data deluge on the Web, and in particular that of online social media, we propose a methodology for automatically summarizing events. This reduces the load on the user for the task of understanding events. Our approach is based on the incorporation of multimedia data into our summaries as well as text. To achieve this we propose a social based selection scheme for multimedia objects (such as video, images, documents and tweets themselves) that compose the information of an event. The main advantage of our method is the use of social information extracted from Twitter, as a key factor of multimedia document importance. This is, for any event disseminated through the social network, we consider as its most important elements those that have generated the most interest from users (e.g., shares and/or replies). Then, using this selected media we build a comprehensive summary of the event.

Our approach is unique, in the sense that unlike other automatic summarization techniques, we do not create textual excerpts from existing documents. Instead, we create our summary as a composition of the most representative elements of each subtopic of the event from a social perspective. This allows us to generate summaries regardless of the type of the element in the summary, facilitating multimodal element combination.

By taking advantage of all of the information people post in the social Web, we see this approach as a form of implicit collaboration. Since every message conforms the description of an event, summarizing this information helps in understanding particular aspects of it. For example, this type of approach can help journalistic inquiry or the evaluation of sociological hypotheses.

In this article we present our system prototype, which uses Last.fm[8] and Google News[9] as event aggregators for concerts and world news, respectively. For each of the identified events our system gathers related social media information using Twitter. We further expand this dataset by including all of the multimedia elements embedded in each tweet. We model each of these multimodal Web elements using a vector representation. This representation is generated aggregating the text from the tweets that reference the multimedia object, identified by its URL. In addition, we clusterize these elements, to identify subtopics within the event. Then, from each subtopic we extract the most important documents according to social information. Even though this is a first implementation of our approach, our use cases show that it can be a very effective and useful summarization technique. We show promising preliminary results from manually inspected case studies that support the soundness of our approach.

This paper is organized as follows: in Section 2 we present relevant related work; in Section 3 we give an overview of our prototype system and methodology; in Section 4 we show some preliminary results, and in Section 5 we discuss conclusions and future work.

## 2 Related Work

Relevant literature for our research topic can be classified as follows: automatic generation of summaries based on real-world events and social data, content selection for event representations and measuring relevance from social messages.

[8] `http://last.fm`
[9] `http://news.google.com`

There are several works on the topic of content summarization using social information, being that of Chakrabarti and Punera [8] one of the most representatives. In this work the authors summarize sport events based on the frequency of tweets in certain periods of time, whose duration is determined automatically. They then summarize each time window using a similarity approach. Summaries are created based on textual information in tweets, from which they select those with higher scores.

By using multimedia data, Del Fabro and Böszörmenyi [2,3] generate image and video summaries from well-known events, such as a royal wedding. The relevance of each video is measured by its popularity on the social video platform YouTube. These summaries are only generated for selected events and do not consider other types of information (just videos). Another approach is that of Sinha and Jain [4], in which they make summaries from personal photography collections. They use content-based image features to create clusters from which they produce the summary.

Our contribution in this area is that of creating summaries by aggregating information from social networks using data mining techniques. Our summaries are composed of multimodal data, which includes video, audio, images, documents and in some cases social media messages. Our work is related to that of Becker et al. [5,6], in the sense that multimedia information is used for summarization. But it differs in the fact that we base our event representation solely on social data, as well as element relevance. We select social features for event representation using the work of Castillo et al. [7] and Duan et al. [9].

## 3 Methodology Overview

We propose a methodology for generating real-world event summaries using a unified representation of multimodal documents and their social relevance. The relevance of each document is measured using social indicators that mined from on-line social networks. This is, the more shared or commented document is, the more *relevant* it is considered compared to the rest. The methodology involves several stages, from data extraction for each event, to generating a summary by gathering the most relevant documents from each subtopic of an event representation. In detail, the steps that compose our methodology are the following (shown in Fig. 1):

1. Event metadata extraction from event aggregators, like news or music events
2. For each event: Document extraction by searching on-line social media sources using the event metadata
3. Event modeling phase, i.e., for each event, aggregate and adapt the data for a proper representation.
4. Event subtopic identification (for each event)
5. Event multimodal summary creation, selecting the most relevant documents from each subtopic.

**Event Metadata Extraction**
In order to obtain descriptive event metadata, we use event aggregators. These aggregators usually have a list of events and their descriptions such as, a title, start and end date of the event, and a short list of keywords or a short description of the event.

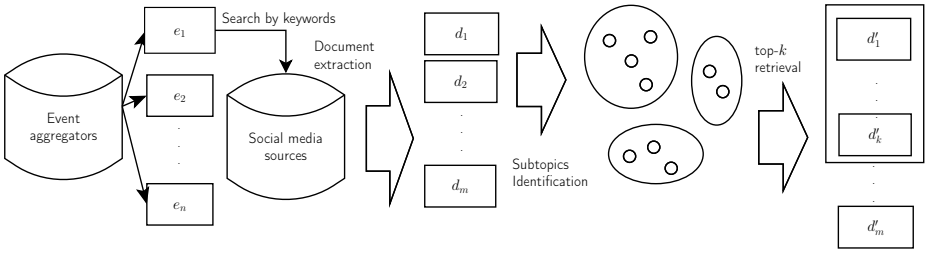


**Fig. 1.** Methodology overview. The steps of the methodology consist in an event extraction phase, social media search for each event, then subtopic identification and finally the selection of the top-$k$ most relevant documents from each subtopic.

For our initial implementation we only considered two types of events: music concerts and world news. For music events (concerts and festivals) we used the Last.fm API for geolocated events in London, Glastonbury, Las Vegas, Stockholm and Santiago. For news event information in Chile and the US we used the Google News API. Using these public APIs we extract relevant data, such as names of artists participating in music events, and the related news titles. We use this data as keywords for each event for the next phase.

**Event Data Extraction and Representation Using Social Sources**
In order to model multimodal documents that are part of a particular event we use only social data. Therefore, for each event, we perform a social media platform (in this case Twitter) search using the metadata extracted in the previous phase. This search is time-sensitive, hence we search the time-frame in which the event occurs (start and end date of the event).

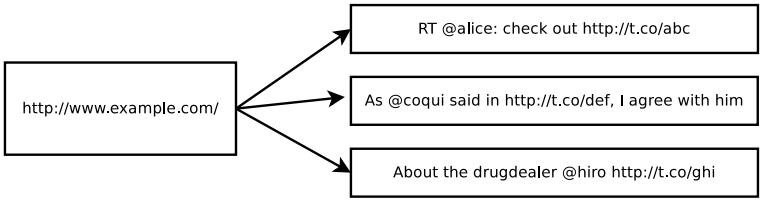


**Fig. 2.** Example of a document representation. After resolving all shortened hyperlinks contained in the messages, we group every tweet with the same resolved URL into a document, with that URL as its identifier. In this example we do not show the tf-idf scaling of the texts for clarity.

Once we obtain all the messages or tweets that discuss the event, we separate those messages which contain hyperlinks from those that do not. Messages which do not include a hyperlink are added to our multimodal document collection, considering the sole tweet text as the content of the document. On the other hand, if a message contains a hyperlink, then this message is not added to our document representation. Instead, we extract all of the hyperlinks mentioned in messages. Next, we resolve each URL, which in many cases has been shortened, to find duplicate links. Each of these unique hyperlinks may reference several multimodal documents, which can be textual documents,

Web pages, images, videos, etc. We add each of these hyperlinks, identified by their URL and use the aggregation all of the tweets which mention this URL as surrogate textual content for the document. Therefore, our implementation does not require that we download or process the actual contents of the multimodal document referenced by the URL. Instead, we use the aggregated short messages that mention the URL as text content for the document. This way, we tackle the issue of processing very short and noisy text messages, like tweets. Figure 2 shows an example of the previously described document representation.

Next, after all of the documents of an event are added to the event collection we apply standard *tf-idf* scaling to their vector-model representation (see [1]).

**Subtopic Identification**
Once we have the set of multimedia document vectors that compose an event in our system, the next step is to identify subtopics. As a first approach, we use the *K-means* clustering algorithm for this purpose.

K-means requires as an input the number of desired subtopics. As this is not simple to estimate, we approximate this by using the number of related news items and the number of artists, for news and concert events, respectively. It should be noted that it is possible to improve the subtopic estimation of cluster numbers by running K-means multiple times with different number of clusters. This parameter was determined empirically for in the use cases we describe in the following section.

**Selection of Relevant Documents**
The final step is the selection of the most relevant documents for each event subtopic. In this initial implementation we select the top-2 or top-3 documents of each subtopic. To select the documents, we use a simple method approach consisting of determining relevance by weighting several social features. As a first approximation we use the number of retweets, the number of times the tweets was marked as favorite, the number of followers/followees of the author, the number of lists that the author belongs to. This weighting scheme was built following the conclusions of Duang et al. [9] work, and most of the indicators detailed in the work of Castillo et al. [7].

Therefore, the summary of each event was created by using the top-2 or top-3 elements of each event subtopic, for the purpose of generating a succint summary of each event.

## 4 Case Studies

In order to assess the soundness of our methodology we inspected two different types of events:

1. `Police arrest suspects in Tel Aviv` (News, 250 hyperlinks, 743 tweets)
2. `New York Philharmonic Dvorak's New World Symphony` (Music concert, 150 hyperlinks, 279 tweets)

We obtain preliminary results for each event, first by determining empirically the appropriate number of clusters. For this we used the inter-cluster and intra-cluster ratio measures.

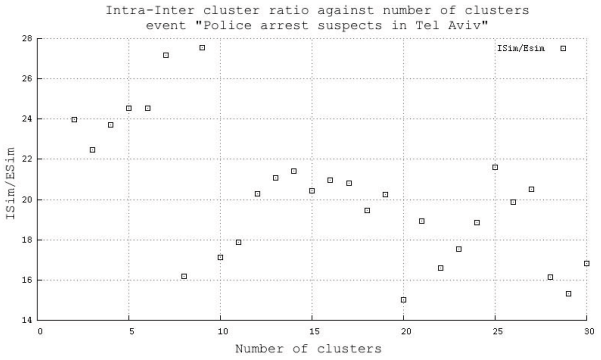


**Fig. 3.** Evaluation of the clustering solution for the event "Police arrest suspects in Tel Aviv". *ISim* represents intra-cluster measure and *ESim* the inter-cluster one. It can be seen that there is a local maximum at 9 clusters.

For our implementation we used `Cluto` (see [10]) to generate a clustering solution with a fixed number of clusters. This program returns a measure of the intra-cluster and inter-cluster evaluation of the clustering. The inter-cluster indicates how similar the elements are of one cluster to the elements of the remainder of them. The intra-cluster indicates how similar the elements of one cluster are to every other element in the same cluster. We divided the intra-cluster measure by the inter-cluster one, by considering that the higher the ratio, the better the clustering.

The Fig. 3 shows the ratio against various clustering runs for the Tel Aviv event. It can be seen that 9 is a proper number of clusters because the Intra-Inter cluster similarity ratio is a local maximum. Figure 4 shows the most relevant documents according to a methodology applied to this event. The summary is mostly composed of Web documents and twitter status messages. The following links constitute the identifiers of the summary documents of the event:

1. *Arrest announced in Tel Aviv bus bombing — National News - WDSU Home*[10]
2. *Shin Bet, police arrest suspects in TA bus bombing — JPost — Israel News*[11]
3. *Tel Aviv Bombing Suspects Arrested - The Daily Beast*[12]
4. *Twitter / panosharitos: Tel Aviv police chasing after ...*[13]

[10] `http://www.wdsu.com/news/national/Arrest-announced-in-Tel-Aviv-bus-bombing/-/9853500/17524408/-/m562vn/-/index.html`
[11] `http://www.jpost.com/Defense/Article.aspx?ID=293140&R=R1`
[12] `http://www.thedailybeast.com/cheats/2012/11/22/tel-aviv-bombing-suspects-arrested.html`
[13] `https://twitter.com/panosharitos/status/271204357654077441`

5. *Twitter / 1stNewsHeds: New York (NY) Times: Police ...*[14]
6. *Twitter / MARKETRISER: Israel arrests suspects in ...*[15]
7. *Previous bomb attacks in Tel Aviv - Yahoo! News*[16]
8. *Twitter / BreakingNews: Israel's army spokesman says ...*[17]
9. *Arrest announced in Tel Aviv bus bombing - CNN.com*[18]


**Fig. 4.** Twitter message with the highest score in the "Police arrest suspects in Tel Aviv" event summary. (Source: Twitter)

For the "New York Philharmonic Dvorak's New World Symphony" event, most of the summary documents are non-textual elements (multimedia). These elements are referenced from Instagram, Twitter (with an embedded photo) or YouTube, as can be seen in Fig. 5.


**Fig. 5.** Some of the resulting documents from the "New York Philharmonic Dvorak's New World Symphony" summary. (Sources: Instagram, Twitter and YouTube, from left to right)

[14] https://twitter.com/1stNewsHeds/status/271752495401934849
[15] https://twitter.com/MARKETRISER/status/271758949043298304
[16] http://news.yahoo.com/previous-bomb-attacks-tel-aviv-153452046.html
[17] https://twitter.com/BreakingNews/status/271718080856592384
[18] http://edition.cnn.com/2012/11/22/world/meast/israel-bus-bombing-arrests/index.html

## 5 Conclusions and Future Work

We presented a novel methodology for generating automatic summaries from multimedia real-world event information. The main contribution of our work is to exploit social information from Twitter to summarize in a simple way events which are extremely rich in information. In addition, by using social information we are able to combine and select multimodal elements for our summaries. Overall, our preliminary inspection of results shown by this approach are very promising and the resulting summaries give comprehensive descriptions of the events. As part of future work we expect to improve subtopic identification algorithms and document selection and ranking. In addition, as this is a work in progress, we are working on a large scale evaluation which will compare with alternative approaches and incorporate other types of events.

**Acknowledgements.** Mauricio Quezada was supported by CONICYT grant CONICYT-PCHA/Magister Nacional/2013 - 22131151 and FONDECYT grant 11121511.

Barbara Poblete was partially supported by FONDECYT grant 11121511 and Program U-INICIA VID 2012, grant U-INICIA 3/0612; University of Chile.

## References

1. Salton, G., Wong, A., Yang, C.-S.: A vector space model for automatic indexing. Communications of the ACM 18(11), 613–620 (1975)
2. Del Fabro, M., Sobe, A., Böszörmenyi, L.: Summarization of real-life events based on community-contributed content. In: The Fourth International Conferences on Advances in Multimedia, MMEDIA 2012 (April 2012)
3. Del Fabro, M., Böszörmenyi, L.: Summarization and presentation of real-life events using community-contributed content. In: Schoeffmann, K., Merialdo, B., Hauptmann, A.G., Ngo, C.-W., Andreopoulos, Y., Breiteneder, C. (eds.) MMM 2012. LNCS, vol. 7131, pp. 630–632. Springer, Heidelberg (2012)
4. Sinha, P., Mehrotra, S., Jain, R.: Summarization of personal photologs using multidimensional content and context. In: Proceedings of the 1st ACM International Conference on Multimedia Retrieval (April 2011)
5. Becker, H., Iter, D., Naaman, M., Gravano, L.: Identifying content for planned events across social media sites. In: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining (February 2012)
6. Becker, H., Naaman, M., Gravano, L.: Learning similarity metrics for event identification in social media. In: Proceedings of the Third ACM International Conference on Web Search and Data Mining (February 2010)
7. Castillo, C., Mendoza, M., Poblete, B.: Information credibility on twitter. In: Proceedings of the 20th International Conference on World Wide Web (March 2011)
8. Chakrabarti, D., Punera, K.: Event summarization using tweets. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (May 2011)
9. Duan, Y., Jiang, L., Qin, T., Zhou, M., Shum, H.Y.: An empirical study on learning to rank of tweets. In: Proceedings of the 23rd International Conference on Computational Linguistics. Association for Computational Linguistics (August 2010)
10. Karypis, G.: CLUTO-a clustering toolkit (No. TR-02-017). Dept. of Computer Science. University of Minnesota (2002), `http://www.cs.umn.edu/~cluto`