Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 00 (2014) 000–000

Procedia Engineering

www.elsevier.com/locate/procedia

Humanitarian Technology: Science, Systems and Global Impact 2014, HumTech2014

Effects of disaster characteristics on Twitter event signatures

Haji Mohammad Saleem[a,*], Yishi Xu[a], Derek Ruths[a]

[a]*School of Computer Science, McGill University, Montreal, QC H3A 0G4, Canada*

Abstract

Twitter has emerged as a platform that is heavily used during disasters. Therefore, as an event unfolds, it generates varying levels of online engagement from victims as well as onlookers (both physical and virtual). Because methods for mining disaster-related content at scale must contend with the problem of filtering out vast numbers of unrelated posts, any prior knowledge about the characteristics of disaster-related content in the live Twitter feed may help improve the recovery of relevant posts. In this study, we consider the relative abundance of a disasters Twitter content over time (both relative to total event-related content and relative to the overall volume of content generated on Twitter). We refer to this time-varying abundance as the events signature. In an analysis of three different disasters, we find that event signatures are qualitatively different. These differences can be explained in terms of several characteristics of disasters: foreknowledge, duration, severity, and news media engagement.

Selection and peer-review under responsibility of the Organizing Committee of HumTech2014.

Keywords: Twitter event analysis; Crisis Informatics; Humanitarian Data Mining; Social media; Hashtags.

1. Introduction

During and in the immediate aftermath of a crisis, many people choose to use Twitter as a means of sharing and collecting information and experiences related to the event (e.g., [1,2]). As a result, many disasters now have a substantial presence on Twitter, in the form of a time-varying progression of tweets that directly or indirectly relate to the event [3]. We call the time-varying volume of disaster-related tweets the disasters Twitter *signature*. Signatures of disasters, themselves, have not been studied, though a better understanding of them could have implications for both disaster response technology as well as sociological understanding of social media engagement, particularly in times of crisis:

- While Twitter is widely regarded as a potentially valuable source of information for responders during crises, the problem of extracting relevant Twitter posts remains a difficult and largely open problem (e.g., [4–6,8–10]). One significant issue with which extraction systems must contend is the sheer volume of unrelated content in the live Twitter stream (called the firehose): even a system with high specificity will still include large numbers of unrelated tweets, simply because of the overwhelming number of true negatives in the Twitter feed. In such circumstances, a model of the time-varying volume of a disasters Twitter content — its signature — can provide the extraction system with more accurate priors on the occurrence of disaster-related content.

* Corresponding author. haji.saleem@mail.mcgill.ca.
1877-7058 © 2014 The Authors. Published by Elsevier Ltd.
Selection and peer-review under responsibility of the Organizing Committee of HumTech2014.

- From a sociological perspective, the shape of a disasters signature may convey information about the nature of the disaster itself, how different populations of Twitter users are engaging in the event, and how exogenous variables such as news media influence the visibility and perception of the event. In fact, the results of this paper speak directly to these ideas.

To our knowledge, no work to date has directly considered the question of disaster (or even more generally, event) signatures in Twitter. Thus, in this paper we provide a careful analysis of disaster signatures across three different events. Our goals are three fold. First, we formally define the notion of an event signature in Twitter. As part of this, we highlight how diurnal variance in Twitter volume itself must be corrected in order to characterize the signature of the event itself. Second, we extract the signatures for three different disasters, providing a first glimpse at the time-varying volume of Twitter content generated by disasters. Third, we show that the three disasters have qualitatively different signatures that can be explained in terms of three features of the events themselves: foreknowledge, duration, severity, and news media coverage. While compelling, we consider the proposed connections between event and signature features to be ideas to be explored and confirmed in future work.

2. Prior Work

The intention of our study was to understand the relationship between the temporal signature of a disaster and its intrinsic characteristics. While outside the scope of crisis informatics, the work most closely related to our notion of event signatures considered the more general phenomenon of temporal variation of content on Twitter [11]. That work focused on the propagation of various hashtags in the Twitter context as a time series data, without keeping in context the specific events these individual hashtags associate with. A related study considered the dynamics of hashtag emergence in terms of growth and persistence on a cumulative scale [12]. They, however, based their studies on different hashtags from the same event — the 2012 U.S Presidential elections.

Crisis informatics studies on Twitter tend to either deeply analyze a singular event or implement a uniform extraction framework on a few curated datasets. Those focused on a particular event occasionally touch on the overall structure of the Twitter feed. One study, for example, plotted the Twitter signature for the Australian Black Saturday bushfire, but only to point out the increased event traffic [14]. Similarly, Sakaki et al. used the signature to point out subevents in the case of multiple earthquakes and use exponential distribution to describe the signature in itself [3]. Doan et al. studied the 2011 Tohuku earthquake and used the signature to study public awareness and anxiety [15]. Finally Terpstra et al. studied Twitter activity due to a storm that hit a music festival in Belgium [16]. They use the signature to analyze tweet content, namely damage reporting and casualty reporting.

3. Datasets

We ran our study on three characteristically different recent disasters, helping us understand the dynamics of public response to crisis in each situation. For each, our goal was to extract and characterize key features in the events Twitter signature. We used the Twitter decahose to collect data streams, spanning the time period during which the event occurred, with a focus on the period of time during which the primary disaster hazards (e.g., hurricanes, tornados, and high waters) were present in the area.

3.1. 2012 Hurricane Sandy

One of the deadliest and most destructive hurricanes in United States history, Sandy started as a tropical storm in the Caribbean Sea on October 22, 2012 and finally made landfall as a category 2 hurricane on the New Jersey coast a week later. The storm was closely monitored throughout its course by the National Hurricane Center (under NWS, NOAA), FEMA and its federal partners, giving out constant updates and public advisories and thus providing significant foreknowledge before landfall. Further, since the Sandy caused widespread flooding, power outages, and the shutdown of city services, its effects persisted over a long period of time. Over this time, the hurricane lead to 117 lost human lives and $65 billion in structural damage [17,18].

In our analysis, we considered Hurricane Sandys landfall on the east coast of the United States as our point of incidence. We analyzed three days – Oct 28, Oct 29 and Oct 30 — to give proper attention to both pre-event and post event tweet volume. During this time frame, 123.2 million (relevant and irrelevant) tweets were collected from the unfiltered Twitter decahose, which was the dataset used in this study.

3.2. 2013 Moore, Oklahoma tornado

The Oklahoma tornado was a category EF5 tornado that struck Moore, OK and adjoining areas on May 20, 2013. People had little concrete foreknowledge of the tornado (as compared to Sandy): the National Weather Service issued a general tornado lookout at 1:10 PM that day, a severe thunderstorm warning at 2:12 PM, which was elevated to a tornado warning at 2:40 PM and finally to a tornado emergency at 3:01 PM. The tornado touched down at 3:01 PM, 16 minutes after the first siren, and reached the town of Moore 15 minutes afterwards (at around 3:16 PM). Thus, residents who were following weather updates would have had roughly 2 hours foreknowledge. Moreover, since the tornado itself was on the ground for less than an hour, the event duration was substantially less than Sandy.

Ultimately, the tornado was responsible for 27 deaths and over $1 billion in structural damage (final figure yet to be released) [18]. However, while the tornado was highly destructive, its impact was highly localized to a few towns, which means that the disaster directly affected far fewer people than Sandy. Since foreknowledge was limited to a few hours before the incident, we only considered the day of the event and the next (i.e. May 20 and May 21) in our analysis, which resulted in a dataset of approximately 98 million tweets.

3.3. 2013 Boston Marathon Bombings

During the Boston Marathon on April 15, 2013, two crudely made bombs exploded at 2:49 pm, near the finish line. The perpetrators were eventually caught 5 days later. Being a terrorist attack, there was absolutely no foreknowledge for this event. Furthermore, since the two bombs went off in a space of 13 seconds and inflicted little structural damage to surrounding buildings, the event duration in itself was very short [13]. Despite this, the attack took the lives of 3 bystanders [19]. Notably, this was immediately recognized as an act of terrorism and, as such, it received immediate and nationwide attention from news organizations.

For the purposes of this event, we concentrated on April 15 and 16, running our study on a total of 99.4 million tweets.

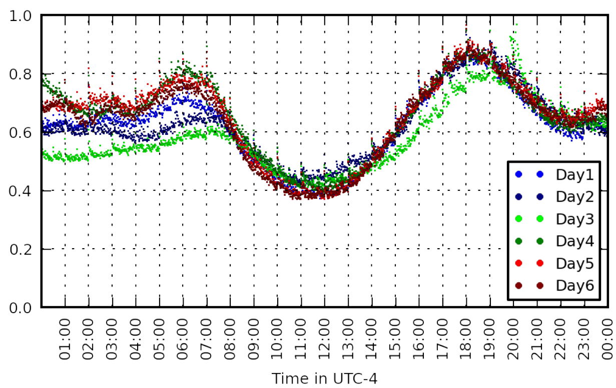

Fig. 1. Volume of tweets received during six randomly chosen days. The timespan selected was from 19th Sept 2013 to 24th Sept 2013.

4. Methods

4.1. Event signature definition

The goal in this study was to estimate and study the Twitter signatures of different disasters. Previously we defined the Twitter signature of an event as the time-varying volume of Twitter content directly or indirectly related to that event. Formally, the signature is a numeric sequence:

$$S = < s_1, s_2, , s_n >$$

where s_i is the volume of event-related Twitter content in time period i. Briefly setting aside the choice of time, consider that quantifying the volume of event-related Twitter content in a given period can be done in at least two ways: by (1) absolute volume ($s_i \in \mathbb{N}$ is the number of event-related tweets generated in time period i) or (2) relative volume ($0 \leq s_i \leq 1$, is the fraction of all tweets generated in time period i that pertain to the event) which we call *time-corrected*.

To see the implications of using absolute volume for event signatures, consider an event which occurs at 10 AM and that 100% of Twitter volume at that time focuses on the event. If the proportion of tweets pertaining to the event has fallen to 75% by 5 PM, note that, by absolute numbers, we will find that the event is more talked about at 5 PM than at 10 AM since 75% of the volume at 5 PM is much more than 100% of the volume at 10 AM.

Of course, this is not necessarily the wrong trend to observe. Certainly, if we are interested in the number of users being exposed to a particular piece of content, absolute numbers will be more important. However, if we assume that fraction of Twitter event-related content in a particular period roughly reflects the fraction of all Twitter users who are interested in the event, then the relative abundance between two periods is a more accurate indication of differences in event engagement across time.

4.2. Disaster-related content collection

Determining the absolute number of relevant tweets in a given time period is a non-trivial task: tweets provide users with nearly unlimited ways of mentioning and discussing a particular event. Despite this fact, it is well established that Twitter users tend to converge on the use of a small number of hashtags when discussing a particular topic or event, with hashtag popularity dependent on contextual features and length and preferential attachment being correlated with adoption [20,21]. Thus, following prior work, we use a set of known hashtags to flag tweets about a given disaster (e.g., [1,7,16]). In order to obtain absolute and relative volume measures for a given time period, we count the number of flagged tweets in the time period (and in the case of relative counts, divide by the total number of tweets in that time period). For each dataset we used the 5 hashtags shown in Table 1 which were among the most commonly used to indicate event-related content.

Table 1. The keywords used to collect event-related tweets for each event.

Hurricane Sandy	Oklahoma Tornado	Boston Bombings
#sandy	#oklahoma	#boston
#hurricanesandy	#moore	#bostonmarathon
#hurricane	#tornado	#bostonbombing
#nyc	#okwx	#watertown
#frankenstorm	#okc	#marathon

These hashtags were then used to calculate topic frequencies for each event, with the presence of any of these hashtags in a tweet being labeled as event-related.

Finally are part of the analysis of each event, we consider the event-specific news media engagement. Specifically, since traditional news outlets have the choice to tweet about anything they want, we take the presence of greater numbers of tweets from news sources to indicate a higher level of media engagement in the event. We select equal time periods around the peaks of the event curves and search for the presence of tweets from the following accounts in the event-related tweets: @CNN, @BBC, @FOX, @CBS, @NBC, and @ABC.

These accounts were selected to proxy for the attention given to a specific event by the media industry. In order to create this list, we extracted the most frequent twitter handles from the entirety of the three datasets followed by selecting the ones that related to news sources.

5. Results

5.1. Effects of time correction

We introduced the measure of time-compressed frequency to reduce the effect of varied user activity at different times of a day. The plots in Figure 2(a,c,e) were generated to study the difference in our measure and regular frequency. For there to be a pronounced effect, we need an event that spans over longer durations, to be actually impacted by “time of day.” As we see in the case of Sandy, the measure boosts the curve during early morning and afternoon, when there is less user activity. The measure is subdued in short span events, since they are less likely to be affected by the vast fluctuations in the Twitter volume frequency, as seen in Figure 2(e), in the case of Boston Bombings.

Though our measure does not change where the curve attains its peak for these events, it does boost the portions of the curve with lower user activity, and therefore represents a better picture in terms of how much of the stream constituted of event-related tweets at a particular instant rather than the absolute numbers.

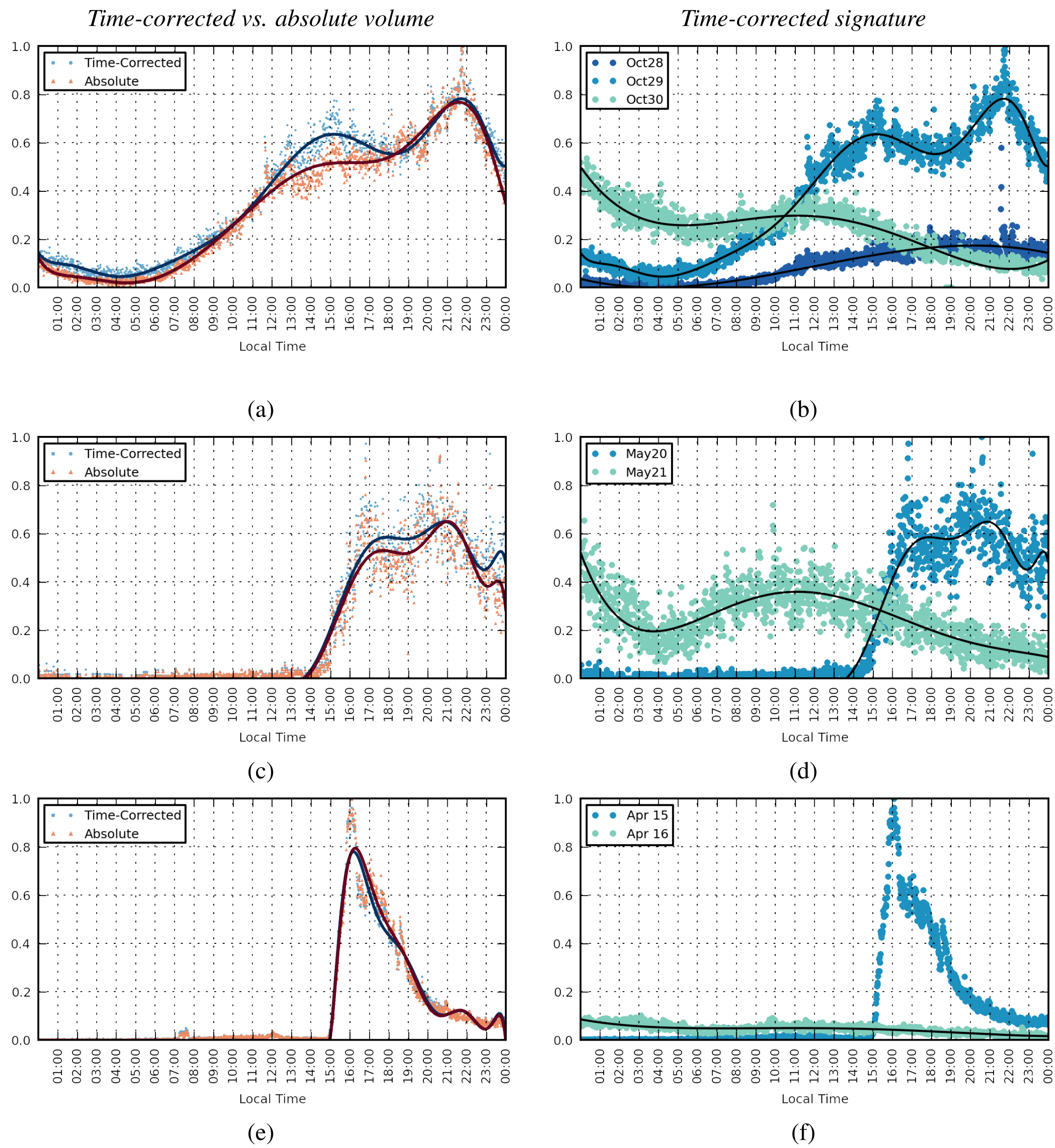

Fig. 2. Twitter signatures for (a,b) Hurricane Sandy, (c,d) Moore tornado, and (e,f) the Boston Bombing. Panels (a), (c), and (e) contrast the absolute measure of volume with the time-corrected formulation. As can be seen, differences emerge, particularly in the case of Sandy which lasted longer than the other two events. Panels (b), (d), and (f) show the complete Twitter signatures for each event.

5.2. Hurricane Sandy

Figure 2(b) displays the event curves for Hurricane Sandy. The point of incidence is at 8:00 PM, Oct 29 (landfall). We see a significant level of event-related Twitter content prior to the incident, due to the anticipation of the storm. Also, notice the gradual accession from the lowest point of the curve to the highest on Oct 29. Studying the span of the curve, we see that high volume of tweets for Sandy is sustained over a considerably large time period, being over 0.5 of the highest individual peak for almost 12 hours. The decline of the curve continued into the next day is, again, gradual, retaining value over 0.2 (almost 25% of the majority peak) over a significant amount of time.

5.3. Oklahoma tornado

In case of the tornado, the point of incidence is at about 3:15 PM, May 20, when the tornado enters Moore. Since there was a general tornado lookout and warning sirens before the incident, we can see some event-related Twitter volume present before the tornado arrived. The high activity persists over a shorter time period than Sandy, falling below 0.5 in a few hours.

5.4. Boston Marathon Bombings

In contrast to the prior two events, the victims of the Boston Bombing possessed no foreknowledge of the attack. Thus, we see virtually no traffic prior to the explosion, followed by a precipitous rise at the point of incidence (twin blasts at 3 PM). The slight earlier traffic is from the marathon itself with a bump at noon coinciding with the men and women winners.

The span of the curve is very small compared to the other two events, going down below 0.5 in almost two hours with an almost negligent plateau. Volume falls quickly such that the next day's volume is dwarved by the peak on the day prior.

5.5. Event peaks and news media engagement

For all three events, we identified the absolute peak of the curve (the highest per-minute volume of tweets that was registered for every event) and computed the number of news organization-originating tweets produced in the hour straddling the peak (see Table 2).

Table 2. At the peak of each event's signature the (a) total number of event-specific tweets and (b) total number of event-specific tweets generated by news organizations. As can be seen, the overwhelming majority of content generated around the peak of the Boston bombings signature was generated by news organizations.

Dataset	Total	News only
Boston Bombings	1237	1289
Hurricane Sandy	623	162
Oklahoma Tornado	105	41

6. Discussion

Our objective in this study was to characterize the shape of disaster-related content on Twitter and connect aspects of the shape to attributes of the events themselves. Since the traffic related to any event is correlated with the amount of user-interest it generates, we consider several factors that could affect the extent to and duration for which users engage with social media over an event. Additionally, we comment on the impact that such factors could have on the design of future event Twitter content collection systems.

6.1. Foreknowledge

The datasets we studied had different levels of available foreknowledge at the point of incidence. Sandy, with the highest level of foreknowledge, had the highest volume of traffic preceding the point of incidents. Moreoever, it also had a gradual rise in volume after the point of incidence. On the other extreme, the Boston bombings afforded no foreknowledge. Thus, the preceding Twitter volume was lowest for the keywords considered and the volume experienced the fastest growth, taking only an hour to reach the peak. We suspect that this steep rise derives, in part, from the level of urgency associated with an event. An anticipated event gives more opportunity for people to prepare, reducing the overall urgency and novelty of information at the time of incidence. This suggests that different data extraction methods may be needed depending on the degree of foreknowledge of the disaster being analyzed.

6.2. Event Duration

Studying event duration over the three datasets, we find an indication that events with longer impact span have a larger sustained plateau in the signature. This implies, somewhat intuitively, that systems should accommodate sustained and elevated levels of event-related content in proportion to the duration of the event.

6.3. Severity

The severity of an event affects the duration of the recovery period: more severe events require longer recovery times. In the event signatures, we find that events with a lower overall severity in terms of number of people affected and immediate damage done to the affected area (e.g., Boston Bombing vs. Hurricane Sandy) declines more quickly. A natural explanation for this observation is that, since the effects of more severe events persist even after the original hazard itself is gone, people tend to keep talking about it, which yields a more gradual decline in interest. This suggests that disaster collection techniques might incorporate a tunable model of event decline as a function of projected severity.

6.4. News Media Engagement

Given the general climate of news media reporting, the coverage of events by news organizations may correlate strongly with the sensational value of the event (rather than the actual dollars of damage, for example). We find support for this in our data as the Twitter signature for the Boston bombing yielded much higher news tweets at the peak than the other events. Certainly, the terrorist and completely unanticipated nature of the event gave the Boston bombing a unique "shock" value, when compared against the other two events.

Somewhat strikingly, this news media engagement also could explain the relative height of the peak in each event. Again, the Boston bombing had a peak that was nearly two times higher than Hurricane Sandy (and 10 times higher than the Oklahoma tornado). In this particular case, this trend clearly derives from the extent of Twitter-based news coverage. For such events, then, content collection systems may benefit from having richer models of the breakdown of volume in terms of source (news vs. personal) which can be used to calibrate the stream to the kind of event.

7. Conclusions

This study is part of a much larger initiative aimed at extracting meaningful information from social media during disasters that is relevant to informing victims, first responders, and the longer-term crisis response planning process. Through the analysis of the Twitter signatures for three different events, we uncovered several ways in which event signatures vary. We have also proposed several factors that explain the variance observed. We consider an important direction for future work to be the further investigation of these signature characteristics and the disaster attributes which drive them — ultimately moving towards a principled typology for disasters which can guide the development of better, more accurate tools for extracting and analyzing disaster-related social media content in real-time.

References

[1] Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010, April). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1079-1088). ACM.

[2] Kavanaugh, A., Yang, S., Sheetz, S., Li, L. T., & Fox, E. (2011). Microblogging in crisis situations: Mass protests in Iran, Tunisia, Egypt. In ACM. CHI.

[3] Sakaki, T., Okazaki, M., & Matsuo, Y. (2010, April). Earthquake shakes Twitter users: real-time event detection by social sensors. In Proceedings of the 19th international conference on World wide web (pp. 851-860). ACM.

[4] IPalen, L., Starbird, K., Vieweg, S., & Hughes, A. (2010). Twitter-based information distribution during the 2009 Red River Valley flood threat. Bulletin of the American Society for Information Science and Technology, 36(5), 13-17.

[5] Brown, S. (2011) Twitter Usage in Times of Crisis. Journal of Digital Research and Publishing, 29.

[6] Bruns, A. (2012). At times of crisis, Twitter shines brightest. Continuity, (Q3), 16-17.

[7] Heverin, T., & Zach, L. (2010). Microblogging for Crisis Communication: Examination of Twitter Use in Response to a 2009 Violent Crisis in the Seattle-Tacoma, Washington, Area. ISCRAM.

[8] Soboroff, I., McCullough, D., Lin, J., Macdonald, C., Ounis, I., & McCreadie, R. (2012). Evaluating real-time search over tweets. Proc. ICWSM, 943-961.

[9] Lau, C. H., Li, Y., & Tjondronegoro, D. (2011). Microblog Retrieval Using Topical Features and Query Expansion. In TREC.

[10] Tao, K., Abel, F., Hauff, C., & Houben, G. J. (2012). What makes a tweet relevant for a topic. Making Sense of Microposts (# MSM2012), 49-56.

[11] Yang, J., & Leskovec, J. (2011, February). Patterns of temporal variation in online media. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 177-186). ACM.

[12] Lin, Y. R., Margolin, D., Keegan, B., Baronchelli, A., & Lazer, D. (2013). # Bigbirds Never Die: Understanding Social Dynamics of Emergent Hashtag.arXiv preprint arXiv:1303.7144.

[13] Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., & Miller, R. C. (2011, May). Twitinfo: aggregating and visualizing microblogs for event exploration. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 227-236). ACM.

[14] Sinnappan, S., Farrell, C., & Stewart, E. (2010). Priceless tweets! A study on Twitter messages posted during crisis: Black Saturday. In 21st Australasian Conference on Information Systems.

[15] Doan, S., Vo, B. K. H., & Collier, N. (2012). An analysis of Twitter messages in the 2011 Tohoku Earthquake. In Electronic Healthcare (pp. 58-66). Springer Berlin Heidelberg.

[16] Terpstra, T., de Vries, A., Stronkman, R., & Paradies, G. L. (2012, April). Towards a realtime Twitter analysis during crises for operational crisis management. In ISCRAM12: Proceedings of the 9th International ISCRAM Conference.

[17] Centers for Disease Control and Prevention (CDC. (2013)). Deaths associated with hurricane sandy-october-november 2012. MMWR. Morbidity and mortality weekly report, 62(20), 393.

[18] Lott, N., Smith, A., Houston, T., Shein, K., & Crouch, J. (2013). Billion Dollar US Weather Disasters, 1980-2013. National Oceanic and Atmospheric Administration.

[19] Department of Homeland Security (DHS. (2013)). Boston Marathon Bombings: Medical Response Activities at the Incident Site. Lessons Learned Information Sharing.

[20] U.S. Security and Exchange Commission (SEC. (2013)). FORM S-1, Twitter, Inc. SEC EDGAR Archives Accession No. 0001193125-13-390321.

[21] Ma, Z., Sun, A., & Cong, G. (2013). On predicting the popularity of newly emerging hashtags in Twitter. Journal of the American Society for Information Science and Technology, 64(7), 1399-1410.

[22] Cunha, E., Magno, G., Comarela, G., Almeida, V., Gonalves, M. A., & Benevenuto, F. (2011, June). Analyzing the dynamic evolution of hashtags on twitter: a language-based approach. In Proceedings of the Workshop on Languages in Social Media (pp. 58-65). Association for Computational Linguistics.