Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media

# Event Detection via Communication Pattern Analysis

**Flavio Chierichetti***
Sapienza University
Rome, Italy
flavio@di.uniroma1.it

**Jon Kleinberg**
Cornell University
Ithaca, NY
kleinber@cs.cornell.edu

**Ravi Kumar**†
Google, Inc.
Mountain View, CA
ravi.k53@gmail.com

**Mohammad Mahdian**†
Google, Inc.
Mountain View, CA
mahdian@google.com

**Sandeep Pandey**†
Twitter, Inc.
San Francisco, CA
sandeep.pandey@gmail.com

**Abstract**

Social media applications such as Twitter provide a powerful medium through which users can communicate their observations with friends and with the world at large. We have witnessed live reporting of many events, from soccer games in Johannesburg to revolutions in Cairo and Tunis, and these reports have in many ways rivaled the content provided by the official media. Tapping into this valuable resource is a challenge, due to the heterogeneity and noise inherent in real-time text, diversity of languages, and fast-evolving linguistic norms. In this paper we seek to analyze a tweet stream to automatically discover points in time when an important event happens, and to classify such events based on the type of the sentiments they evoke, using *only* non-textual features of the tweeting pattern. This results not only in a robust way of analyzing tweet streams independent of the languages used; it also provides insights about how users behave on social media websites. For example, we observe that users often react to an exciting external event by decreasing the volume of communication with other users. We explain this effect through a model of how users switch between producing information or sentiments and sharing others' news or sentiments. We develop and evaluate our models and algorithms using several Twitter data sets, focusing in particular on the tweets sent during the soccer World Cup of 2010. This data set has the feature that the underlying ground truth is well-defined and known whereby goals serve as events.

## Introduction

Understanding how a large population reacts to a major event in real-time is a fundamental question that, until very recently, was extremely difficult to approach in a large-scale quantitative fashion. With the growth of real-time social information systems such as Twitter, however, it becomes possible to analyze the behavior of large groups as they observe and participate in such events.

Watching how Twitter has been used during episodes such as sporting events, large gatherings, political protests, and emergency situations, it is clear that the medium is employed by different people during an unfolding event for a very broad range of purposes: it is used for reporting both by participants at the site of the event and by observers from far away; it is used for communication and coordination among people involved in the event; and it is used to express collective reactions to new developments as they shape the course of the event. While earlier work has considered the structure of communication within Twitter across longer time-scales (Golder and Yardi 2010; Huberman, Romero, and Wu 2009; Kwak et al. 2010; Romero and Kleinberg 2010), this combination of rapidly evolving real-time events with the behavior of large populations involved in the event represents an important and distinctive role for Twitter.

The central questions we consider are how to identify new developments in an event stream of tweets, and how these new developments in the event influence users' discussion, reporting, and communication behavior. We study how to extract the sequence of key events in a news story from the raw numbers of tweets and retweets that take place during these events. We propose an efficient linear classifier for this task. We then study how the evolution of the event affects the users' activity, in particular, the balance between producing new information and forwarding existing information and the level of communication among individuals.

For our analysis, we focus on a comprehensive collection of tweets spanning three episodes of varying lengths: the month-long 2010 Soccer World Cup, the 2011 Academy Awards presentation, and the 2011 Super Bowl. These datasets provide an ideal testing ground for studying the global reaction to an evolving event: the constituent parts of the event in our case are known, with exact time-stamps (e.g., the starts and ends of each game and key events such as goals within each game); there were strong emotions and active communication associated with the event; and different segments of the population were strongly supporting divergent outcomes (one team winning versus another). All of these are ingredients that one expects to see, potentially in reduced forms, in a wide range of global events that have a significant projection on Twitter.

**Primary and secondary information**. The dynamics of an event unfold at many different time-scales. In our case, for example, the full World Cup was a month-long event, with games comprising short, intense sub-events nested inside the

*Part of this work was done while the author was at Cornell and part while visiting Yahoo! Research and Google. This work was partially supported by a Google Faculty Award.

†Part of this work was done while the author was at Yahoo! Research.

longer whole, and with goals and other pivotal moments in the games serving as a further level of sub-event nested within the games. We begin by considering how user behavior is altered during these intense sub-events; we can approach this question at two different levels of scale by considering (a) games nested within the full World Cup and also (b) short time windows around a goal nested within a game.

We find that intense sub-events produce a fundamental shift in the generation of *secondary* forms of information on Twitter. We view retweets (the forwarding of information) and communication via messaging as forming this layer of secondary information, since they consist of a body of partly social activity operating on top of a base level of tweets being generated by users. We refer to all other tweets as *primary information*. During a significant sub-event, a characteristic pattern emerges in which the generation of secondary information is diminished during the sub-event itself, but then secondary information appears at a temporarily elevated rate during a window of time following the sub-event. There is an intuitive basis for this kind of "heartbeat" pattern: as the sub-event is actually unfolding, users are devoting more of their time to reporting on and discussing the sub-event, and hence have less time for producing secondary information. Once the sub-event has subsided, however, there is a glut of new tweets that can be retweeted, as well as communication opportunities for discussing the sub-event retrospectively, and so the generation of secondary information rapidly increases. This trajectory thus suggests a complex complementarity/substitutability relationship between the volume of primary tweets and the volume of (secondary) social interactions.

We further argue that this heartbeat pattern can be an effective component of applications to detect sub-events and estimate their intensity. While sub-events generally involve spikes in the volume of tweets, there tend to be many spikes on Twitter over the course of a short period of time, and so searching for spikes directly is not a very discriminating test. Tracking the balance of primary and secondary information, on the other hand, makes for a more powerful filter, since it requires not just a spike in volume, but a simultaneous drop in the level of secondary information. The extent to which these effects move in opposite directions can be further used to measure the intensity of the reaction to the sub-event.

We build a mathematical model that formalizes the intuitive picture for how the heartbeat phenomenon occurs. In the model, every user has the same probabilities of tweeting or retweeting in the absence of an unusual event. When an unusual event happens, each user becomes interested in it independently by flipping a coin. An interested user will tweet or retweet about the event before tweeting or retweeting about something else. We show how our model is able to generate the aggregate behavior we observe in the temporal vicinity of sub-events, i.e., we show that our model naturally produces the heartbeat pattern that we observe consistently in the datasets. Intuitively, this happens because people who are interested in the event will need to produce primary information about the new event (new tweets), before becoming able to share existing secondary information about the same event (retweets).

## Related work

There have been several recent papers on automatically building event reports as witnessed by users from their tweets. Sakaki, Okazaki, and Matsuo (2010) showed that tweets can be used to detect earthquakes. They proposed an algorithm to detect a target event, where their algorithm is based on classification and a spatiotemporal model; see also the recent work of Qu et al. (2011) on earthquakes in China. Petrovic, Osborne, and Lavrenko (2010) considered the problem of detecting new events in a stream of tweets and Sankaranarayanan et al. (2009) identified news topics and clustered tweets for each topics. Chen and Roy (2009) and Luo, Tang, and Yu (2007) considered similar problems on other social media applications such as Flickr. Most of these work are in an unsupervised setting.

A few research papers have also studied tweets during specific events, including sporting events. Shamma, Kennedy, and Churchill (2009) studied tweet usage during the 2008 Presidential Debates and showed that Twitter activity serves as a predictor of topic changes in the media event. Chakrabarti and Punera (2011) used HMM-based methods to summarize a sequence of tweets produced during a sporting event. Most recently, Zhao et al. (2011) considered the problem of inferring, from a stream of tweets, the touchdowns during an American football game; they show that key events can be recognized to within 40 seconds of their occurrence. While their work is the closest to ours in terms of the domain, they are more focused on real-time event recognition. We obtain a much higher precision (sometimes to within 15 seconds of a goal).

Becker, Naaman, and Gravano (2010) considered the problem of developing similarity metrics to help clustering of media to events; they work in an unsupervised fashion and focus on developing the similarity metric rather than try to align a set of tweets to a set of events. In another work (Becker, Naaman, and Gravano 2011a), they consider the problem of detecting real-world events in tweets; see also (Becker, Naaman, and Gravano 2009). They also study the problem of selecting high-quality event content from tweets (Becker, Naaman, and Gravano 2011b).

The topic of new event detection in a time series has been studied for a while; see the work Allan, Papka, and Lavrenko (1998). Kleinberg (2002) formalized the concept of event 'burstiness' and showed how one can select the "bursty" words in a stream of text using a version of the Viterbi algorithm. For additional references on these topics, see the surveys (Allan 2002; Kleinberg 2004).

The social dynamics behind Twitter continues to be extensively studied. Yardi and boyd (2010) and Conover et al. (2011) studied the effect of polarization on Twitter. For some early papers investigating the social network of Twitter, see the work of Java et al. (2007) and Krishnamurthy, Gill, and Arlitt (2008). Huberman, Romero, and Wu (2009) studied the `@` posts in tweets and boyd, Golder, and Lotan (2010) studied the retweeting phenomenon.

## Experimental setup

The dataset for our experiments comes from the Twitter Firehose, which contains all the tweets during the entire lifetime of Twitter. Each tweet has a lot of important metadata associated with it: the text, the geographic location of the tweet and of the user, and the time-stamp. If the tweet is produced in response to another tweet, then this information is also included. The tweet text itself contains a wealth of information and embodies certain conventions adopted by the Twitter community. For example, retweets are characterized by the symbol @ followed by a user name, who is the originator of the tweet. Special strings (called *hashtags*) are represented by prefixing them a # symbol; several applications such as Twitter search treat such tokens specially. In fact, we will heavily depend on hashtags in our analysis.

As one can imagine, the amount of total data is staggering. During the period of interest, on average, there were more than 100M tweets per day; this amounts to a total of tens of billions tweets that we have to analyze. Processing this massive data is only possible with the use of a map-reduce system. All of our analyses heavily use the power of distributed processing to extract various pieces of information.

**Data**. We focus on three major *social episodes*: the 2010 soccer World Cup held in South Africa (denoted WORLDCUP), the 2011 Academy awards held in Hollywood (denoted OSCARS), and 2011 Super Bowl XLV, which took place in Arlington, Texas (denoted SUPERBOWL). These three datasets cover a broad spectrum of social episodes, including different geographic localization (city to country), different time periods (single day to almost half a year), multiple sub-episodes (WORLDCUP) vs. a single episode (OSCARS), and different genres (sporting and entertainment).

The datasets were collected using the following methodology. For each dataset, we first assembled the following pieces of information by hand.

(i) Timeline: the start and end time of the episode.

(ii) Events: a list of all events in the episode including the features for each of the events, with some of them identified as *key events*. In all our cases, each event featured at least one person, denoted by the first and last names.

(iii) Hashtags: a list of all hashtags that could have been used to refer to the episode. As we mentioned earlier, these hashtags will be used to identify all the tweets that are related to the episode. Of course, since hashtags are only a convention, there will be tweets about the episode that may not use one of our listed hashtags. We will not consider such tweets, and this does not appear to pose a significant limitation due to the total volume of tweets.

Using the hashtags, we obtain all the tweets about the social episode. Table 1 provides more details about the dataset. In some cases, we perform additional processing to identify users who participate a lot in tweeting about the episode. We define a user to be *active* if he/she has used at least 10 episode-related tags during at least one of the sub-episodes. We then do a manual check on the most frequently tweeting users to see if they represented a real human or if they could be bots/spam. We obtain a threshold from this scan and eliminate users who tweet more than this threshold during the period.

From the sequence of tweets sent by the users, we extract various time-series such as the volume of tweets, frequency of usage of various words, and other indicators. We also obtain these time-series on the general population of Twitter users; this way, we can normalize the data and avoid artifacts such as the time-of-day effects. We also extract time-series about the social interactions by the users. As we mentioned, that there are two kinds of social interactions in Twitter: mentioning of another user (which can be a retweet) and replying to another user.

**Methodology**. We now describe the methodology used to assemble the datasets that we use in the paper. Note that in all the cases, the absolute time-stamps of the key events will be used in our evaluation.

WORLDCUP. The duration and events for World Cup 2010 are available publicly (soccerstand.com). There were 64 games with non-key events such as 253 yellow cards and 17 red cards. Each of 32 countries participating in the tournament was assigned its own hashtag (e.g., Netherlands was denoted by #ned, Uruguay was denoted by #uru); in addition, a generic hashtag of #worldcup was also used. By convention, to refer to the Netherlands–Uruguay game, most users would use one or more of #ned, #uru, #worldcup tags while tweeting; this is especially so during the game. Some of the games were held concurrently.

OSCARS. For the Academy awards, we assembled the events from oscars.nytimes.com/dashboard, which contains the time when a particular award was announced. By analyzing the top hashtags that were used on the day of the Oscars, we were able to find out all the tags that were related to the ceremony.

SUPERBOWL. For the Superbowl, we assembled the events from blogs.wsj.com/dailyfix/. Unlike the other sporting events, NFL is more challenging since the official data does not contain the absolute time-stamps, which are necessary to align them against the tweets; this difficulty was also noted in a recent paper (Zhao et al. 2011). In addition to `#sb45`, `#sbxlv`, `#superbowl`, we also used the names of the two competing teams, `#packers` and `#steelers`.

## Key events and tweet volume

We start by studying simple non-textual statistics that we can readily extract from the set of users and their tweets. In particular, we focus on the volume of information generated by these users. We also ask if it is possible to align the key events by just considering the volume of tweets immediately after the event.

For WORLDCUP, Figure 1 shows a graph that aggregates the average number of tweets over all games, scaling the time to ensure that the length of each game is precisely 105 minutes. The green line represents the absolute number, and the red line represents the number divided by the total number of tweets sent by any twitter user at that minute (therefore, any potential time-of-day effect is alleviated in the red

| Data | Start (GMT) | End (GMT) | #Tweets | #Key events | Sample hashtags | Sample events |
|---|---|---|---|---|---|---|
| WorldCup | Jun 11, 2010 | July 12, 2010 | 342M | 159 | `#worldcup`, team tags | goal |
| SuperBowl | Feb 6, 2011 | Feb 7, 2011 | 1.49M | 7 | `#sb45`, `#superbowl` | touchdown |
| Oscars | Feb 12, 2011 | Feb 13, 2011 | 1.61M | 24 | `#oscars`, `#redcarpet` | award |

Table 1: Details of the datasets. Key events are boldfaced.

curve). We plot the curves corresponding to the number of words and the number of characters tweeted, and they both look very similar to the number of tweet curves: the volume of tweets quickly increases as the game is about to start, stays at about the same level during the first half, drops during the half-time, and then returns to an even higher level as the second half starts, and keeps increasing with a sharp peak at the end of the game. After the game, the volume drops quickly, but to a level still above the level before the game (post-game chatter). Figure 2 shows the time-series

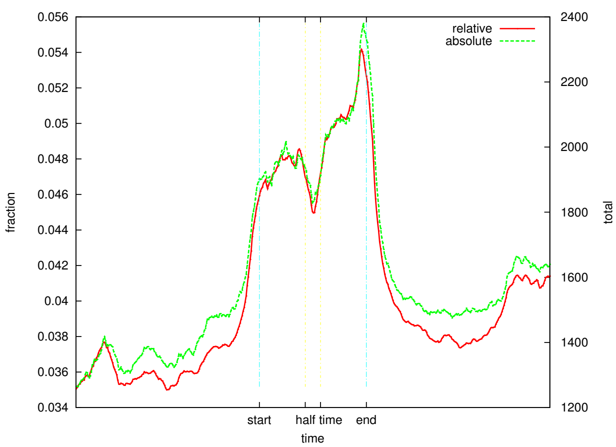


Figure 1: Average tweets volume during a World Cup game.

for SuperBowl, with the corresponding events (touchdowns, goals) marked and the time-series for Oscars, with the corresponding key events (awards) marked. Unlike the WorldCup case, the average volume goes down after the game/ceremony is over, when compared to the beginning. This is presumably due to the "build-up" caused by TV and online media and the buzz associated with it.

Note that in both cases, each key event causes a peak in the volume of tweets. But, it is not the case that each peak corresponds to a key event. Also, the volume significantly increases during the half-time for SuperBowl. Furthermore, since the interval between the key events in Oscars is very short, it is hard to accurately align each peak with the corresponding key event.

## Information production vs. social interaction

We now turn to the pattern of social interactions among the users. First, we study the average number of messages replied to during a game. Figure 3 shows the plot. It is illustrative to compare this plot against Figure 1. The relative

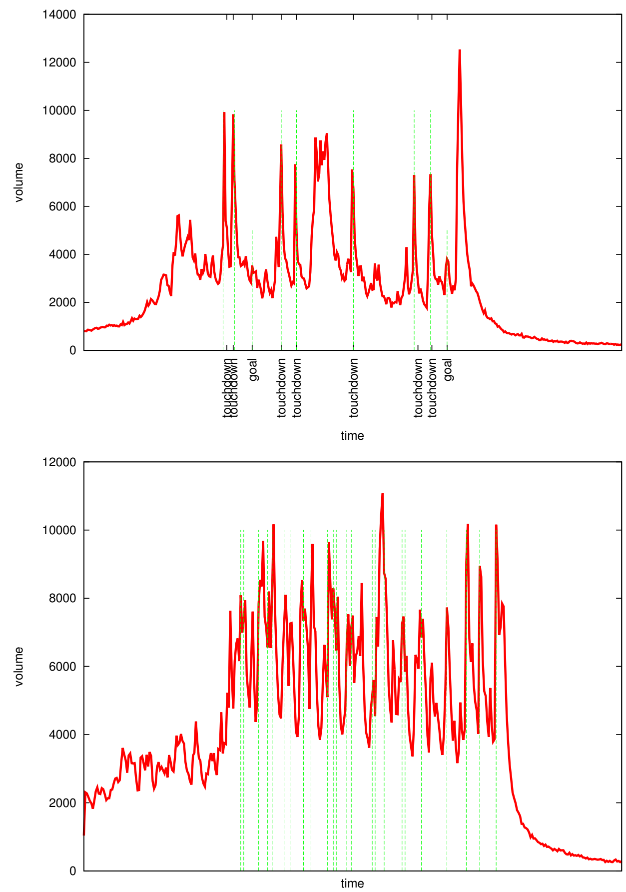


Figure 2: Events and tweet volume during SuperBowl and the Academy awards.

numbers are almost a mirror image of those in Figure 1. We investigate this phenomenon more closely by looking at similar patterns around a key event (`goal`); see Figure 4.

The pattern that we observe here is quite surprising: at a time when a key event happens, we see that the users become *less social* rather than more social. However, quickly after the event, the users get back to socializing, this time at a higher level. This is a pattern similar to the "heartbeat" pattern in electrocardiographs.[1]

[1]This segment of the heartbeat pattern is known as the *QRS complex*. There is a significant body of literature on automatically detecting QRS complexes in electrocardiographs. However, the algorithms used in this literature usually rely on the periodicity of the heartbeat, and therefore are not useful in detecting similar patterns during a game.

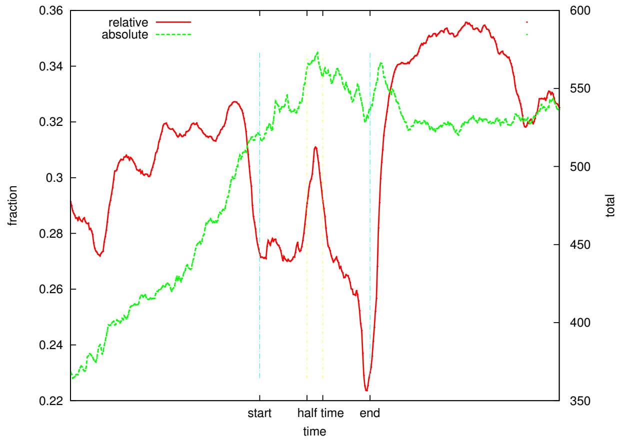


Figure 3: Average number of messages replied to during a World Cup game.

An intuitive explanation for this phenomenon is as follows: information generation (i.e., new tweets) and social interaction (i.e., replying to someone's tweet) are in a way both complement and substitute activities. They are complements since in order to reply to a tweet, that tweet must have been generated in the first place. They are substitutes since a user has a limited amount of time/attention, and the more time she spends tweeting, the less time she will have to reply to others' tweets. This can cause the volume of the social interaction to decrease at the moment that a new event has happened and users are busy tweeting about it, while after some time, it will increase the volume of social interactions. Later we will formalize this intuition and build a model that can generate patterns very similar to what we observed in this section.

Figure 6 shows similar plots for SUPERBOWL and OSCARS. Clearly, as in the case of WORLDCUP, we observe the heartbeat pattern: at the moment of the key event, the users becomes less social rather than more social.

## Event detection

In this section we consider the problem of finding key events in a tweet stream using only the tweet and retweet counts. We show that a simple logistic regression approach allows us to pinpoint most of the goals in our World Cup dataset, with a precision of 15 seconds. The point of this exercise is to show that there is plenty of signal in non-textual features such as the pattern of the tweet and retweet volumes for detecting events. Most notably, the pattern of the retweet volume plays an important role in improving the accuracy of prediction.

This dataset has 159 positive instances (windows of 15 seconds containing a goal) and 38,070 negative ones (windows of 15 seconds not containing a goal during or around one of the games). Our classifier returns 66 false negatives, and only 17 false positives. The five-fold cross-validated error rate of this classifier is about $0.197$ percent. To put this number in perspective, the error rate of a classifier that classifies every instance as negative is $0.414$ percent. This is

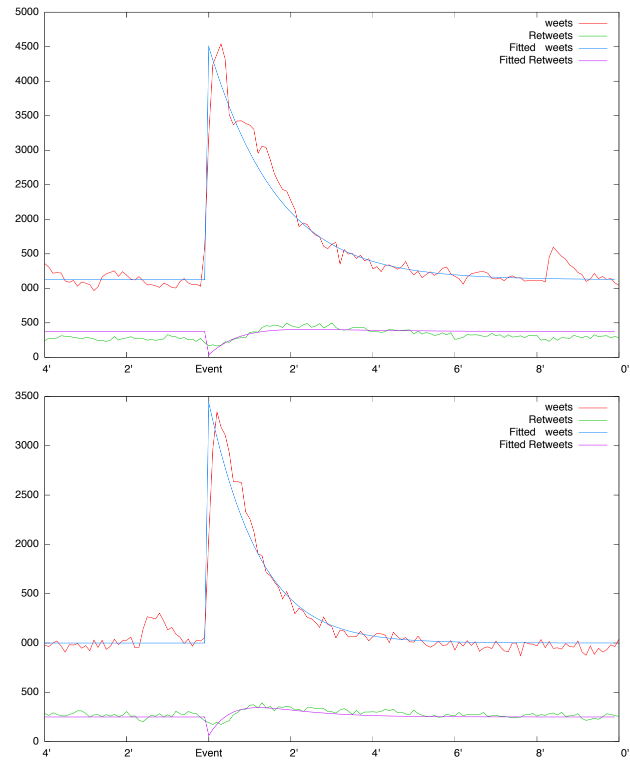


Figure 4: Number of tweets and retweets around two goal events in the World Cup (the one above is the first goal that Brazil scored against North Korea, in their 15 June 2010 game; the one below is the only goal scored by Mexico against Argentina, in their 27 June 2010 game.)

more than a fifty percent improvement in accuracy.

A better-scaled measure of the quality of the classifier is the so called *Matthews correlation coefficient*. We recall that the Matthews correlation coefficient of a binary classification that produces $P$ true positives, $N$ true negatives, $p$ false positives, and $n$ false negatives is equal to

$$\frac{P \cdot N - p \cdot n}{\sqrt{(P+p)(P+n)(N+p)(N+n)}}.$$

This functional is always in the range $[-1, 1]$, a value of 1 corresponds to a perfect classification (i.e., $p + n = 0$), and a value of $-1$ to a completely wrong one (i.e., $P + N = 0$). Predicting always positive, always negative, or at random results in a Matthews correlation coefficient of zero (or concentrated around zero). The Matthews correlation coefficient of our classifier is close to $0.707$, which is quite large.

We now describe the simple, yet very effective, linear classifier that we used. As already mentioned, our classifier only uses tweet and retweet counts, in particular, the number of tweets and the number of retweets in each time window. In fact, the classifier even uses only a tiny part of this information. To classify a window $i$ as an *eventful*, or *non-eventful*, the classifier only uses the counts $T(i)$, $R(i)$ of the window $i$, and those of the windows $i-2, i-1, i+1$ and $i+2$, i.e., the classifier uses only 10 integers per window to

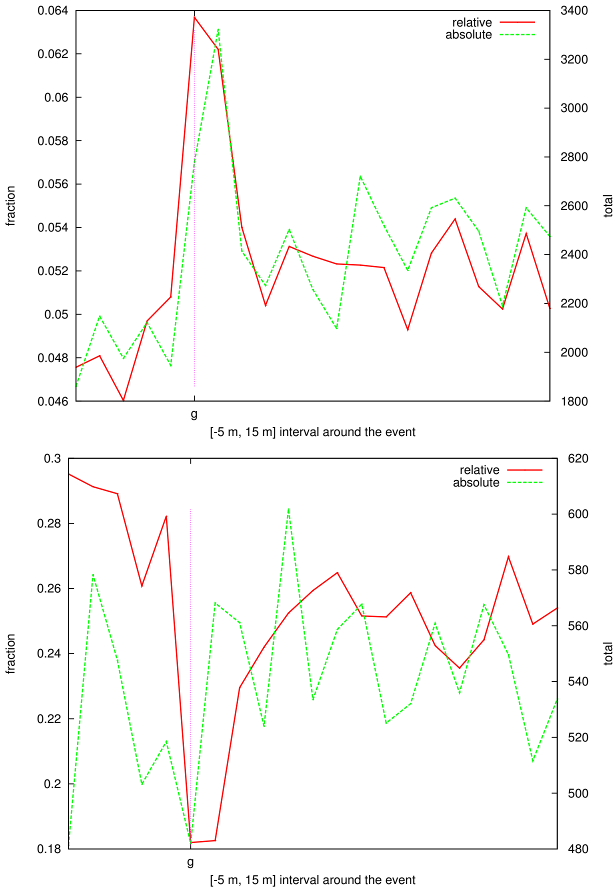


Figure 5: Average number of tweets and messages replied to around a red card event in World Cup.

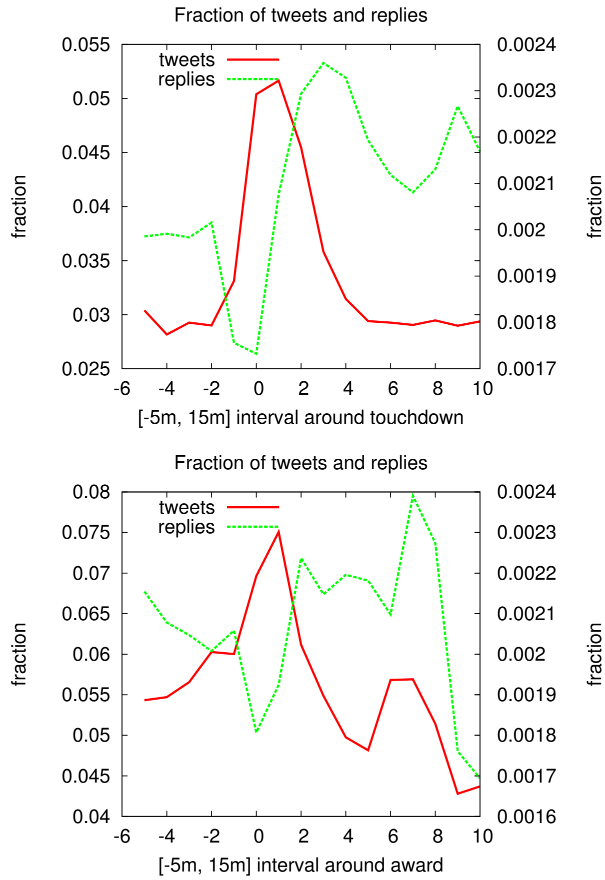


Figure 6: Fraction of tweets and replies around key events for SUPERBOWL and OSCARS.

return its guess.

Moreover, our classifier is linear and is thus very efficient: it classifies a window $i$ as eventful if and only if a linear inequality on the 10 integers holds true. The inequality's coefficients were obtained by running a logistic regression on our 159-goals dataset. The logistic regression produced the following inequality:

$$(-2.31, -56.14, +24.64, +71.76, +11.9) \cdot \begin{pmatrix} T(i-2) \\ T(i-1) \\ T(i) \\ T(i+1) \\ T(i+2) \end{pmatrix}$$

$$+ (-0.70, +80.17, +32.21, -8.46, +38.54) \cdot \begin{pmatrix} R(i-2) \\ R(i-1) \\ R(i) \\ R(i+1) \\ R(i+2) \end{pmatrix}$$

$$\geq 39.25.$$

The heartbeat-shaped curve that we have observed around goals is reflected in the coefficients of the above inequality. Indeed, we know that the number of tweets spikes up right after a goal; correspondingly, the coefficients of $T(i)$, $T(i+1)$, and $T(i+2)$ are all positive and quite high. At the same time, the number of retweets goes down for a little while right after a goal, and the regression chose a negative coefficient for $R(i+1)$.

Using weighted logistic regression and varying the weight of the positive instances, we can explore the tradeoff between precision and recall. The resulting precision-recall graph is shown in Figure 7.

Finally, we note that even though the false positives reported by our classifier are not goal moments, they exhibit tweeting/retweeting patterns similar to a goal moment, and therefore can be considered "important moments" during the game. We do not have any ground truth to evaluate this claim, but manually looking at the set of false positives supports this claim. For example, the non-goal moment scored highest by our classifier is a few minutes before the end of the final game, when Xavi missed a free kick. The second highest-scored non-goal moment is the end of the first game of the World Cup (between South Africa and Mexico), the third is the end of the extra time in the Japan–Paraguay game (which was decided in the penalties).

## Event labeling

We now mention our results for a task related to the previous one: after we detect that a goal event happened, if we know

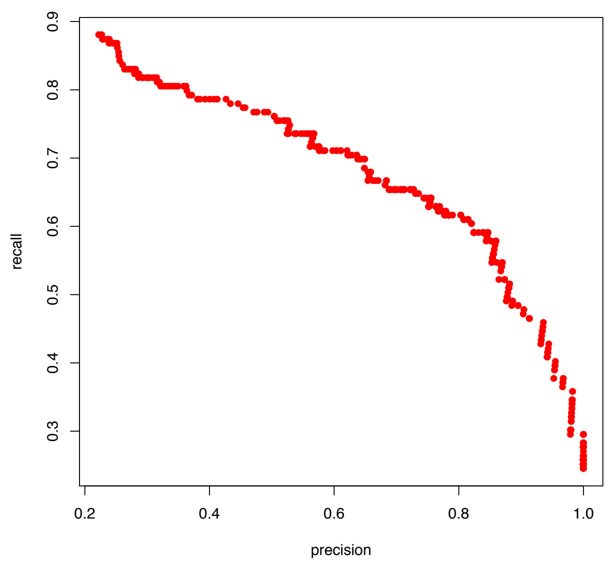


Figure 7: Precision-recall curve for our goal detection task.

that it happened while Team $A$ and Team $B$ were playing, and if we have a partition of users into supporters of Team $A$ and Team $B$, can we determine which of $A$ or $B$ scored the goal? Once again, we wanted to make this prediction without looking at the textual information contained into (re)tweets.

As one can expect, though, this second task cannot be carried out satisfactorily without being able to guess which users are supporters of which team. Therefore, we relaxed our non-text constraint as follows. We used geographic and language information of each user, as well as their hashtag usage[2] to produce an assignment of the users to the 32 teams participating in the World Cup. A user who is assigned to a team is considered a supporter of that team. Users who cannot be detected using these signals to be a fan of one of the teams will remain unassigned.

As it turns out, the tweet volumes are heavily skewed toward the winners. This hints at the following *baseline* classifier for picking the team that scored the goal: just use the numbers of tweets produced by supporters of team $A$ and supporters of team $B$ in that window, and claim that the goal was scored by the one team whose supporters produce a larger number of tweets in a window around the goal time (we used a window of 20 seconds here).

The baseline classifier has an error rate of $19.80\%$ in the dataset (i.e., it got right more that 4 goals out of 5.) It is often not easy to improve over such a good success rate, but as in the case of event detection, using logistic regression with features that capture the pattern of tweeting and retweeting around the goal, we obtained a classifier with a $16.17\%$ error rate. This is almost an $18\%$ relative improvement over the accuracy of the baseline classifier.

[2]For each user, we counted how many times the user used the official hashtag of each single team; we have also guessed the user language through a dictionary approach.

Our classifier used the following features: the difference in the volume of tweets by supporters of team $A$ and supporters of team $B$ two buckets before and two buckets after the goal (i.e., five buckets in total, including the bucket at the time of the goal), and the similar features for the volume of communication tweets (i.e., retweets, replies, and mentions). We picked each bucket to be 20 seconds long. These features are enough to capture the pattern around the goal.

The number of features used to classify a goal was then $2 \times 5 = 10$. The error rate of $16.17\%$ of this classifier was calculated using the Leave-One-Out Cross-Validation (LOOCV) method. The Matthews correlation coefficient of the classifier is about $0.62\%$.

## Model

In this section we present a simple theoretical model that captures the information generation and propagation pattern we observed earlier. Our model is based on the following principle: it is hard to create secondary information if there is only little primary information. In our context, it translates to the following: if only a few users have tweeted about an event, it is unlikely for users to retweet about the event.

In our model, there are $N$ users. The model will support two types of users, namely, *concerned* and *unconcerned*. The behavior of the model is as follows.

(i) An *unconcerned* user, at any point in time, will tweet or retweet something unrelated to the event with probabilities $t_g$ and $r_g$ respectively.

(ii) An event happens (i.e., a goal is scored) at time 0; $n \leq N$ users will care about the event and will become *concerned* at time 0. If a user is concerned she will tweet or retweet about the event before performing any other action. Once a concerned user (re)tweets about the event, she will return to the unconcerned state.

(iii) A concerned user who has not yet posted anything about the event will *tweet* about the event with probability $t_e$ and will *decide to retweet* with probability $r_e$. In the latter case, she will look at a twitter profile chosen uniformly at random: if she finds a tweet or a retweet about the event in that profile, she will retweet it (i.e., will retweet about the event). Otherwise, she will not do anything at that point in time.

The aim of this process is to capture the fact that the time series representing the number of tweets has a spike at the time when the event happens. This spike is induced by choosing the event tweet probability to be larger than the general tweet probability. Moreover, the process also captures the dip of retweets after the event: there will be a few event tweets in the first few seconds after the event and hence the users interested in the event have a small probability of retweeting for some time after the event. After enough time, however, the number of users will have gone back to the unconcerned state (and will therefore behave as before the event) and the users who are still concerned and who decide to retweet will have a large probability of retweeting about the event (since a larger number of users will have tweeted or retweeted about the event).

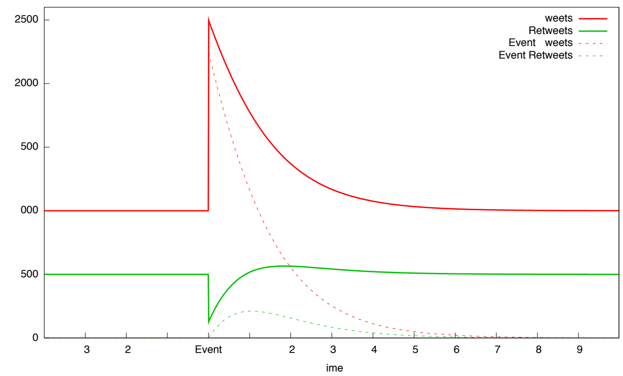


Figure 8: The curves $T'(t)$ and $R'(t)$ (in bold), and $\tau'(t)$ and $\rho'(t)$ (dashed), with $N = 5000$, $n = 3/4 \cdot N$, $t_g = 0.2$, $r_g = 0.1$, $t_e = 0.6$ and $r_e = 0.3$.

We now analyze the model in detail. Let $\tau(t)$ (resp., $\rho(t)$) be the number of people who have tweeted (retweeted) about the event at time $t$. (We assume the event happened at time 0.) We assume that every person posts at most one message (be it a tweet or a retweet) about the event. Therefore $\tau(t)$ ($\rho(t)$) also count the number of tweets (retweets) about the event at time $t$. Let $T(t)$ ($R(t)$) be the total number of tweets (retweets) at time $t$. Hence, the derivative $T'(t)$ of $T(t)$ represents the number of tweets that are posted at time $t$ (e.g., the red and the blue curves of Figure 4). Analogously, the derivative $R'(t)$ of $R(t)$ represents the number of retweets posted at time $t$ (the green and the purple curves of Figure 4.) Finally, the derivatives $\tau'(t)$ ($\rho'(t)$) represent the number of tweets (retweets) about the event at a given time. We do not have an empirical plot that shows them (since they are, by nature, hidden); to show how they can look like, we have plotted them in Figure 8.

A version of our model is captured by the following system of differential equations. For $t \geq 0$,

$$\begin{aligned}
\tau'(t) &= t_e \cdot (n - \tau(t) - \rho(t)), \\
\rho'(t) &= r_e \cdot (n - \tau(t) - \rho(t)) \cdot \frac{\tau(t) + \rho(t)}{N}, \\
T'(t) &= \tau'(t) + t_g \cdot (N - n + \tau(t) + \rho(t)), \\
R'(t) &= \rho'(t) + r_g \cdot (N - n + \tau(t) + \rho(t)). \qquad (1)
\end{aligned}$$

We are interested in the functions $T'(t)$ and $R'(t)$, i.e., the tweet and retweet curves. As boundary conditions, we choose $\tau(0) = \rho(0) = 0$, i.e., no tweet and no retweets about the event happened before the event time $t = 0$. For completeness, we also define $T'(t) = t_e N, R'(t) = r_e N$ for $t < 0$.

Next, we solve the differential equation system and prove some of its properties, summarized in Table 2.

**Lemma 1.** *Let $A = t_e N + r_e n$. If $\tau(0) = \rho(0) = 0$, then there is a unique solution to the system (1) of differential equations that satisfies:*

$$\begin{aligned}
\tau'(t) &= \frac{t_e n \cdot A e^{-\frac{A}{N} t}}{t_e N + r_e n \cdot e^{-\frac{A}{N} t}} \\
\rho'(t) &= \frac{t_e r_e n^2 \cdot A e^{-\frac{A}{N} t} \left(1 - e^{-\frac{A}{N} t}\right)}{\left(t_e N + r_e n \cdot e^{-\frac{A}{N} t}\right)^2} \\
T'(t) &= t_g N + \tau'(t) \left(1 - \frac{t_g}{t_e}\right) \\
R'(t) &= \rho'(t) + r_g \left(N - \frac{\tau'(t)}{t_e}\right).
\end{aligned}$$

*Moreover, Table 2 contains some limiting values of the functions in the system.*

*Proof.* If we only consider the variables $\tau(t)$ and $\rho(t)$, which are independent of $T(t)$ and $R(t)$, then by integration, we get that the solutions of System (1) satisfy, for any two constants $\alpha$ and $\beta$:

$$\begin{aligned}
\tau(t) &= \frac{t_e}{r_e} N \cdot \ln \frac{t_e N (t_e N + r_e n)}{r_e \left(\alpha N \cdot e^{-(t_e + r_e \frac{n}{N})t} + \beta (t_e N + r_e n)\right)} \\
\rho(t) &= n - \tau(t) - \frac{\tau'(t)}{t_e}.
\end{aligned}$$

Therefore we have

$$\tau(t) + \rho(t) = n - \frac{\tau'(t)}{t_e}.$$

Hence, the solutions to system (1) satisfy:

$$\begin{aligned}
\tau(t) &= \frac{t_e}{r_e} N \cdot \ln \frac{t_e N (t_e N + r_e n)}{r_e \left(\alpha N \cdot e^{-(t_e + r_e \frac{n}{N})t} + \beta (t_e N + r_e n)\right)} \\
\rho(t) &= n - \tau(t) - \frac{\tau'(t)}{t_e} \\
T'(t) &= \tau'(t) + t_g \left(N - \frac{\tau'(t)}{t_e}\right) \\
R'(t) &= \rho'(t) + r_g \left(N - \frac{\tau'(t)}{t_e}\right). \qquad (2)
\end{aligned}$$

We use the boundary conditions $\tau(0) = \rho(0) = 0$ (Note that these are equivalent to saying that no (re)tweet about the event was published before the event happened, i.e., before time 0.)

Define $A = t_e N + r_e n$. Forcing $\tau(0) = 0$ gives us

$$\alpha = A \left(\frac{t_e}{r_e} - \frac{\beta}{N}\right).$$

Then,

$$\tau(t) = \frac{t_e}{r_e} N \cdot \ln \frac{t_e}{r_e \left(\left(\frac{t_e}{r_e} - \frac{\beta}{N}\right) \cdot e^{-\frac{A}{N} t} + \frac{\beta}{N}\right)}.$$

Therefore,

$$\tau'(t) = \frac{t_e \left(\frac{t_e}{r_e} - \frac{\beta}{N}\right) \cdot A e^{-\frac{A}{N} t}}{r_e \left(\left(\frac{t_e}{r_e} - \frac{\beta}{N}\right) \cdot e^{-\frac{A}{N} t} + \frac{\beta}{N}\right)}.$$

This gives us an expression for $\rho(t)$; setting $\rho(0) = 0$, gives us:

$$\rho(0) = \frac{\beta r_e A - (t_e N)^2}{t_e r_e N} = 0.$$

Therefore,

$$\beta = \frac{(t_e N)^2}{r_e A},$$

and therefore,

$$\alpha = A\left(\frac{t_e}{r_e} - \frac{t_e^2 N}{r_e A}\right) = t_e n.$$

Under the boundary conditions, we then get

$$\tau(t) = \frac{t_e N}{r_e} \cdot \ln \frac{A}{t_e N + r_e n \cdot e^{-\frac{A}{N} t}}.$$

Therefore, $\tau(0) = 0$ and $\tau(\infty) = \frac{t_e N}{r_e} \cdot \ln\left(1 + \frac{r_e n}{t_e N}\right)$. (I.e., a $\frac{\tau(\infty)}{n}$ fraction of the people interested in the event, will eventually tweet about the event.) Also,

$$\tau'(t) = \frac{t_e n \cdot A e^{-\frac{A}{N} t}}{t_e N + r_e n \cdot e^{-\frac{A}{N} t}}.$$

Observe that $\tau'(0) = t_e n$ and $\tau'(\infty) = 0$.

Then,

$$\rho(t) = n - \frac{t_e N}{r_e} \ln \frac{A}{t_e N + r_e n \cdot e^{-\frac{A}{N} t}} - \frac{n \cdot A e^{-\frac{A}{N} t}}{t_e N + r_e n \cdot e^{-\frac{A}{N} t}}.$$

Observe that $\rho(0) = 0$ and $\rho(\infty) = n - \frac{t_e N}{r_e} \cdot \ln\left(1 + \frac{r_e n}{t_e N}\right)$. (Obviously we have $\tau(\infty) + \rho(\infty) = n$, i.e., all and only the people interested in the event will eventually tweet or retweet about it.) We also have,

$$\rho'(t) = \frac{t_e r_e n^2 \cdot A e^{-\frac{A}{N} t}\left(1 - e^{-\frac{A}{N} t}\right)}{\left(t_e N + r_e n \cdot e^{-\frac{A}{N} t}\right)^2}.$$

The limiting values of $\rho'(t)$ are then $\rho'(0) = \rho'(\infty) = 0$.

Finally, using the expressions we obtained for $\tau'(t)$ and $\rho'(t)$, and System (2), we can compute the explicit expressions for $T'(t)$ and $R'(t)$ that are in statement of the Lemma. □

We have fit our model to the goals of WORLDCUP. The fitting is quite convincing for most of the goals (see Figure 4). There are a handful of exceptions: for instance, when two goals are scored within two to three minutes of each other (Figure 9) the fitting procedure fails to produce a good fit for the parameters. Enhancing our model to accommodate closely occurring events is an interesting future direction.

## Conclusions

Twitter provide a powerful medium through which users can communicate their observations not only with their friends, but also with the world at large. This is especially true when a user is closely following an event online. We studied the problems of identifying key events in a tweet stream and

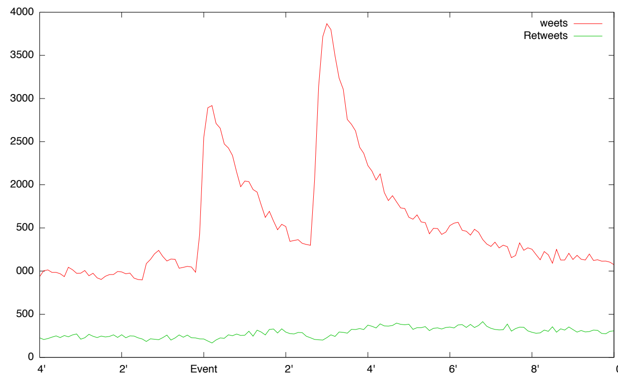


Figure 9: The two peaks in this figure correspond to the third and fourth goals that Germany scored against Australia in June 13, 2010.

| | |
|---|---|
| # tweets at the event time | $T'(0) = t_e n + t_g(N - n)$ |
| # tweets at time $\infty$ | $T'(\infty) = t_g N$ |
| # retweets at the event time | $R'(0) = r_g(N - n)$ |
| # retweets at time $\infty$ | $R'(\infty) = r_g N$ |
| # event tweets at the event time | $\tau'(0) = t_e n$ |
| # event tweets at time $\infty$ | $\tau'(\infty) = 0$ |
| # event retweets at the event time | $\rho'(0) = 0$ |
| # event retweets at time $\infty$ | $\rho'(\infty) = 0$ |
| total # event tweets | $\tau(\infty) = \frac{t_e N}{r_e} \ln\left(1 + \frac{r_e n}{t_e N}\right)$ |
| total # event retweets | $\rho(\infty) = n - \frac{t_e N}{r_e} \ln\left(1 + \frac{r_e n}{t_e N}\right)$ |

Table 2: Some properties of the solution to (1) under the boundary conditions $\tau(0) = \rho(0) = 0$.

the tweet production/consumption patterns around the key event. We observed, across many datasets, a surprising and robust "heartbeat" phenomenon: when a key event happens, the users become less social but quickly after the event, they get back to socializing, this time at a higher level. We explained this phenomenon with a natural model, and we used it to obtain a simple classifier that, by only looking at the tweet/retweet volume and without using any textual information, is able to detect key events more accurately than strong baseline methods.

## References

Allan, J.; Papka, R.; and Lavrenko, V. 1998. On-line new event detection and tracking. In *SIGIR*, 37–45.

Allan, J. 2002. Introduction to topic detection and tracking. In Allan, J., ed., *Topic Detection and Tracking – Event-based*

*Information Organization*. Kluwer Academic Publisher. 1–16.

Becker, H.; Naaman, M.; and Gravano, L. 2009. Event identification in social media. In *WebDB*.

Becker, H.; Naaman, M.; and Gravano, L. 2010. Learning similarity metrics for event identification in social media. In *WSDM*, 291–300.

Becker, H.; Naaman, M.; and Gravano, L. 2011a. Beyond trending topics: Real-world event identification in Twitter. In *ICWSM*, 438–441.

Becker, H.; Naaman, M.; and Gravano, L. 2011b. Selecting quality Twitter content for events. In *ICWSM*, 442–445.

boyd, d.; Golder, S.; and Lotan, G. 2010. Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In *HICSS*, 1–10.

Chakrabarti, D., and Punera, K. 2011. Event summarization using tweets. In *ICWSM*, 66–73.

Chen, L., and Roy, A. 2009. Event detection from Flickr data through wavelet-based spatial analysis. In *CIKM*, 523–532.

Conover, M.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Flammini, A.; and Menczer, F. 2011. Political polarization on Twitter. In *ICWSM*, 89–96.

Golder, S. A., and Yardi, S. 2010. Structural predictors of tie formation in Twitter: Transitivity and mutuality. In *SocialCom/PASSAT*, 88–95.

Huberman, B. A.; Romero, D. M.; and Wu, F. 2009. Social networks that matter: Twitter under the microscope. *First Monday* 14(1).

Java, A.; Song, X.; Finin, T.; and Tseng, B. 2007. Why we twitter: Understanding microblogging usage and communities. In *WebKDD/SNA-KDD*, 56–65.

Kleinberg, J. 2002. Bursty and hierarchical structure in streams. In *KDD*, 91–101.

Kleinberg, J. 2004. Temporal dynamics of on-line information streams. In Garofalakis, M.; Gehrke, J.; and Rastogi, R., eds., *Data Stream Management: Processing High-Speed Data Streams*. Springer.

Krishnamurthy, B.; Gill, P.; and Arlitt, M. 2008. A few chirps about Twitter. In *WOSP*, 19–24.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. B. 2010. What is Twitter, a social network or a news media? In *WWW*, 591–600.

Luo, G.; Tang, C.; and Yu, P. S. 2007. Resource-adaptive real-time new event detection. In *SIGMOD*, 497–508.

Petrovic, S.; Osborne, M.; and Lavrenko, V. 2010. Streaming first story detection with application to Twitter. In *HLT-NAACL*, 181–189.

Qu, Y.; Huang, C.; Zhang, P.; and Zhang, J. 2011. Microblogging after a major disaster in China: A case study of the 2010 Yushu earthquake. In *CSCW*, 25–34.

Romero, D. M., and Kleinberg, J. M. 2010. The directed closure process in hybrid social-information networks, with an analysis of link formation on Twitter. In *ICWSM*, 138–145.

Sakaki, T.; Okazaki, M.; and Matsuo, Y. 2010. Earthquake shakes Twitter users: Real-time event detection by social sensors. In *WWW*, 851–860.

Sankaranarayanan, J.; Samet, H.; Teitler, B. E.; Lieberman, M. D.; and Sperling, J. 2009. Twitterstand: News in tweets. In *GIS*, 42–51.

Shamma, D. A.; Kennedy, L.; and Churchill, E. 2009. Tweet the debates. In *WSM*.

Yardi, S., and boyd, d. 2010. Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society* 30.

Zhao, S.; Zhong, L.; Wickramasuriya, J.; and Vasudevan, V. 2011. Human as real-time sensors of social and physical events: A case study of Twitter and sports games. Technical Report 1106.4300, ArXiV.