

# Hip and Trendy: Characterizing Emerging Trends on Twitter

**Mor Naaman**

*School of Communication and Information, Rutgers University, 4 Huntington St., New Brunswick, NJ 08901.  
E-mail: mor@rutgers.edu*

**Hila Becker and Luis Gravano**

*Computer Science Department, Columbia University, 1214 Amsterdam Ave., New York, NY 10027.  
E-mail: {hila, gravano}@cs.columbia.edu*

Twitter, Facebook, and other related systems that we call *social awareness streams* are rapidly changing the information and communication dynamics of our society. These systems, where hundreds of millions of users share short messages in real time, expose the aggregate interests and attention of global and local communities. In particular, emerging temporal trends in these systems, especially those related to a single geographic area, are a significant and revealing source of information for, and about, a local community. This study makes two essential contributions for interpreting emerging temporal trends in these information systems. First, based on a large dataset of Twitter messages from one geographic area, we develop a taxonomy of the trends present in the data. Second, we identify important dimensions according to which trends can be categorized, as well as the key distinguishing features of trends that can be derived from their associated messages. We quantitatively examine the computed features for different categories of trends, and establish that significant differences can be detected across categories. Our study advances the understanding of trends on Twitter and other social awareness streams, which will enable powerful applications and activities, including user-driven real-time information services for local communities.

## Introduction

In recent years, a class of communication and information platforms we call *social awareness streams* (SAS) has been shifting the manner in which we consume and produce information. Available from social media services such as Facebook, Twitter, FriendFeed, and others, these hugely popular networks allow participants to post streams of lightweight content artifacts, from short status messages to links, pictures, and videos. These SAS platforms have already

shown considerable impact on the information, communication, and media infrastructure of our society (Johnson, 2009), as evidenced during major global events such as the Iran election or the reaction to the earthquake in Haiti (Kwak, Lee, Park, & Moon, 2010), as well as in response to local events and emergencies (Shklovski, Palen, & Sutton, 2008; Starbird, Palen, Hughes, & Vieweg, 2010).

SAS allow for rapid, immediate sharing of information aimed at known contacts or the general public. The content of the often-public shared items ranges from personal status updates to opinions and information sharing (Naaman, Boase, & Lai, 2010). In aggregate, however, the postings by hundreds of millions of users of Facebook, Twitter, and other systems expose global interests, happenings, and attitudes in almost real time (Kwak et al., 2010).

These interests and happenings as reflected in SAS data change rapidly. The strong temporal nature of SAS information allows for the detection of significant events and other temporal trends in the stream data. Such trends may reflect a varied set of occurrences, including local events (e.g., a baseball game or “fire on 34th street”), global news events (e.g., Michael Jackson’s death), televised events (e.g., the final episode of ABC’s *Lost*), Internet-only and platform-specific memes (e.g., a “fad” of users describing various things they object to using the #idonotsupport keyword), and hot topics of discussion (e.g., healthcare reform or the tween idol Justin Bieber).

Most related SAS research so far has focused on Twitter, due to its wide global reach and popularity, and because its contents are mostly public and are easily downloaded with automated tools. Several research efforts focused on characterizing or analyzing content from individual events on Twitter (Diakopoulos, Naaman, & Kivran-Swaine, 2010; Nagarajan, Gomadam, Sheth, Ranabahu, Mutharaju, & Jadhav, 2009; Sakaki, Okazaki, & Matsuo, 2010; Starbird et al., 2010; Shamma, Kennedy, & Churchill, 2010; Yardi & boyd, 2010). Other research efforts have addressed the problem of detecting and identifying trends in Twitter and other

---

Received July 30, 2010; revised December 20, 2010; accepted December 21, 2010

© 2011 ASIS&T • Published online 7 March 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21489

SAS data. “Bursts” of interest and attention can be detected in this data in hindsight (Becker, Naaman, & Gravano, 2010; Chen & Roy, 2009; Kleinberg, 2003; Rattenbury, Good, & Naaman, 2007) or in almost real time (Sakaki et al., 2010; Sankaranarayanan, Samet, Teitler, Lieberman, & Sperling, 2009). Most recently, some work has focused on characterizing aggregate general trend characteristics, for example, showing a power law distribution of participation for manually identified terms that correspond to events (Singh & Jain, 2010).

Indeed, SAS systems in general, and Twitter in particular, reflect an ever-updating live image of our society. However, the lack of a well-established structure and semantics for this data limits its utility. Our interest in this article is in characterizing the features that can help identify and differentiate the types of trends that we can find on Twitter. Better understanding of the semantics of SAS trends could provide critical information for systems that build on this emerging data. The outcome will be a more robust and nuanced reflection of emerging trends that captures key aspects of relevance and importance.

We focus here on content that is produced and shared within a specific geographic community and trends detected in that content. The relationship between geography and neighborhood and community has been long studied and argued (Campbell, 1990; Hampton & Wellman, 2003; Tilly, 1974), particularly in view of the Internet’s effect on local community ties (Hampton & Wellman, 2003; Putnam, 2000). It is clear, though, that social ties are still more likely between geographically proximate individuals (Mok, Carrasco, & Wellman, 2010; McPherson, Smith-Lovin, & Cook, 2001), and those patterns persist in online networks as well (Scellato, Mascolo, Musolesi, & Latora, 2010). On Twitter in particular, Scellato et al. (2010) and Takhteyev, Gruzd, and Wellman (2010) show that a significant proportion of the connections are local, although significant “global” patterns of connections exist. Beyond the higher likelihood of connections and ties, people living in the same geographic area are more similar (McPherson et al., 2001), and likely to share interests and information needs (Yardi & Boyd, 2010). Therefore, we posit that trends that appear in content produced by individuals in a geographic community can be critical and useful to detect or report to others in this community. On the other hand, this type of information can also become distracting and meaningless if these interests are not reported or harvested correctly. In this work, the focus on a specific geographic community helps us effectively reason about emerging trends with global and local impact.

This article offers the following contributions:

1. A taxonomy of trends that can be detected from Twitter for a specific geographic community using popular, widely accepted methods.
2. A characterization of the data associated with each trend along a number of key characteristics, including social network features, time signatures, and textual features.

This improved understanding of emerging information on Twitter in particular, and in SAS in general, will allow researchers to design and create new tools to enhance the use of SAS as information systems in different contexts and applications, including the filtering, search, and visualization of real-time SAS information as it pertains to local geographic communities.

To this end, we begin with an introduction to Twitter and a review of related efforts and background to this work. We then formally describe our dataset of Twitter trends and their associated messages. Later, we describe a qualitative study exposing the types of trends found on Twitter. Finally, in the bulk of this article we identify and analyze emerging trends using the unique social, temporal, and textual characteristics of each trend that can be automatically computed from Twitter content.

## Twitter

Twitter is a popular SAS service, with tens of millions of registered users as of June 2010. Twitter’s core function allows users to post short messages, or *tweets*, which are up to 140 characters long. Twitter supports posting (and consumption) of messages in a number of different ways, including through Web services and “third party” applications. Importantly, a large fraction of the Twitter messages are posted from mobile devices and services, such as Short Message Service (SMS) messages. A user’s messages are displayed as a “stream” on the user’s Twitter page.

In terms of social connectivity, Twitter allows a user to follow any number of other users. The Twitter contact network is directed: user A can follow user B without requiring approval or a reciprocal connection from user B. Users can set their privacy preferences so that their updates are available only to each user’s followers. By default, the posted messages are available to anyone. In this work, we only consider messages posted publicly on Twitter. Users consume messages mostly by viewing a core page showing a stream of the latest messages from people they follow, listed in reverse chronological order.

The conversational aspects of Twitter play a role in our analysis of the Twitter temporal trends. Twitter allows several ways for users to directly converse and interact by referencing each other in messages using the @ symbol. A *retweet* is a message from one user that is “forwarded” by a second user to the second user’s followers, commonly using the “RT @username” text as prefix to credit the original (or previous) poster (e.g., “RT @justinbieber Tomorrow morning watch me on the today show”). A *reply* is a public message from one user that is a response to another user’s message, and is identified by the fact that it starts with the replied-to user @username (e.g., “@mashable check out our new study on Twitter trends”). Finally, a *mention* is a message that includes some other username in the text of the message (e.g., “attending a talk by @informor”). Twitter allows users to easily see all recent messages in which they were retweeted, replied to, or mentioned.

Finally, Twitter supports a hashtag annotation format so that users can indicate what their posted messages are about. This general “topic” of a tweet is, by convention, indicated with the hash sign, #. For example, #iranelections was a popular hashtag with users posting about the Iran election events.

## Related Work and Background

The general topic of studying Twitter trends, as well as Twitter content related to real-life events, has recently received considerable research interest. Research efforts often examined a small number of such trends to produce some descriptive and comparative characteristics of Twitter trends or popular terms. Cheong and Lee (2009) looked at four trending topics and two control terms, and a subset of the messages associated with each, commenting on features such as the time-based frequency (volume of messages) for each term and the category of users and type of devices used to post the associated messages. Yardi and boyd (2010) examined the characteristics of content related to three topics on Twitter, two topics representing geographically local news events and one control topic. The authors studied the messages posted for each topic (i.e., messages containing terms manually selected by the authors to capture related content) and the users who posted them. Among other findings, the authors suggest that local topics feature denser social connectivity between the posting users. Similarly, Sakaki et al. (2010) suggest that the social connectivity for breaking events is lower, but have only examined content related to two manually chosen events. Singh and Jain (2010) examine Twitter messages with select hashtags and show that the content for each such set follows a power-law distribution in terms of popularity, time, and geo-location. Kwak et al. (2010) show that different trending terms on Twitter have different characteristics in terms of the number of replies, mentions, retweets, and “regular” tweets that appear in the set of tweets for each term, but do not reason about why and how exactly these trends are different. Some of the metrics we use here for characterizing trends are similar to those used in these studies, but we go further and perform a large-scale analysis of trends according to manual assignments of these trends to distinct categories.

On a slightly larger scale, Kwak et al. (2010) also examined the time series volume data of tweets for each trending term in their dataset, namely, a sample of 4,000 of the trending terms computed and published by Twitter. The authors based their analysis on the findings of Crane and Sornette (2008), which analyzed time series viewing data for individual YouTube videos. Crane and Sornette observed that YouTube videos fall into two categories, based on their view patterns. When a time series shows an immediate and fast rise in a video’s views, Crane and Sornette assert that the rise is likely caused by external factors (i.e., attention was drawn to the video from outside the YouTube community) and, therefore, dub this category of videos “exogenous.” When there is no such rise, the authors suggest that a video’s popularity

is due to “endogenous” factors. Videos are also classified as “critical” or “sub critical,” again according to the time series data. Kwak et al. (2010) use these guidelines to classify the Twitter trends in each of these two categories, showing how many trends fit each type of time-series signature. However, the two groups of authors never verified that the trends or videos labeled as exogenous or endogenous indeed matched their labels. Here we use the time series data (among other characteristics) while manually coding identified trends as exogenous or endogenous in order to observe whether these categories show different time series effects.

While trend and event detection in news and blog posts has been studied in depth (Allan, 2002; Kleinberg, 2003; Sayyadi, Hurst, & Maykov, 2009), the detection of trends on Twitter is a topic that is still in its infancy (Petrovic, Osborne, & Lavrenko, 2010; Sakaki et al., 2010). For example, Sankaranarayanan et al. (2009) use clustering methods to identify trending topics—corresponding to news events—and their associated messages on Twitter. Looking at social text stream data from blogs and email messages, Zhao, Mitra, and Chen (2007) detect events using textual, social, and temporal document characteristics in the context of clustering with temporal segmentation and information flow-based graph cuts. Other research considers event and trend detection in other social media data, such as Flickr photographs (Becker et al., 2010; Chen & Roy, 2009; Rattenbury et al., 2007).

The related problem of information dissemination has also attracted substantial attention. As a notable example, recent work studies the diffusion of information in news and blogs (Gruhl, Guha, Liben-Nowell, & Tomkins, 2004; Leskovec, Backstrom, & Kleinberg, 2009). As another example, Jansen, Zhang, Sobel, and Chowdury (2009) study word-of-mouth activity around brands on Twitter. Trends identified in the Twitter data are, of course, both products and generators of information dissemination processes.

Several recent efforts attempt to provide analytics for trends and events detected or tracked on Twitter. Sakaki et al. (2010) study social, spatial, and temporal characteristics of earthquake-related tweets, and De Longueville, Smith, and Luraschi (2009) describe a method for using Twitter to track forest fires and the response to the fires by Twitter users. Starbird et al. (2010) described the temporal distribution, sources of information, and locations in tweets from the Red River Valley floods of April 2009. Nagarajan et al. (2009) downloaded Twitter data for three events over time and analyzed the topical, geographic, and temporal importance of descriptors (e.g., different keywords) that can help visualize the event data. Finally, Shamma et al. (2010), Diakopoulos and Shamma (2010), and Diakopoulos, Naaman, and Kivran-Swaine (2010) analyze the tweets corresponding to large-scale media events (e.g., the United States President’s annual State of the Union speech) to improve event reasoning, visualization, and analytics. These tasks may all be improved or better automated with the enhanced understanding of the Twitter trends that is the result of the work presented here.



FIG. 1. Trending terms, on the dark blue (middle) banner, on Twitter's home page.

## Trends on Twitter

Because of the quick and transient nature of its user posts, Twitter is an information system that provides a “real time” reflection of the interests and thoughts of its users, as well as their attention. As a consequence, Twitter serves as a rich source for exploring the mass attention of millions of its users, reflected in “trends” that can be extracted from the site.

For the purposes of this work, a *trend* on Twitter (sometimes referred to as a *trending topic*) consists of one or more terms and a time period, such that the volume of messages posted for the terms in the time period exceeds some expected level of activity (e.g., in relation to another time period or to other terms). According to this definition, trends on Twitter include our examples above, such as Michael Jackson’s death (with terms “Michael” and “Jackson,” and time period June 25, 2009), the final episode of *Lost* (with terms “Lost” and “finale,” and time period May 23, 2010), and the healthcare reform debate (with term “HCR” and time period May 25, 2010). This definition conveys the observation by Kleinberg (2003) that the “appearance of a topic in a document stream is signaled by a burst of activity, with certain features rising sharply in frequency as the topic emerges” but does not enforce novelty (i.e., a requirement that the topic was not previously seen). In Twitter’s own (very informal) definition, trends “are keywords that happen to be popping up in a whole bunch of tweets.” Figure 1 captures Twitter’s home page with several trending topics displayed at the top.

In this article, each trend  $t$  is then identified by a set  $R_t$  of one or more terms and a time period  $p_t$ . For example, Figure 1 highlights one trend  $t$  that is identified by a single

term, iOS4 (referring to the release of Apple’s mobile operating system). To analyze a trend  $t$ , we study the set  $M_t$  of associated messages during the time period that contain the trend terms (in our example, all messages with the string “iOS4”). Note that, of course, alternative definitions and representations of trends are possible (e.g., based on message clustering; Sankaranarayanan et al., 2009). However, for this work we decided to concentrate on the above term-based formulation, which reflects a model commonly used in other systems (e.g., by Twitter as well as other commercial engines such as OneRiot).

While detecting trends is an interesting research problem, we focus here instead on characterizing the trends that can be detected on Twitter with existing baseline approaches. For this, we collect detected trends from two different sources. First, we collect local trends identified and published hourly by Twitter; the trends are available via an application programmer interface (API) from the Twitter service. Second, to complement and expand the Twitter-provided trends, we run a simple burst-detection algorithm over a large Twitter dataset to identify additional trends. We describe these two trend-collection methods next.

## Collecting Trend Data

In this section we describe the two methods we use to compile trends on Twitter, and also how we select the set of trends for analysis and how we get the associated messages, or tweets, for each trend. The set of trends  $T$  that we will analyze in this article consists of the union of the trends compiled

using both methods below. We use two methods in order to control, at least to some degree, for bias in the type of trends that may be detected by one system, but not another. While other algorithms for trend detection exist, we strongly believe our selected methods will provide a representative sample of the type of trends that can be detected. The set of detected trends might be skewed towards some trend types in comparison to other methods, but this skewness does not affect the analysis in this work. We further address this issue in the limitations discussion below.

In subsequent sections we qualitatively examine a subset  $T_{Qual}$  of the trends in  $T$  to extract the key types of trends that are present in Twitter data and develop a set of dimensions according to which trends can be categorized. We then use the categories to compare the trends in (a different) subset of  $T$ ,  $T_{Quant}$ , according to several features computed from the data associated with each trend, such as the time dynamics of each trend and the interaction between users in the trend's tweets. We examine whether trends from different categories show a significant difference in their computed features.

#### *Tweets Dataset*

The “base” dataset used for our study consists of over 48,000,000 Twitter messages posted by New York City users of Twitter between September 2009 and March 2010. This dataset is used in one of our methods described below to detect trends on Twitter (i.e., to generate part of our trend set  $T$ ). The dataset is also used for identifying the set of tweets  $M_t$  for each trend  $t$  in our trend set  $T$ . (Recall that  $T$  consists of all the trends that we analyze, compiled using both methods discussed below.) We collected the tweets via a script for querying the Twitter API. We used a “whitelisted” server, allowed to make a larger number of API calls per day than the default quota, to continuously query the Twitter API for the most recent messages posted by New York City users (i.e., by Twitter users whose location, as entered by the users and shown on their profile, is in the New York City area). This querying method results in a highly significant set of tweets, but it is only a subsample of the posted content. First, we do not get content from New York users who did not identify their home location. Second, the Twitter search API returns a subsample of matching content for most queries. Still, we collected over 48,000,000 messages from more than 855,000 unique users.

For each tweet in our dataset, we record its textual content, the associated timestamp (i.e., the time at which the tweet was published), and the user ID of the user who published the tweet.

#### *Trend Dataset I: Collecting Twitter's Local Trending Terms*

As mentioned above, one of our trend datasets consists of the trends computed by, and made available from, the Twitter service. Twitter computes these trends hourly, using an unpublished method. This source of trend data is

commonly used in research efforts related to trends on Twitter (e.g., Kwak et al., 2010; Cheong & Lee, 2009).

The Twitter-provided trends are computed for various geographic scales and regions. For example, Twitter computes and publishes the trends for New York City, as well as for the United States, and across all the Twitter service (e.g., those shown in Figure 1). From the data, we can observe that location-based trends are not necessarily disjoint: for example, New York City trends can reflect national trends or overlap with other cities' trends.

We collected over 8,500 trends published by Twitter for the New York City area during the months of February and March of 2010. The data included the one or two terms associated with each published trend, as well as the trend's associated time period, expressed as a date and time of day. We use the notation  $T_{Tw}$  (for “Twitter”) to denote this set of trends.

#### *Trend Dataset II: Collecting Trends With Burst Detection*

We derived the second trend dataset using a simple trend-detection mechanism over our Tweets dataset described above. This simple approach is similar to those used in other efforts (Nagarajan et al., 2009) and, as noted by Phelan, McCarthy, and Smyth (2009), it “does serve to provide a straightforward and justifiable starting point.” The trend-detection mechanism relies conceptually on the TF-IDF score (Salton, 1983) of terms, highlighting terms that appear in a certain time period much more frequently than expected for that time of day and day of the week. We tune this approach so that it does not assign a high score to weekly recurring events, even if they are quite popular, to ensure that we include a substantial fraction of trends that represent “one-time,” nonrecurring events, adding to the diversity of our analysis.

Specifically, to identify terms that appear more frequently than expected, we will assign a score to terms according to their deviation from an expected frequency. Assume that  $M$  is the set of all messages in our Tweets dataset,  $R$  is a set of one or more terms to which we wish to assign a score, and  $h$ ,  $d$ , and  $w$  represent an hour of the day, a day of the week, and a week, respectively. We then define  $M(R, h, d, w)$  as the set of every Twitter message in  $M$  such that (1) the message contains all the terms in  $R$  and (2) the message was posted during hour  $h$ , day  $d$ , and week  $w$ . With this information, we can compare the volume in a specific day/hour in a given week to the same day/hour in other weeks (e.g., 10 AM on Monday, March 15, 2010, vs. the activity for other Mondays at 10 AM).

To define how we score terms precisely, let  $Mean(R, h, d) = (\sum_{i=1, \dots, n} |M(R, h, d, w_i)|) / n$  be the number of messages with the terms in  $R$  posted each week on hour  $h$  and day  $d$ , averaged over the weeks  $w_1$  through  $w_n$  covered by the Tweets dataset. Correspondingly,  $SD(R, h, d)$  is the standard deviation of the number of messages with the terms in  $R$  posted each week on day  $d$  and hour  $h$ , over all the weeks. Then, the score of a set of terms  $R$  over a specific

TABLE 1. Summary of event datasets.

Notation	Data source	Selection for analysis
$T_{tw}$	Twitter's own trends as retrieved from the Twitter API	Selected from complete set of trends published by Twitter
$T_{tf}$	Trends computed from raw Twitter data using term frequency measures	Selected from top-scoring terms for each day

TABLE 2. Sample trends and their explanation.

Trend terms	Explanation	Date	# Tweets
#TEDxNYED	A New York City conference on media, technology, and education	March 6, 2010	556
Sparklehorse	The suicide of Mark Linkous, of the band Sparklehorse	March 8, 2010	230
Burger	Reaction to a tweet by Lady Gaga: "once you kill a cow, you gotta make a burger"	March 12, 2010	3,249
Masters	Tiger Woods's announcement of his return to golf at the Masters	March 16, 2010	693
itsreallyannoying	Twitter meme: users sharing their annoyances	March 23, 2010	2,707
Seder	Passover-eve meal	March 28, 2010	316
iPad	Launch of the Apple iPad	March 29, 2010	1,714

hour  $h$ , day  $d$ , and week  $w$  is defined as  $score(R, h, d, w) = (|M(R, h, d, w)| - Mean(R, h, d))/SD(R, h, d)$ .

Using this definition, we computed the score for every individual term in our dataset (in other words, we computed the scores for all  $R$  sets where each  $R$  is a set with a single 1-gram that appears in  $M$ ). We computed the score for each  $R$  over all  $h$ ,  $d$ , and  $w$  values for the weeks covered by our Tweets dataset. For each day  $d$  and week  $w$ , we identified the  $R$  and  $h$  pairs such that (1)  $M(R, h, d, w)$  contains at least 100 messages and (2) the  $score(R, h, d, w)$  value is among the top-30 scores for day  $d$  and week  $w$  across all term-hour pairs. Each selected pair defines a trend with set of terms  $R$  and associated time period specified by  $h$ ,  $d$ , and  $w$ . (Note that certain terms could be repeated if they scored highly for multiple hours in the same day; such repetition is also possible for the trend set  $T_{tw}$ . We compute a trend's "real" peak after we choose the trends for analysis, as described below.) We use the notation  $T_{tf}$  (for "term frequency") to denote the resulting set of 1,500 trends.

For reference, the sources and properties of the event datasets are summarized in Table 1.

### Selecting Trends for Analysis

After identifying the above two sets of trends, namely,  $T_{tw}$  and  $T_{tf}$ , our goal is to perform both a quantitative and a qualitative analysis of these trends. To be meaningful, this analysis will rely on a manual coding of the trends, but an exhaustive manual processing of all trends in  $T_{tw}$  and  $T_{tf}$  would, unfortunately, be prohibitively expensive. Therefore, our analysis will focus on a carefully selected subset of the two trend sets (see the Trend Taxonomy and Dimensions section). This selection of trends should (1) reflect the diversity of trends in the original sets and (2) include only trends that could be interpreted and understood by a human, through inspection of the associated Twitter messages.

For both sets  $T_{tw}$  and  $T_{tf}$ , one of the authors performed a random selection of trends to serve as an initial dataset. For each trend in this initial selection we attempted to identify the topic reflected in the trend by inspecting associated

messages (posted on the corresponding day, and with the corresponding terms). If we could not identify the topic or reason for the trend, we removed it from the selected set to satisfy condition (2). In addition, after the first round of coding trends according to the categories described below, we manually inspected the trends from the initial sets  $T_{tw}$  and  $T_{tf}$  that were not yet selected for analysis. Instead of randomly choosing among them, we randomly chose a date and then purposefully selected additional trends from that date from underrepresented categories, satisfying condition (1). Note that we attempted to create a comprehensive, but not necessarily proportional, sample of trends in the data. In other words, some types of trends may be over- or underrepresented in the selected trends dataset. At the same time, the sample of trends in each category is representative of trends in the category overall. Our aim here is to provide insight about the categories of trends and features of trends in each category, rather than discuss the magnitude of each category in the data, a figure likely to shift, for example, with changes to the detection algorithms.

The result of this process was a set of trends  $T$  that combines trends from both  $T_{tw}$  and  $T_{tf}$ . We split the set  $T$  into two subsets. The first subset of selected trends,  $T_{Qual}$ , consisting of trends in  $T$  through February 2010, was used for the qualitative analysis described next. The second subset of  $T$ ,  $T_{Quant}$ , consisting of trends in  $T$  from March 2010, was used for the quantitative analysis described below. Table 2 lists several of the trends selected for the analysis: for each trend  $t$ , we list its description, time period, and number of associated messages (i.e., the cardinality of  $M_t$ ). Next, we explain how we identify  $M_t$  for each trend  $t$ .

Table 3 provides a summary of the datasets described in this section, along with their respective size.

### Identifying Tweets Associated with Trends

For our statistical analysis of trend features, for each trend  $t$  in  $T_{Quant}$  we need to know the set of tweets  $M_t$  associated with  $t$ . Each trend includes the terms that identify the trend

TABLE 3. Details of the trend datasets produced and used in this work.

Initial set of Twitter trends ( $T_{tw}$ )	>8,500*
Initial set of burst detection trends ( $T_{bf}$ )	>1,500*
Selected trends for qualitative analysis ( $T_{Qual}$ )	50
Selected trends for quantitative analysis ( $T_{Quant}$ )	200

\*Including duplicate trending terms in different hours.

and the associated time period, as discussed (e.g., a trend might consist of term “Passover” on March 29, 2010, for the hour starting at 4 PM). To define  $M_t$ , we first collect every message in our Tweets dataset that contains all of  $t$ ’s terms and such that it was posted up to 10 days before or after the time period for  $t$ . We sort these messages according to the time at which they were posted and we aggregate them into hourly bins. Since the identifying term(s) may be popular at various times (e.g., as is the case for a trend that persists for several hours), we identify the peak time for the trend by selecting the bin with the largest number of messages. Finally, after anchoring the trend in its new associated time period, we retrieve all messages posted up to 72 hours before or after the new time period; this set is  $M_t$ , the set of messages associated with trend  $t$ . On average, the set  $M_t$  for each trend in  $T_{Quant}$  consists of 1,350 tweets, and the median cardinality of  $M_t$  is 573.

Note again that other methods exist for trend detection that may associate content with trends not only by simple term matching as we do here (Becker et al., 2010). However, most current systems rely on term matching to identify related content. In addition, many of the characteristics we extract for each trend’s content would directly apply, or apply with minor changes, to sets of content collected via other methods. We further discuss this issue as part of our limitations below.

### Trend Taxonomy and Dimensions

We now describe the qualitative analysis that we performed to characterize the Twitter trends in the  $T_{Qual}$  set of trends described above. The analysis was geared to identify the different types of trends that occur in Twitter data from one metropolitan area and relies on a taxonomy of the trends.

Many Twitter trends correspond to events that are reflected on Twitter by its users. The definition and characterization of “event” has received substantial attention across academic fields, from philosophy (“Events,” 2002) to cognitive psychology (Zacks and Tversky, 2001). Media events have been characterized by Dayan and Katz (1992) into three generic types of scripts that these events tend to follow, namely, “contest,” “conquest,” and “coronation,” for events such as a presidential debate, an unfolding visit by a leader to a foreign state, and a leader’s funeral, respectively. Boll and Westermann (2003) present discussion of events in the area of personal multimedia collections. In information retrieval, the concept of events has prominently been addressed in the area of topic detection in news events (Allan, 2002; Kleinberg, 2003; Yang, Pierce, & Carbonell, 1998). To summarize, the aforementioned research from multiple disciplines is closely related to our work. However, the taxonomies available in these literatures do not capture the variety of trends that

emerge in a social information system such as Twitter, which is our focus here.

Our qualitative analysis of trends is based on a variation of the affinity diagram method, an inductive process (LeCompte & Schensul, 1999) to extract themes and patterns from qualitative data. For this analysis we used sticky notes to represent each trend in  $T_{Qual}$  and recorded the terms and the explanation of the trend if needed, which happened when the terms associated with the trend did not immediately offer an idea of the content. Two of the authors of this article then put together the different items into groups and categories in an iterative process of comparing, contrasting, integrating, and dividing the grouped trends. According to the affinity process, we considered the relationship between categories as well as the items that are grouped and linked together.

Indeed, the categories that emerged could be described and differentiated according to one key dimension: whether the trends in the category are exogenous or endogenous. Trends in exogenous categories capture an activity, interest, or event that originated outside of the Twitter system (e.g., an earthquake). Trends in endogenous categories are Twitter-only activities that do not correspond to external events (e.g., a popular post by a celebrity). Having this dimension at the top level of the taxonomy reflects and highlights the substantial differences on Twitter between exogenous and endogenous trends regarding their importance and use scenarios. The top level of the taxonomy thus separates nonvirtual external events from activities that only pertain to the Twitter system.

The groups of trends that emerged are described below, with sample trends to illustrate each category.

#### Exogenous Trends

- Broadcast-media events:
  - Broadcast of local media events: “fight” (boxing event), “Ravens” (football game).
  - Broadcast of global/national media events: “Kanye” (Kanye West acts up at the MTV Video Music Awards), “Lost Finale” (series finale of Lost).
- Global news events:
  - Breaking news events: “earthquake” (Chile earthquake), “Tsunami” (Hawaii Tsunami warning), “Beyoncé” (Beyoncé cancels Malaysia concert).
  - Nonbreaking news events: “HCR” (health care reform), “Tiger” (Tiger Woods apologizes), “iPad” (toward the launch of Apple’s popular device).
- National holidays and memorial days: “Halloween,” “Valentine’s.”
- Local participatory and physical events:
  - Planned events: “marathon,” “superbowl” (Super Bowl viewing parties), “patrick’s” (St. Patrick’s Day Parade).
  - Unplanned events: “rainy,” “snow.”

#### Endogenous Trends

- Memes: #in2010 (in December 2009, users imagine their near future), “November” (users marking the beginning of the month on November 1).

- Retweets (users “forwarding” en masse a single tweet from a popular user): “determination” (users retweeting LL Cool J’s post about said concept).
- Fan community activities: “2pac” (the anniversary of the death of hip-hop artist Tupac Shakur).

Needless to say, the above set of categories might not be comprehensive (i.e., other trends that are not in our data might not comfortably fit in any of these categories). However, we developed this set of categories after an exhaustive, thorough analysis of a large-scale set of trends, as described above. Therefore, we believe that this categorization is both sufficiently broad and, at the same time, simple enough to enable a meaningful study of the “trends in trend data.”

Our quantitative analysis below focuses on a limited number of dimensions extracted from the taxonomy that capture key differences between trends. We identified the dimensions to focus on according to two criteria: (a) significance, or the importance of being able to extract differences between the selected trend categories, and (b) the likelihood that these categories will result in measurable differences between trends.

The first dimension we examine is the high-level exogenous and endogenous categories of trends. Such comparison will allow us to reason about this most distinguishing aspect of any Twitter trend.

Within exogenous trends, in this work we chose to concentrate on two important dimensions. First, whether the exogenous activity falls into the *local participatory and physical events* category above. These “local trends” represent physical events, located in one geographic area (e.g., the New York marathon) that are currently underrepresented in the detected trends, but naturally play an important role in local communities. The second dimension chosen is whether the exogenous trends are *breaking news events*, global news events that are surprising and have not been anticipated (e.g., an earthquake), as opposed to all other events and trends that are planned or expected (e.g., a vote in the Senate, or a holiday). This dimension will allow us to separate “news-worthy” versus “discussion-worthy” trends, which may lead to a different manner in which systems use and display these different trend types.

Similarly, within endogenous trends in this work we chose to investigate the differences between trends in the two main categories of this group of events, namely, *memes* and *retweets*, as explained above.

Next, these dimensions help us guide the quantitative study of the trends detected in Twitter data, as we label each trend according to categories derived from the dimensions above.

## Characterizing Trends

The next step in our analysis is to characterize each Twitter trend using features of its associated messages. These features are later used to reason about differences between the various trend dimensions described in the previous section. For this analysis, we use the trend set  $T_{Quant}$  defined above. For each trend  $t$  we compute features automatically,

based on its associated set of tweets  $M_t$ . These features range from aggregate statistics of the content of each individual message (e.g., number of hashtags, URLs) to social network connections between the authors of the messages in  $M_t$  and the temporal characteristics of  $M_t$ .

### Content Features

Our first set of features (Table 4a) provides descriptive characteristics for a trend  $t$  based on the content of the messages in  $M_t$ . These features include aggregate characteristics such as the average length of a message in  $M_t$  and the percentage of messages with URLs or hashtags, or measures of the textual similarity of the tweets in  $M_t$ .

### Interaction Features

The interaction features (Table 4b) capture the interaction between users in a trend’s messages as indicated on Twitter by the use of the @ symbol followed by a username. These interactions have somewhat different semantics on Twitter, and include “retweets” (forwarding information), replies (conversation), or mentions of other users.

### Time-Based Features

The time-based features (Table 4c) capture different temporal patterns of information spread that might vary across trends. To capture these features for a trend  $t$ , we fit a family of functions to the histogram describing the number of Twitter messages associated with the trend over the time period spanned by the tweet set  $M_t$  (by construction, as discussed,  $M_t$  has the matching messages produced up to 72 hours before and after  $t$ ’s peak). We aggregate all messages in  $M_t$  into hourly bins. We refer to all bins before the peak as the head of the time period, while all bins after the peak are the tail of the time period.

We proceed to fit the bin volume data, for both the head and the tail of the time period, separately, to exponential and logarithmic functions. Using the least squares method, we compute logarithmic and exponential fit parameters for the head and tail periods for each trend, considering the full time period of 72 hours, which we refer to as the *Log72* fit and the *Exp72* fit, respectively. We proceed in the same manner for a limited time period of 8 hours before and after the peak, which we refer to as the *Log8* fit and the *Exp8* fit, respectively. The focus on the shorter time periods will allow us to better match rapidly rising or declining trends (Leskovec et al., 2009).

In sum, our features for each trend thus include the fit parameters for 8-hour and 72-hour spans for both the head and the tail periods; and for each period and span we calculate the logarithmic and exponential fit parameters. In addition, for each combination we also computed the  $R^2$  statistic, which measures the quality of each fit.

### Participation Features

Trends can have different patterns of participation, in terms of authorship of messages related to the trend. The



TABLE 4. Features for a trend  $t$ . As usual,  $M_t$  denotes the set of tweets associated with  $t$ .

	Explanation
<b>(a) Content features</b>	
Average number of words/characters	Let $words(m)$ be the number of words in a tweet $m$ and let $char(m)$ be the number of characters in tweet $m$ . Then the average number of words per message is $\frac{\sum_{m \in M_t} words(m)}{ M_t }$ , and the average number of characters per message is $\frac{\sum_{m \in M_t} char(m)}{ M_t }$ .
Proportion of messages with URLs	Let $U_t \subseteq M_t$ be the set of messages with URLs out of all messages for trend $t$ . Then the proportion of messages with URLs is $ U_t / M_t $ .
Proportion of unique URLs	Let $URL(m)$ be the set of URLs that appear in tweet $m$ . The set of unique URLs for $t$ is $ UU_t $ , where $UU_t = \{u: u \in URL(m) \text{ for a message } m \in M_t\}$ , and the proportion of unique URLs is $ UU_t / M_t $ (Note that the set semantics ensures that each unique URL is only counted once.)
Proportion of messages with hashtags	Let $H_t \subseteq M_t$ be the set of messages with hashtags in $M_t$ . Then the proportion of messages with hashtags is $ H_t / M_t $ .
Proportion of messages with hashtags, excluding trend terms	Let $H'_t \subseteq M_t$ be the set of messages with hashtags in $M_t$ , excluding messages where the hashtag is a term in $R_t$ , the set of terms associated with trend $t$ . Then the proportion of messages with hashtags excluding the trend's terms is $ H'_t / M_t $ .
Top unique hashtag?	Whether there is at least one hashtag that appears in at least 10% of the messages in $M_t$ . This measure captures agreement on the terms most topically related to the trend.
Similarity to centroid	We represent each message $m \in M_t$ as a TF-IDF vector (Salton, 1983), where the IDF value is computed with respect to all messages in the <i>Tweets</i> dataset. We compute the average TF-IDF score for each term across all messages in $M_t$ to define the centroid $C_t$ . Using $C_t$ , we then compute the average cosine similarity (Salton, 1983) $\frac{\sum_{m \in M_t} sim(C_t, m)}{ M_t }$ as well as the corresponding standard deviation. These features help indicate content cohesiveness within a trend.
<b>(b) Interaction features</b>	
Proportion of retweets	Let $RT_t \subseteq M_t$ be the set of messages in $M_t$ that are “retweets” (i.e., these messages include a string of the form “RT @user”). Then the proportion of retweets is $ RT_t / M_t $ .
Proportion of replies	Let $RP_t \subseteq M_t$ be the set of messages in $M_t$ that are “replies” (i.e., these messages begin with a string of the form “@user”). Then the proportion of replies is $ RP_t / M_t $ .
Proportion of mentions	Let $MN_t \subseteq M_t$ be the set of messages in $M_t$ that are “mentions” (i.e., these messages include a string of the form “@user” but are <i>not</i> replies or retweets as defined above). Then the proportion of mentions is $ MN_t / M_t $ .
<b>(c) Time-based features</b>	
Exponential fit (head)	Best fit parameters $(p_0, p_1, p_2)$ and goodness of fit $R^2$ for function $M(h) = p_1 e^{-p_0 h } + p_2$ , where $M(h)$ represents the volume of messages during the $h$ -th hour before the peak. Computed for 72- and 8-hour periods before the peak.
Logarithmic fit (head)	Best fit parameters $(p_0, p_1)$ and goodness of fit $R^2$ for function $M(h) = p_0 \log(h) + p_1$ , where $M(h)$ represents the volume of messages during the $h$ -th hour before the peak. Computed for 72- and 8-hour periods before the peak.
Exponential fit (tail)	Similar to above, but over 72- and 8-hour periods <i>after</i> the trend's peak.
Logarithmic fit (tail)	Similar to above, but over 72- and 8-hour periods <i>after</i> the trend's peak.
<b>(d) Participation features</b>	
Messages per author	Let $A_t = \{a: a \text{ is an author of a message } m \in M_t\}$ . Then the number of messages per author is $ M_t / A_t $ .
Proportion of messages from top author	We designate $a' \in A_t$ as the top author if $a'$ produced at least as many messages in $M_t$ as any other author. Then the proportion of messages from top author is $ \{m: m \in M_t \text{ and } m \text{ was posted by } a'\} / M_t $ .
Proportion of messages from top 10% of authors	Let $A10_t$ be the set of the top 10% of the authors in terms of the number of messages produced in $M_t$ . Then the proportion of messages from top 10% authors is $ \{m: m \in M_t \text{ and } m \text{ was posted by } a \in A10_t\} / M_t $ .
<b>(e) Social network features</b>	
Level of reciprocity	The fraction of reciprocal connections out of the total number of connections $ E_t $ , where authors $a_1, a_2 \in A_t$ form a reciprocal connection if $(a_1, a_2) \in E_t$ and $(a_2, a_1) \in E_t$ .
Maximal eigenvector centrality	The eigenvector centrality of an author measures the importance of this author in $A_t$ by computing the eigenvector of the largest eigenvalue in the adjacency matrix of the network graph. We pick the author with the highest eigenvector centrality value over all $a \in A_t$ . A high value suggests the existence of a dominant node in the network.
Maximal degree centrality	The degree centrality of an author $a \in A_t$ is the fraction of authors it is connected to. We compute the highest degree centrality value over all $a \in A_t$ . A high value suggests the existence of a dominant node in the network.

(Continued)

TABLE 4. (Continued)

	Explanation
Transitivity	Authors $a_1, a_2, a_3 \in A_t$ form a <i>triad</i> if one of them shares an edge with the other two authors. If each author shares an edge with both remaining authors, then they also form a <i>triangle</i> . The transitivity of the graph is computed as $3 \cdot \text{triangles}(G_t) / \text{triads}(G_t)$ , and is a measure of the connectivity of the authors in the network.
Density	The ratio of the number of edges $ E_t $ in the author social network, out of the total number of possible edges $ A_t  \cdot ( A_t  - 1)$ . This is an indication of how closely-knit the group of authors $A_t$ is.
Average component size	A strongly connected component of $G_t$ is a maximal set of authors $AC \subseteq A_t$ such that for each $a_1, a_2 \in AC$ there exists a directed path from $a_1$ to $a_2$ and vice versa. The average component size is the number of authors (i.e., nodes in the network) divided by the number of strongly connected components. This is a normalized measure of how “broken apart” the network of authors is.
Proportion of largest component	Let $AC'$ be the largest connected component. Then, $ AC' / A_t $ is a measure of how connected the network of authors is.

participation features (Table 4d) characterize a trend using statistics about the participation of authors that produced the trend’s associated messages; in particular, we capture the skew in participation (i.e., to which extent a small portion of authors produced most of the content).

#### Social Network Features

Our final group of features for a trend  $t$  focuses on the set  $A_t$  of the authors of the messages in  $M_t$ . Specifically, the social network features (Table 4e) capture the properties of the social network  $G_t$  of authors. To model this network, we used the Twitter API to collect the list of followers for each author, consisting of other Twitter users in  $A_t$  that subscribe to the author’s message feed. (We ignore followers that are not among the  $A_t$  authors. We also ignore followers of authors who restrict access to this information and those who have suspended Twitter accounts.) In other words, our social network graph is a directed graph  $G_t(A_t, E_t)$ , such that there exists an edge  $e \in E_t$  from  $a_1$  to  $a_2$  if and only if  $a_1$  is a follower of  $a_2$  on Twitter. We computed various features of the social network graph  $G_t$  for each trend  $t$ , capturing the connectivity and structure of connections in the graph (Wasserman and Faust, 1994).

#### Categorizing Trends in Different Dimensions

In addition to the automatically extracted features, we manually categorized the trends in  $T_{Quant}$  according to the dimensions picked for analysis (e.g., whether the trend belongs to the “exogenous” or “endogenous” category). We manually associated every trend with one category in each dimension. Later, we examined how the categories differ according to the automatically computed features described above.

We required a content description of each trend in order to properly label it according to the categories introduced in the previous section. The trend detection methods only output the trend terms and a time period. This type of output (e.g., “Bacall on March 8th”) was often not enough to discern the

content of the trend to correctly assign it to different categories. One of the authors examined each of the trends to generate a short description. The sources used for this examination were, first, the actual Twitter messages associated with the trend. If that examination did not prove informative enough, we used news search tools (e.g., Google News) to inspect corresponding news reports for that day and those terms. At the end of the process we had a description for 200 of the trends in our trend dataset  $T_{Quant}$ , after removing 29 trends that could not be resolved (e.g., “challenging” on March 14, 2010) from our dataset. We computed these features for the 200 resolved trends in our trend dataset  $T_{Quant}$ . This data is the basis for our analysis, described below.

We mapped each of the trends into categories based on the dimensions for analysis. Two of the authors independently annotated each of the trends. If an annotator could not assign a value for some dimension, either a “not applicable” or an “unknown” label was used. In each dimension, after removing trends marked “not applicable” or “unknown” by at least one of the annotators, the inter-annotator agreement of the labeled trends in each dimension was very high (the remaining number of trends for each dimension is reported below, in the analysis). For the final analysis in each dimension we removed all “not applicable” and “unknown” entries for that dimension, as well as any remaining disagreements between the annotators. In other words, we ignored those trends for which we had reason to doubt the assignment to a category.

#### Quantitative Analysis of Trends

The main drivers for our analysis are the coded categories of trends, as detailed above. In other words, we compared the samples of trends according to their categorization in different dimensions (e.g., exogenous vs. endogenous) and according to the features we computed from the data (e.g., the percentage of messages with URLs). Our hypotheses, listed below, are guided by intuitions about deviations in the characteristics of trends in different categories, and are geared towards confirming the expected deviations between the trend categories. Such confirmation would allow, later on, for the

development of automated systems to detect the trend type or provide better visualizations or presentation of the trend data. We continue by listing the key hypotheses that guided our analysis.

#### *Exogenous vs. Endogenous Trends*

**H1.** We hypothesize that exogenous and endogenous trends will have different quantitative characteristics. In particular:

- H1.1: Content features of exogenous trends will be different than those of endogenous trends; in particular, they will have a higher proportion of URLs and a smaller proportion of hashtags in tweets.
- H1.2: Interaction features of exogenous trends will be different than those of endogenous trends; in particular, exogenous trends will have fewer retweets (forwarding), and a similar number of replies (conversation).
- H1.3: Time features of exogenous trends will be different for the head period before the trend peak but will exhibit similar time features in the tail period after the trend peak, compared to endogenous trends.
- H1.4: Social network features of exogenous trends will be different than those of endogenous trends, with fewer connections (and less reciprocity) in the social network of the trend authors.

#### *Breaking News vs. Other Exogenous Trends*

**H2.** We hypothesize that breaking news events will have different quantitative characteristics compared to other exogenous trends. In particular:

- H2.1: Interaction features of breaking events will be different than those of other exogenous trends, with more retweets (forwarding), but fewer replies (conversation).
- H2.2: Time features of breaking events will be different for the head period, showing more rapid growth, and a better fit to the functions' curve (i.e., less noise) compared to other exogenous trends.
- H2.3: Social network features of breaking events will be different than those of other exogenous trends.

#### *Local Events vs. Other Exogenous Trends*

**H3.** We hypothesize that local participatory and physical events will have different quantitative characteristics compared to other exogenous trends. In particular:

- H3.1: Content features of local events will be different than those of other exogenous trends.
- H3.2: Interaction features of local events will be different than those of other exogenous trends; in particular, local events will have more replies (conversation).
- H3.3: Time features of local events will be different than those of other exogenous trends.
- H3.4: Social network features of local events will be different than those of other exogenous trends; in particular, local events will have denser networks, more connectivity, and higher reciprocity.

#### *Memes vs. Retweet Endogenous Trends*

**H4.** We hypothesize that memes will have different quantitative characteristics compared to retweet trends. In particular:

- H4.1: Content features of memes will be different than those of retweet trends.
- H4.2: Interaction features of memes will be different than those of retweet trends; in particular, retweet trends will have significantly more retweet (forwarding) messages (this hypothesis is included as a "sanity check" since the retweet trends are defined by having a large proportion of retweets).
- H4.3: Time features of memes will be different than those of retweet trends.
- H4.4: Participation features of memes will be different than those of retweet trends.
- H4.5: Social network features of memes will be different than those of retweet trends; in particular, meme trends will have more connectivity and higher reciprocity than retweet trends.

#### *Method*

We performed our analysis on the 200 resolved trends in  $T_{Quant}$ . The analysis was based on a pairwise comparison of trends according to the trends categorization in different dimensions, following our hypotheses above. For each such pair we performed a set of two-tailed  $t$ -tests to show whether there are differences between the two sets of trends in terms of the dependent variables, namely, our automatically extracted trend features. However, since each sub-hypothesis involved multiple dependent variables (e.g., we computed seven different social network features), we controlled for the multiple  $t$ -tests by using the Bonferroni correction, which asks for a significance level of  $\alpha/n$  when conducting  $n$  tests at once. We thus only report here results with significance level of  $p < 0.008$ .

As is common in studies of social-computing activities, many of our dependent variables were not normally distributed, but rather they were most often skewed to the right. Following Osborne (2002), we used logarithms (adding a small constant to handle zero values as needed) or square root functions to transform these variables in order to improve their normality. For most variables, such transformation indeed generated a normal distribution. In the cases where we performed a variable transformation, whenever we find significant differences between the transformed means in the analysis we also report here the original variable means and medians. For variables that were still skewed after the transformation, we performed the Mann–Whitney test for nonnormal distributions, and note when that is the case. For the one dependent variable in our data that was nominal, we used the CHI-square test.

Finally, following Asur and Huberman (2010), in the analysis we considered the temporal features only for trends that peaked on a US weekday (Monday through Friday), as the

TABLE 5. Summary of results.

	Categories compared				
	Content	Interaction	Time	Participation	Social
Exogenous vs. endogenous					
Hypothesis:	H1.1	H1.2	H1.3	None	H1.4
Found:	Yes* ✓	Yes* ✓	No ✗	No	Yes ✓
Breaking news vs. other exogenous					
Hypothesis:	None	H2.1	H2.2	None	H2.3
Found:	No	Yes ✓	No ✗	No	No ✗
Local vs. other exogenous					
Hypothesis:	H3.1	H3.2	H3.3	None	H3.4
Found:	No ✗	Yes* ✓	No ✗	No	No ✗
Memes vs. retweets (endogenous)					
Hypothesis:	H4.1	H4.2	H4.3	H4.4	H4.5
Found:	Yes ✓	Yes ✓	Yes* ✓	Yes ✓	Yes ✓

\*Starred entries represent partial findings or findings that diverged somewhat from the detailed hypothesis.

temporal aspects in particular might be influenced by the different patterns of Twitter usage during weekends.

## Results and Discussion

We report below the results from our analysis. For convenience, an overview of the results and findings as they related to the hypothesis is provided in Table 5.

### *Exogenous vs. Endogenous Trends*

Exogenous trends were found to be different than endogenous trends in content, interaction, and social features, supporting most of the hypotheses under H1 as shown in Table 5. In our dataset we had 115 exogenous trends and 55 endogenous trends (for some parts of the analysis the numbers are lower due to missing data). The detailed numerical results are shown in Table 6a,b. In terms of content features (H1.1), exogenous trends had a higher proportion of messages with URLs than endogenous trends (results were similar for the proportion of unique URLs appearing in the trend's content). In addition, the average term length for exogenous trends was somewhat shorter than the length of terms used in endogenous trends. We found only some differences in the presence of hashtags in the content: exogenous trends did not have a higher proportion of messages with hashtags, even when excluding the trending terms. However, fewer exogenous trends had a unique hashtag appearing in at least 10% of the messages compared to endogenous trends. This finding indicates less agreement between authors of exogenous trends on the ad-hoc "semantics" of the trend (in other words, the chosen community representation for what that trend content is about), which may stem from the fact that exogenous trends are seeded at once from many users who choose different hashtags to represent the trend.

In terms of interaction features (H1.2), we found that exogenous trends had a smaller proportion of retweets in the trend's tweets compared to endogenous trends. This finding

suggests that users created more original content based on exogenous sources, rather than retransmit and forward content that was already in the "system" as often happens for endogenous trends. Interestingly, we also found that exogenous trends tend to have more "conversation": the proportion of replies in exogenous trends was higher than endogenous ones.

In terms of time features (H1.3), the hypothesis was not supported: our data did not show exogenous trends to have different time features for the head period. The tail period time parameters were, as we hypothesized, not found to be different for exogenous and endogenous trends.

Finally, in terms of social network features (H1.4), we found differences in the level of reciprocity between exogenous and endogenous trends. Social network connections in exogenous trends had less reciprocity than those of endogenous trends. Other differences were found but with marginal significance.

### *Breaking News vs. Other Exogenous Trends*

Trends corresponding to breaking events were found to have different interaction characteristics from other exogenous trends, but no other differences were found, giving only partial support to hypothesis H2 (Table 5). In our dataset, we had 33 breaking events and 63 other exogenous events (for some parts of the analysis the numbers are lower due to missing data). The detailed numerical results are shown in Table 6c.

In terms of interaction features (H2.1), we found that breaking exogenous trends have a larger proportion of retweet messages than other exogenous events. Breaking trends also have a smaller proportion of reply messages than other exogenous events. These findings show the informational nature of breaking events, which focus on information transmission rather than conversation.

Hypotheses H2.2 and H2.3 were not confirmed, however, finding no significant differences in time features

TABLE 6a. Quantitative analysis results of exogenous/endogenous categories.

	URL Proportion*		Top Unique Hashtag		Term Length (chars) <sup>†</sup>		Retweet Proportion‡		Reply Proportion§	
	Exo	Endo	Exo (Y/N)	Endo (Y/N)	Exo	Endo	Exo	Endo	Exo	Endo
N	115	55	47/68	36/19	115	55	115	55	115	55
Mean	0.34	0.144	—	—	5.31	6.13	0.32	0.47	0.094	0.083
Median	0.307	0.058	—	—	—	—	0.26	0.38	0.081	0.028

\*log-transformed,  $t = 6.117$ ,  $p < 0.001$ .  $\chi^2 = 9$ ,  $p < 0.002$ . <sup>†</sup> $t = -5.119$ ,  $p < 0.001$ . <sup>‡</sup>sqrt-transformed,  $t = -3.865$ ,  $p < 0.001$ . <sup>§</sup>log-transformed,  $t = 2.98$ ,  $p < 0.003$ .

TABLE 6b. Quantitative analysis results of exogenous/endogenous categories.

	Reciprocity <sup>†</sup>	
	Exo	Endo
N	114	54
Mean	0.26	0.33
Median	—	—

<sup>†</sup> $t = -6.87$ ,  $p < 0.001$ .

between breaking exogenous events and other exogenous events. It is noted, however, that we found that the  $R^2$  quality of fit on the *Exp72* time fit parameters for the tail period was significantly different between breaking and other events, with breaking events having better fit on average than other events. Similar yet marginal differences were found for *Log72* fit parameters. This difference might suggest that the breaking events, after the peak, are less noisy than other exogenous events with discussion levels dropping more “smoothly.”

#### Local Events vs. Other Exogenous Trends

We found limited support that local events have different characteristics than other exogenous trends (H3). In particular, our data surfaced differences between interaction features of local events and other exogenous trends (Table 5). In our dataset we had 12 local events and 96 other exogenous trends (for some parts of the analysis the numbers are lower due to missing data). We note that the analysis was limited by the small number of local events in our trends dataset. The detailed numerical results are shown in Table 6d.

We could confirm only one difference in terms of interaction features between local and other exogenous trends (H3.2), where local events have a smaller proportion of messages that are retweets than other exogenous trends. In addition, our analysis suggests that local events might be more conversational, in terms of the proportion of messages that are replies, than other exogenous trends; however, the result for replies is not significant at the level we require for reporting in this paper, and thus cannot be fully confirmed. We thus can provide only partial support to H3.2.

Finally, we found no support for H3.3, as the differences in time features between local events and other exogenous trends could not be confirmed.

#### Memes vs. Retweet Endogenous Trends

Looking at endogenous trends, “retweet” trends were found to be different than “meme” trends (H4) in content, interaction, time, participation, and social network features (Table 5). In our dataset, we had 29 memes and 27 retweet trends (for some parts of the analysis the numbers are lower due to missing data). The detailed numerical results are shown in Table 6e–g.

In terms of content (H4.1), we found several differences between the retweet trends and meme trends. Retweet trends have a larger proportion of messages with URLs than meme trends, and a higher proportion of unique URLs. More meme trends have a single hashtag that appears in more than 10% of the trend’s messages. Accordingly, meme trends have a larger proportion of hashtags per message than retweet trends, but are not different when we remove the trending terms from consideration (indeed, memes are often identified by the hashtag that the relevant messages contain). Finally, retweet trends have more textual terms in the tweets than meme trends, and the retweet trend tweets are longer on average than meme trend tweets, and are even longer when counting characters in URLs (not reported here). However, these differences may be attributed to the “RT @username” phrase added to individual retweet messages, which were more common, of course, for retweet trends (this observation was true at the time the data was collected; since then, Twitter has changed its format for retweet messages, so that “RT @username” does not always appear in the data).

Supporting H4.2, retweet trends naturally have a significantly greater proportion of messages that are retweets than meme trends. Retweet trends also have a greater portion of replies, showing that they are slightly more conversational than meme trends. The retweet and meme trend categories are not different in their proportions of mentions.

Regarding time features, addressing H4.3, we found one difference between the time fit parameters: the head fit parameter *Log8* head  $p_1$ , indicating different growth for the retweet trends. This finding may suggest that retweet trends develop in a different manner than meme trends.

TABLE 6c. Quantitative analysis results of breaking/other categories.

	Retweet Proportion*		Reply Proportion <sup>¶</sup>		Exp72 tail $R^{2\dagger}$		Log72 tail $R^{2§}$	
	Breaking	Other	Breaking	Other	Breaking	Other	Breaking	Other
N	33	63	33	63	26	37	26	39
Mean	0.39	0.28	0.058	0.11	0.0119	0.0041	0.589	0.386
Median	0.35	0.21	0.049	0.101	—	—	—	—

\*sqrt-transformed,  $t = 3.2$ ,  $p < 0.002$ . <sup>¶</sup>log-transformed,  $t = -2.7$ ,  $p < 0.008$ . <sup>†</sup> $t = 3.554$ ,  $p < 0.001$ . <sup>§</sup> $t = 4.508$ ,  $p < 0.001$ .

TABLE 6d. Quantitative analysis results of local/other categories.

	Retweet Proportion*	
	Local	Other
N	12	96
Mean	0.18	0.345
Median	0.148	0.2

\*sqrt-transformed,  $t = -4.82$ ,  $p < 0.001$ .

Looking at the participation features (H4.4), we found a number of significant differences between retweet and meme trends, supporting the hypothesis. Meme trends have more messages per author on average than retweet trends in a statistically significant manner (we performed the Mann–Whitney test due to the nonnormal distribution of this parameter). In addition, meme trends have a higher proportion of messages from the single top author than retweet trends, as well as a higher proportion of messages from the top 10% of authors than retweet trends. These results show a significant difference in participation between these types of trends, where meme trends are more skewed, and a limited number of users are responsible for a fairly significant part of the content, and retweet trends are more “democratic” and participatory.

Finally, retweet trends were significantly different than meme trends in a number of social network features, confirming H4.5. In terms of the proportion of reciprocation in the trends’ authors social network, retweet trends had a lower level of reciprocated ties than meme trends. Meme trends also had a higher average size of strongly connected components than retweet trends. These findings suggest that retweet trends are supported by a network that, while showing the same density, builds on directional, informational ties more than meme trends that are supported by communication and reciprocity.

### Discussion

The results of our quantitative analysis provide a strong indication that we can use the characteristics of the messages associated with a trend to reason about the trend, for example, to better understand the trend’s origin and context.

In particular, we found that exogenous trends, originating from outside the Twitter system but reflected in the activity

of users in the system, are different in a number of important features from endogenous trends, which start and develop in the Twitter “universe.” Connections between the authors of messages in endogenous trends tend to be more symmetrical (i.e., with higher reciprocity) than in exogenous trends, suggesting perhaps that endogenous trends require stronger ties to be “transmitted.” However, we also expected the density of the endogenous trend networks and the average degree of their nodes to be higher, but did not find any such differences. Differences between these two categories of trends were not evident in the temporal features, where the results did not support our hypothesis of more rapid curve leading to the peaks of the exogenous trends. However, the differences between these categories are further supported by the deviations in content and interaction features between the categories: more URLs, unique URLs, and unique hashtags, as well as a smaller proportion of retweets, show that exogenous trends generate more independent contributions than endogenous trends do.

In a deeper examination of the differences between categories of exogenous trends, we found only interaction differences between trends representing “breaking” events and other type of exogenous trends—breaking events are, naturally perhaps, more “informational” and less “conversational” in nature than other trends. Significantly, we could not confirm the hypothesis from Sakaki et al. (2010) that breaking events will be more disconnected, as multiple contributors will independently contribute messages with less in-network coordination. However, one possible reason for not seeing this effect in the data is the long period of content (72 hours before and after a trend’s peak) over which we calculate the author networks. Perhaps focusing on the connection between authors in the first hours of a trend would capture these differences between breaking events and other trends.

Trends capturing local events were found to be only slightly different than other exogenous trends mainly with respect to the interaction features. People discuss more, and forward information less, in the context of local events as compared to other exogenous trends. Note again that we have a low number of local events represented in our trend dataset and these findings should be considered tentative. Yet, it is possible that differences between local events and other trends would be even more pronounced when more data is available.

Finally, we have shown that even endogenous trends, which grow and develop from within the Twitter system

TABLE 6e. Quantitative analysis results of meme/retweet categories.

	URL Proportion*		Unique URL Proportion <sup>¶</sup>		Hashtag Proportion <sup>†</sup>		Length (term) <sup>‡</sup>		Length (chars) <sup>§</sup>	
	Memes	RTs	Memes	RTs	Memes	RTs	Memes	RTs	Memes	RTs
N	27	22	27	22	27	22	27	22	27	22
Mean	0.044	0.24	0.035	0.064	0.989	0.275	13.19	17.83	80.3	106.1
Median	0.032	0.103	0.029	0.06	0.998	0.092	—	—	—	—

\*log-transformed,  $t = -4.231$ ,  $p < 0.001$ . <sup>¶</sup>log-transformed,  $t = -2.759$ ,  $p < 0.008$ . <sup>†</sup> $t = 5.552$ ,  $p < 0.001$ . <sup>‡</sup> $t = -5.156$ ,  $p < 0.001$ . <sup>§</sup> $t = -4.621$ ,  $p < 0.001$ .

TABLE 6f. Quantitative analysis results of meme/retweet categories.

	Term Lengths (chars)*		Top Unique Hashtag <sup>¶</sup>		Retweet Proportion <sup>†</sup>		Reply Proportion <sup>‡</sup>		Log8 head $p_1$ <sup>§</sup>	
	Memes	RTs	Memes (Y/N)	RTs (Y/N)	Memes	RTs	Memes	RTs	Memes	RTs
N	27	22	27/0	7/15	27	22	27	22	7	15
Mean	6.65	5.57	—	—	0.309	0.692	0.029	0.079	0.055	-0.109
Median	—	—	—	—	0.277	0.739	0.018	0.0565	—	—

\* $t = 4.017$ ,  $p < 0.001$ . <sup>¶</sup> $\chi^2 = 27.99$ ,  $p < 0.001$ . <sup>†</sup>sqrt-transformed,  $t = -8.633$ ,  $p < 0.001$ . <sup>‡</sup>log-transformed,  $t = -3.704$ ,  $p < 0.001$ . <sup>§</sup> $t = 3.549$ ,  $p < 0.002$ .

TABLE 6g. Quantitative analysis results of meme/retweet categories.

	Messages/author*		Messages/top author <sup>¶</sup>		Messages/top-10% author <sup>†</sup>		SCC size (avg) <sup>‡</sup>		Reciprocated Ties <sup>§</sup>	
	Memes	RTs	Memes	RTs	Memes	RTs	Memes	RTs	Memes	RTs
N	27	22	27	22	27	22	27	22	27	22
Mean	2.067	1.171	0.042	0.019	0.382	0.182	1.861	1.332	0.363	0.3
Median	—	—	0.018	0.01	0.383	0.157	1.582	1.239	—	—

\*Mann-Whitney  $Z = 5.44$ ,  $p < 0.001$ . <sup>¶</sup>log-transformed,  $t = 2.793$ ,  $p < 0.008$ . <sup>†</sup>sqrt-transformed,  $t = 8.814$ ,  $p < 0.001$ . <sup>‡</sup>log-transformed,  $t = 3.53$ ,  $p < 0.001$ . <sup>§</sup> $t = 3.936$ ,  $p < 0.001$ .

and are not a reflection of external events, could have different categories that are different in a number of key features. Retweet trends, where users respond and forward a message from a single popular user, are different in many characteristics (including content, interaction, time, participation, and social characteristics) than meme trends.

## Limitations and Other Considerations

Before we conclude, we list several important considerations about our study, acknowledging a few limitations and biases in the work. One limitation is in the dataset used in this work, which is incomplete for two reasons. First, we generated the initial set of trends to analyze using two specific, albeit well-established methods. (As we discussed, the focus of this article is not on the trend detection but rather on the analysis of the trends.) However, better methods for trend detection and identification of related content, exploiting both textual and nontextual information (Becker et al., 2010), might assist in capturing additional trends. At the same time, we believe that the sample of trends reflects the span of trend categories that can reasonably be detected by any

method. The second reason why our dataset is incomplete relates to the selection of the tweets that we used for both trend extraction and characterization: specifically, we only included content from New York City users who disclosed their hometown location in their profile and hence excluded content from other local users without an explicit profile location. (Automatically matching locations and users with no explicit geographical information in their profiles is the subject of interesting future work.) In addition, we defined each trend using terms, and we retrieved the messages associated with each trend via simple keyword search. In the future, more complete message sets for each trend could be assembled by using clustering, so that we could associate a message with a trend even if the message and the trend do not include the same terms (Becker et al., 2010).

Furthermore, our analysis focuses on a single system (i.e., Twitter) and a single location (i.e., New York City). Other dynamics and trend characteristics may exist in other systems and locations (e.g., involving Facebook data and concerning users based in Paris, France). Indeed, the dynamics we observed, and some of the characteristics we extracted, are unique to Twitter. However, Twitter is an important

communication and information service that has already made considerable impact on our society, and is important to study regardless of generalization to other SAS systems. Moreover, we have no reason to believe that, other than message volume, trends involving New York City users are significantly different from trends for other locations.

The metrics that we have used to characterize trends can be extended or further developed. For example, for the time-based characterization, one could experiment with different fitting functions, identifying peaks in different ways (e.g., considering the expected volume of tweets for each time of the day), using different time periods before and after the peak, and so forth. In another example, the social network characteristics could consider the social network of authors that appeared in the first 24 hours of the trends, following Yardi and boyd (2010), which might produce networks of different characteristics. These different methods could expose more pronounced differences between trend categories. However, we believe the wide-ranging set of metrics presented here can serve as a good starting point for analysis, and has already helped identify key differences between types of trends.

## Conclusions

“If Twitter had trending topics for Portland, #rain would be our Justin Bieber.”

—@waxpancake on Twitter

Temporal patterns in data on social awareness streams such as Twitter are becoming increasingly important to our society’s information and communication landscape (Yardi & boyd, 2010; Takhteyev et al., 2010). These temporal patterns—for example, emerging trends in SAS data for local communities—are thus deserving of in-depth understanding and analysis. In this work we categorized and characterized Twitter trends (or “trending topics”) for one geographic area, New York City, and showed that not all trends are created equal. There are different types and categories of trends that are reflected in the data for this local community, and these trends are different in a number of ways that can be automatically computed. Our findings suggest directions for automatically distinguishing between different types of trends, perhaps using machine learning or model-based approaches, utilizing the trend characteristics we propose above as well as others. In particular, perhaps most interesting to us is the identification of local happenings and events that might be underrepresented in other sources, such as traditional news media, but which are of great interest to local communities. Given a robust classification of trends, which could follow from the work described above, we can improve prioritization, ranking, and filtering of extracted trends on Twitter and other SAS, as well as provide a more targeted and specialized visualization of content associated with each trend. Such a set of automated tools will significantly increase the utility of social awareness streams for individuals, organizations, and communities.

## Acknowledgments

This material is based on work supported by the National Science Foundation under Grants IIS-0811038, IIS-1017845, and IIS-1017389, and by two Google Research Awards. In accordance with Columbia University reporting requirements, Professor Gravano acknowledges ownership of Google stock as of the writing of this article.

## References

- Allan, J. (2002). Introduction to topic detection and tracking. In J. Allan (Ed.), *Topic detection and tracking—event-based information organization* (pp. 1–16). Amsterdam: Kluwer Academic Publisher.
- Asur, S., & Huberman, B.A. (2010). Predicting the future with social media. Technical Report HPL-2010-53, HP.
- Becker, H., Naaman, M., & Gravano, L. (2010). Learning similarity metrics for event identification in social media. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining (WSDM '10)* (pp. 291–300). New York: ACM Press.
- Boll, S., & Westermann, U. (2003). MediaEther: An event space for context-aware multimedia experiences. In *Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence (ETP '03)* (pp. 21–30). New York: ACM Press.
- Campbell, K.E. (1990). Networks past: A 1939 Bloomington neighborhood. *Social Forces*, 69(1), 139–155.
- Chen, L., & Roy, A. (2009). Event detection from Flickr data through wavelet-based spatial analysis. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM '09)* (pp. 523–532). New York: ACM Press.
- Cheong, M., & Lee, V. (2009). Integrating Web-based intelligence retrieval and decision-making from the Twitter trends knowledge base. In *Proceedings of the Second ACM Workshop on Social Web Search and Mining (SWSM '09)* (pp. 1–8). New York: ACM Press.
- Crane, R., & Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41), 15649–15653.
- Dayan, D., & Katz, E. (1992). *Media events: The live broadcasting of history*. Cambridge, MA: Harvard University Press.
- De Longueville, B., Smith, R.S., & Luraschi, G. (2009). “OMG, from here, i can see the flames!”: A use case of mining location based social networks to acquire spatio-temporal data on forest fires. In *Proceedings of the 2009 International Workshop on Location Based Social Networks (LBSN '09)* (pp. 73–80). New York: ACM Press.
- Diakopoulos, N.A., Naaman, M., & Kivran-Swaine, F. (2010). Diamonds in the rough: Social media visual analytics for journalistic inquiry. In *Proceedings of the IEEE Symposium on Visual Analytics Science & Technology (VAST '10)* (pp. 115–122). Washington, DC: IEEE Press.
- Diakopoulos, N.A., & Shamma, D.A. (2010). Characterizing debate performance via aggregated Twitter sentiment. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI '10)* (pp. 1195–1198). New York: ACM Press.
- Events (2002). In *Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/entries/events/>
- Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). Information diffusion through blogspace. In *Proceedings of the 13th International Conference on World Wide Web (WWW '04)* (pp. 491–501). New York: ACM Press.
- Hampton, K., & Wellman, B. (2003). Neighboring in netville: How the Internet supports community and social capital in a wired suburb. *City & Community*, 2(4), 277–311.
- Jansen, B.J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60(11), 2169–2188.
- Johnson, S. (2009). How Twitter will change the way we live. *Time Magazine*. Retrieved from <http://www.time.com/time/business/article/0,8599,1902604,00.html>



- Kleinberg, J. (2003). Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery*, 7(4), 373–397.
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)* (pp. 591–600). New York: ACM Press.
- LeCompte, M., & Schensul, J. (1999). *Designing and conducting ethnographic research*. Walnut Creek, CA: Altamira Press.
- Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). Meme-tracking and the dynamics of the news cycle. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '09)* (pp. 497–506). New York: ACM Press.
- McPherson, M., Smith-Lovin, L., & Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.
- Mok, D., Carrasco, J.-A., & Wellman, B. (2010). Does distance still matter in the age of the Internet? *Urban Studies*, 47(13), 2747–2783. DOI: 10.1177/0042098010377363
- Naaman, M., Boase, J., & Lai, C.-H. (2010). Is it really about me? Message content in social awareness streams. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)* (pp. 189–192). New York: ACM Press.
- Nagarajan, M., Gomadam, K., Sheth, A.P., Ranabahu, A., Mutharaju, R., & Jadhav, A. (2009). Spatio-temporal-thematic analysis of citizen sensor data: Challenges and experiences. In *Web Information Systems Engineering (WISE 2009)*. Lecture Notes in Computer Science, 5802, 539–553.
- Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research and Evaluation*, 8(6).
- Petrovic, S., Osborne, M., & Lavrenko, V. (2010). Streaming first story detection with application to Twitter. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 181–189). East Stroudsburg, PA: ACL.
- Phelan, O., McCarthy, K., & Smyth, B. (2009). Using Twitter to recommend real-time topical news. In *Proceedings of the Third ACM Conference on Recommender Systems (RecSys '09)* (pp. 385–388). New York: ACM Press.
- Putnam, R.D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon and Schuster.
- Rattenbury, T., Good, N., & Naaman, M. (2007). Towards automatic extraction of event and place semantics from Flickr tags. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '07)* (pp. 103–110). New York: ACM Press.
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)* (pp. 851–860). New York: ACM Press.
- Salton, G. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.
- Sankaranarayanan, J., Samet, H., Teitler, B.E., Lieberman, M.D., & Sperling, J. (2009). Twitterstand: News in tweets. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '09)* (pp. 42–51). New York: ACM Press.
- Sayyadi, H., Hurst, M., & Maykov, A. (2009). Event detection and tracking in social streams. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM '09)*.
- Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010). Distance matters: Geo-social metrics for online social networks. In *Proceedings of the 3rd Workshop on Online Social Networks (WOSN '10)* (p. 8). Berkeley, CA: USENIX.
- Shamma, D.A., Kennedy, L., & Churchill, E. (2010). Statler: Summarizing media through short-message services. Paper presented at the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10), Savannah, GA.
- Shklovski, I., Palen, L., & Sutton, J. (2008). Finding community through information and communication technology in disaster response. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)* (pp. 127–136). New York: ACM Press.
- Singh, V.K., & Jain, R. (2010). Structural analysis of the emerging event-web. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)* (pp. 1183–1184). New York: ACM Press.
- Starbird, K., Palen, L., Hughes, A.L., & Vieweg, S. (2010). Chatter on the red: What hazards threat reveals about the social life of microblogged information. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)* (pp. 241–250). New York: ACM Press.
- Takhteyev, Y., Gruzd, A., & Wellman, B. (2010). Geography of Twitter networks. Manuscript submitted for publication. Available from <http://takhteyev.org/papers/Takhteyev-Wellman-Gruzd-2010.pdf>
- Tilly, C. (1974). *An urban world*. Boston: Little, Brown.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge, MA: Cambridge University Press.
- Yang, Y., Pierce, T., & Carbonell, J. (1998). A study of retrospective and on-line event detection. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '98)* (pp. 28–36). New York: ACM Press.
- Yardi, S., & boyd, d. (2010). Tweeting from the town square: Measuring geographic local networks. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM '10)* (pp. 194–201). Menlo Park, CA: AAAI.
- Zacks, J.M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127.
- Zhao, Q., Mitra, P., & Chen, B. (2007). Temporal and information flow based event detection from social text streams. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI '07)* (pp. 1501–1506). Menlo Park, CA: AAAI.