Online Communication Dynamics During Natural Hazard Events

Abstract

Informal exchange of information occurs continually throughout daily life. These pre-existing communication patterns are vital during non-routine circumstances such as emergencies and disasters. In recent years, informal communication channels have been transformed by the widespread adoption of social media technologies and mobile devices. Although the potential to exploit this capacity for disaster response is increasingly recognized by practitioners, relatively little is known about the dynamics of informal online communication in response to exogenous hazard events. To address this gap, we employ a longitudinal and comparative approach to examine the dynamics of online communication and information exchange during tornado events in the U.S.

Keywords: informal communication, social media, natural hazards, social vulnerability

1 Introduction

Informal exchange of information occurs continually throughout daily life. Casual conversations, informal greetings with strangers, gossip among friends, and rumoring are characteristic human behaviors (Dunbar, 1997). A large portion this informal interaction occurs via pre-establish social connections and everyday communication channels. Individuals utilize these ties to obtain information about both important and mundane topics (Granovetter, 1974). While individuals can theoretically communicate or interact with any other human being, in practice communication channels are constrained by many different barriers (Latané et al., 1995). Recently, however, developments in social media technologies and mobile devices have transformed informal communication channels allowing individuals to reach a larger number of contacts across much greater distances than previously possible. These new communication technologies are widely used by the general public for a variety of purposes including information exchange, emotional expression, and collective sensemaking (Bordia and DiFonzo, 2004; Choudhury et al., 2012; Lansdall-Welfare et al., 2012).

Pre-existing informal channels are often the first to deliver hazard-related information. Findings from the disaster literature reveal a strong tendency for actors in crisis settings to use their social networks to obtain factual information regarding imminent hazards (Drabek, 1968; Perry and Green, 1982). This can be vital during non-routine circumstances, where information is time sensitive, possibly life-saving, and often critical for decision making. In fact, in cases where official sources are unavailable (or insufficiently timely), social ties serve as the primary conduits of information (**?**Turner, 1994). Indeed, such conduits generally outpace official sources (Erickson et al., 1978; Sutton, 2010). This research is supported by observations that actors will use their social ties in order to obtain information regarding negative or anxiety-provoking events (Rosnow et al., 1988; Kapferer, 1989; Walker and Blaine, 1991; Sutton et al., 2008). Indeed, researchers have already observed heightened use of many social media technologies during crisis situations (Vie, 2010; Sutton, 2010).

Within the emergency management community, there is a growing appreciation for the potential value of online communication networks for disseminating emergency warnings and alerts. Moreover such platforms have also gained attention as a tool for information solicitation. Although the potential to exploit the capacity of these new systems for disaster response is increasingly recognized by practitioners, relatively little is known about the dynamics of informal communication during extreme events, particularly in online environments. It is this gap that we address with this research. A better understanding these social processes has both theoretical and practical implications. One can evaluate existing and develop new theories of individual and group behavior during collective stress situations. Such work could also directly impact response and recovery strategies potentially reducing human and economic losses, and mitigating long-term damage to disaster affected communities.

2 Informal Communication During Crisis Events

Literature on "improvised news" suggests that crisis events arouse popular excitement leading to the formation of a "*public*, consisting of those who are in some way concerned with an event that has disturbed the routine of organized life" (Shibutani, 1966, p. 37). This public, or set of followers, can have changing membership, be diverse in their reactions, and be spatially distributed, but they are tied together because of their common focus of attention – the crisis event. In such a situation, the public demands *news* – information that allows one to make adjustments to changing circumstances. When formal channels fail to deliver such information, individuals turn to informal sources and rumors arise.

Though there is much debate on the definition of a *rumor*, used in this context it refers to communication surrounding facts or events of topical interest that does not occur as part of a formal, institutionalized process. No judgement is made on the accuracy of the information in this case; rumor is not synonymous with false information. As such rumors can arise in many different settings. Indeed, informal exchange of information in face-to-face settings has been studied by many different disciplines. For example, research on group dynamics outlines some of the pressures that contribute to a communication act with others, including social norms, emotional states, and desires to establish or change one's position (Festinger, 1950). Beyond influences on individual participants in the communication process, many of the classical theories of rumor focus on the transmission of information and the factors that influence this process (Allport and Postman, 1947).

Despite notable previous work on rumoring behavior, comparative studies of social processes such as informal communication during disasters or crisis events are rare. This is understandable; the disasters context is extremely hard to study. For example, one fundamental social process we want to understand is information exchange – how does information flow through a population during a disaster? This is an old topic, but it is very difficult to examine. In one of the earliest studies of its kind, Scanlon (2007) set out to explore this topic in a series of field studies. His team was interested in how information about emergency events spreads through a population by word of mouth. It is almost impossible, to collect data on this kind of exchange by administering a survey to an entire community, though people have tried this as well. Instead, Scanlon and colleagues designed a series of tracing studies. Given a particular piece of information they asked each individual in the event-affected community if they had heard it. For every individual who answered yes, they then asked from whom they heard it. For all cases of interpersonal transmission they then tracked down that next individual and repeated the process. After considerable effect and time investment, they were able to trace many information chains back to the original source.

While this method may work for a single case study and limited information scope, it does not scale to large-populations nor a large corpus of informational content of interest. Online environments, however, can offer a window into informal communication during extreme events, allowing for the evaluation of classical theories of how these kinds of social processes occur. More generally, social media and social network platforms have been the subject of new research on social processes during disasters.

3 The Role of Social Media

Increased use of social media and other information and communication technologies (ICTs), by both the public and organizational actors, has led to information production and consumption on a massive scale. Moreover the speed at which information can be produced and retrieved means that one can now receive up-to-the-minutes updates about events occurring around the globe. When crises occur these familiar communication channels are appropriated for exchanging emergency warnings/alerts, event-related information, and checking in with family and friends (Starbird et al., 2010; Sutton, 2010). Indeed, social media platforms such as the microblogging service

Twitter, have gained significant attention in recent years as a widely used and viable tool for disaster response and recovery. Even government agencies are producing guidelines for use of social media during disasters and other emergency situations.

Despite the rapidly growing use of social media during crisis, research on the use of social media and ICTs during disasters in still limited. Preliminary work in this area has focused on descriptive analyses of current usage, demonstrating, for example, that Twitter has been used during crises as a mechanism for resource mobilization, collaboration, citizen reporting, as well as an important source of life-safety information (Vie, 2010; Starbird and Palen, 2010; Sutton, 2010; Liu, 2008). Researchers have also observed that many of the common behavioral patterns known to occur in disaster settings are mirrored online. This includes mass convergence of volunteers and aid, increased attention on local actors, and emergent group formation. This recognition is particularly important because it implies that online environments can provide researchers an additional view of post-event collective behaviors that were previously difficult to explore. Many of the behaviors associated with collective response behaviors have become "more noticeable in light of increasingly pervasive information and communication technology (ICT)" Vieweg et al. (2008).

New technologies have also expanded the ways in which people can become involved in disaster response and recovery. Social media tools, for example, have been used as communication tools, venues for collective information gathering, collective sensemaking, and platforms for raising awareness of volunteering, aid, and support opportunities by population distributed across the globe. Groups such as Humanity Road for example, rely on Internet and mobile communications technology to "collect, verify and route information online during sudden onset disaster"[1]; they are a group of coordinated digital volunteers. Despite the growing body of work in this area, research on large-scale communication dynamics during such a context is rare. We aim to fill this gap.

[1] humanityroad.org

4 Classical Rumor Theory

There are many theories within sociology, social psychology, communication, and the social sciences more generally that make specific predictions about the dynamics of informal communication processes during non-routine situations. In the following sections we explore some of these theories in depth and develop a set of hypotheses motivated by their claims.

Early work on rumor was motivated by observations about their importance during war where conditions were characterized by extreme uncertainty and anxiety along with a lack of information. Allport and Postman (1947) and Caplow (1947) are two of the seminal works in this area. Much of the work on rumor during this period focused on the distortion of information as a result of transmission; scholars were particularly interested in how information could become so drastically different from its original form as it diffused through a population. Allport and Postman (1947) describe this process as one of *leveling, sharpening, and assimilation.* Information not only looses detail during the transmission process, but certain details are selected over others. Individuals may also subconsciously distort information by imputing details that fit with common experience.

Another notable contribution to literature on rumoring is the work of sociologist Shibutani (1966), who conceived of rumoring as a collective process in which individuals attempt to comprehend ambiguous situations by filling gaps in their knowledge. For Shibutani, optimal conditions of rumoring behavior are those where the demand for information exceeds the available information. He points to the following contributing factors: "the size of geographic distribution of those who make up the public [attending to the event],...availability of communication channels...[and level of] collective excitement [or tension]" (Shibutani, 1966, p.165).

Many others have contributed theories of rumor transmission; taken together previous research on rumoring identifies three important factors that affect to extent of rumoring and the probability that an individual will pass along information heard. These factors include (1) the perceived importance of the events, (2) the degree of cognitive unclarity surrounding the event, and (3) the relevance to behavior of the event. While not specific to non-routine situations, the disaster context is certainly one in which we expect to observe rumoring processes. Each determinant of rumoring offers a direct mechanism by which the environment and characteristics of the exogenous events may systematically produce changes in information exchange.

5 The Disaster Context

Though many early studies of rumoring did use disaster settings as the context for exploring informal communication processes, they are not the sole area of focus. For disaster researchers, one the other hand, crisis situations are at the core of their investigations but the kinds of social behavior under study is diverse (Drabek and McEntire, 2003). In fact, sociologists have long been interested in studying emergencies and disasters; exploring human responses to such events "reveal key values and structures that define communities and the societies they comprise...Thus, both core behavior patterns and the social factors that constrain them may be illuminated by the study of disaster" (Drabek, 2005). Both formal and informal communication systems are of primary concern.

Barton (1969), for example, explores a number of different individual and collective phenomena during collective stress situations in his book *Communities in Disaster*; communication is an important thread throughout. Notable for our purpose here are Barton's theories about how event and population characteristics are associated with the amount of communication surrounding the hazard. Barton's treatment of the complex interactions between many different features of both the physical and social environments is extensive, but a few features stand out. In a model of communication, Barton points to the following as influential factors: "the severity of the impact, its suddenness, [and] its randomness with respect to social categories" (Barton, 1969, p.217). For each of these factors, Barton suggests that the mechanism involved is contact with a victim; for example, high impact events both directly affect large numbers of people and draw mass media attention increasing victim exposure and driving communication.

The last of Barton's factors considers the characteristics of the hazard affected population. Disaster researchers have long recognized that the social and demographic characteristics of the communities that experience disaster events is a significant factor in determining disaster losses and resiliency (). While Barton (1969) suggests that diversity in the affected population leads to increased attention and discussion surrounding the event, other scholars posit alternate theories for how communication dynamics are related to population characteristics. In particular, we consider both theories of social vulnerability and social status.

Individuals in society are not equal; some individuals have more power, status, and influence than others (Tilly, 1998). Research on social stratification continues to reveal which aspects of society stratify the population along status lines (Davis and Moore, 1945; Tumin, 1953; Rosenbaum, 2011). In the United States, there are certain characteristics that are associated with social status – wealth and income, gender, education, and race for example (Cutter et al., 2003). Research has shown in many cases that negative outcomes, such as a destructive tornado, that occur to high status groups get more attention then if they were to occur to low status groups (Mileti et al., 1999). The mechanisms behind this differential attention may be complex; high status people may have more social contacts or more spatially diverse social ties. High status groups typically have more access to new communication technologies. In any case, status characteristics of the affected population could be associated with communication rates during hazard events.

Status effects are all about inequality; research suggest that the public behavioral response – in our case informal communication – is very different given the same outcome or event but different affected populations. Another dimension of inequality in disasters is about differential outcomes. Social factors also play an important role here. Indeed, disaster scholars have focused significant attention on exploring the relationship between social and demographic characteristics and risks. Theories of social vulnerability demonstrate that racial and ethnic minorities, for example, experience disproportionate losses after disasters – "mainly because they live in lower-quality housing and disasters exacerbate poverty" (Mileti et al., 1999).

Despite long-standing research on social processes during crisis events, previous work lacks extensive empirical support for theories of communication. Observing and collecting data on informal communication is difficult in routine situations; the disaster context only makes it more difficult. As such, early studies were limited due to two factors (1) they typically only consider a single case (or very few) limiting cross event comparison and generalizability, and (2) they focused on successful rumors (because they were possible to observe) potentially leading to biased conclusions. We aim to address both these issues. First, we compare informal communication dynamics across many events of a particular hazard type - here we focus on the case of tornado events. Moreover, we consider hazard-related communication longitudinally; this includes both pre and post-event communication as well as across events that earned national attention and those that did not.

5.1 Informal Online Communication

There are many different online communication channels that might be considered for collecting hazard-related content indicative of informal conversation pertaining to an event. In this work, we use data from the popular microblogging service Twitter. Twitter is a social networking site designed to allow individuals to find out what their friends are doing. It is one of the most popular microblogging sites currently in use, with over 500 million active users in 2012 who produced over 340 millions tweets per day[2].

Twitter data used in this research were collected by Butts et al. (2011). The data consists of public tweets containing both tornado and "control" related words. Control keywords were chosen randomly from Ogden's Basic English word list (e.g. chalk, house, secretary). Raw data shown in Table 5.1.

Informal Online Communication				
Keyword	**Num. Tweets**	**Mean Hourly Rate**	**Median Hourly Rate**	**Max. Hourly Rate**
collapsed	449,763	66	5	20,822
damaged	1,167,926	114	82	3,600
destroyed	2,433,140	230	151	4,395
funnel cloud	30,597	7	2	713
killed	18,880,360	2,222	1,383	182,571
lost	36,518,284	7,356	4,693	116,200
shelter	2,579,778	227	166	5,550
tornado	2,378,708	454	146	41,521
twister	960,071	94	63	4,871
wind	26,206,862	2,283	1,981	51,791
Total	> 89 million			

Table 1: Descriptive statistics for tornado keywords. Hourly rates estimated using data collected from January 01, 2010 until December 31st, 2011.

Tornado keywords, along with basic descriptive statistics of the frequency of use, i.e. rate of use, are shown in Table 5.1. The dataset allows us to estimate the rate of usage for each keyword at hourly resolution. This results in ten unique keyword-based time series, each spanning the 730 day period from January 1st, 2010 until December 31st, 2011. As seen, words vary drastically in their typical rates of usage. Words like "wind" and "lost" are used on average a few thousand times per hour, while words like "funnel cloud" are used less than 10 times per hour on average during our observation period. In addition, rates of usage for a particular keyword can be quite variable. Table 5.1 also demonstrates that extreme spikes in usage are observed. Keywords like "tornado" for example can see a 100-fold increases in their rate of usage. While individual rates of usage over time are important to note, ultimately we are interested in measuring the underlying topic-related communication as will be discussed presently.

[2] http://en.wikipedia.org/wiki/Twitter

5.2 Physical Events and Impact

Our primary interest in this work is to explore the impact of *context* on online communication; in particular we consider the relationship between the occurrence of a severe natural hazard such as a tornado in the environment and the level of informal online communication surrounding such events on Twitter. Events that occur in the physical environment can be shocks to the social system; a severe tornado, for example, is likely to be associated with a perturbation in the typical level of communication surrounding tornadoes – people talk about what is going on around them. In this research we are concerned with evaluating theories that suggest how the volume of communication may change during these extreme events.

To explore this question we collect data on a set of physical hazard events. Here we focus on tornado events in the United States. Data on tornado warnings and events is available from the U.S. National Oceanic and Atmospheric Administration (NOAA). We collect all issued tornado warnings during the period of interest, along with a number of different estimates of impact of realized events (those where a funnel cloud actually touches down) such as the EF-scale[3], injuries and fatalities, economic loss, et cetera.

In practice, many of the impact-related measures – fatalities, injuries, and loss estimates – are highly correlated. We develop an overall measure of overall *impact* by using the scores on the first principal component of these individual impact measures. In addition to event-level statistics we also consider the annual frequency of tornado events in each state in the country. This gives an approximation of the "surprise" of tornado warning or events. As discussed in the previous sections, geographic variability in tornado events could make tornado warning routine in some areas – perhaps indicating higher levels are preparedness. A simple scaled count of annual frequency of tornado events serves as our measure of event "surprise."

[3]The EF-scale is used to assign a tornado a severity rating based on estimated wind speeds and related damage. http://www.crh.noaa.gov/arx/efscale.php

5.3 Social and Demographic Characteristics of Affected Populations

Each of the physical tornado events described above is associated with a particular geographic location. typically a county or set of counties. As a result, it is possible to obtain social and demographic characteristics about the affected population at the level of the county by matching the event data with other data sources, namely the U.S. Census and American Community Survey. Typical data available includes average and median household income, population statistics, racial composition, et cetera.

Population-level characteristics give an overall picture of the people affected by different tornado events. These population features have also been used to estimate the *social vulnerability* of different populations to environmental hazards, such as tornadoes. Accordingly, we also obtain data on the most well-known of this type of measure - the Social Vulnerability Index (SoVI) developed by Cutter et al. (2003) and colleagues at the Hazards and Vulnerability Research Institute at the University of South Carolina. The SoVI estimates are available for all counties in the U.S.m, matching the level of aggregation of the Census and NOAA data.

6 Methods

Addressing the research questions outlined previously requires the ability to estimate the prevalence of discussion surrounding a given hazard topic – in this case tornadoes. A key indicator of topic-related conversation is the frequency with which posts using topic-relevant terms appear within the broader stream of micro-blog posts. While the presence of a given keyword may be neither necessary nor sufficient to conclude that a specific post speaks to a specific topic, aggregate changes in keyword frequency for multiple topic-related keywords provide a strong indication of topic-related discussion prevalence. Our approach utilizes a latent factor model, which captures this joint variability in the many different observed keyword-based data streams, providing an

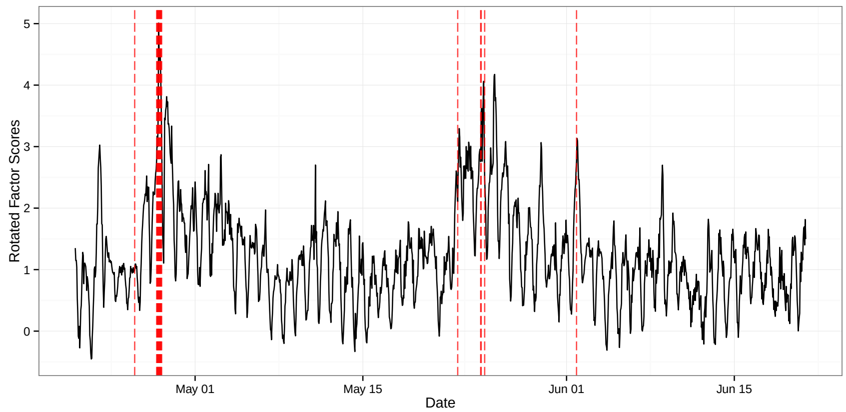


Figure 1: Observed scores from factor model for period of May and June, 2011. Each vertical red lines represents a major tornado event – one that resulted in at least one fatality.

overall estimate of the volume of topic-related communication.

The factor analysis method results in a time series of factor scores that serves as the response of interest. Figure 1 show these scores for the months of May and June, 2011. Each of the dashed red lines shows a major tornado event – one that resulted in at least one fatality. Diurnal patterns are seen in the systematic wave-like oscillations which most likely indicate time of day effects on communication volume. The last week in April had extremely high frequency of tornado events. At first glance there seems to be an association between increased communication (higher observed factor scores) and the these tornado events. The second major increase in communication during the third week of May, immediately following the Joplin, Missouri tornado. Again, in this particular case, the post-event period has relatively high observed communication levels. This variation in topic-related conversation over time and its relation to exogenous hazard events is our primary concern. To test some of the previously discussed theories of rumor and social vulnerability we want to be able to measure the association between the volume of communication and event characteristics overt time; it is natural to approach this problem within the general framework of time series analysis.

7 Results

In April 2011, 748 tornadoes touched down in the United States; 300 occurred in the last week of the month – a record was set for the most events in a 24 hour period. The tornado that hit Joplin, Missouri on May 22, 2011 was the deadliest in the US since 1947, killing 158 people and causing more than 2.5 billion dollars in damage. Overall, the period from January 2010 through December 2011 was an active one in terms of major tornado events in the United States, making it an interesting period of study for this research.

To model the observed data an autoregressive framework is used; the observed rate at time t is assumed to be a linear function of the rate at time t-1, seasonality terms, covariates about possible warnings and events that occur at time t and random noise. We estimate models using standard estimation techniques, utilizing the `arima` function in the `stats` package in the **R** statistical software environment. We evaluate candidate models using the Akaike Information Criterion corrected for small sample size (AICc). The set of candidate models is small enough that an exhaustive comparison is possible.

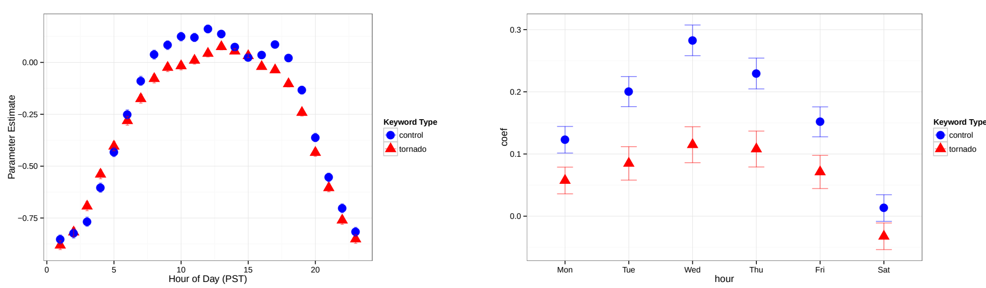
Table 2 shows the parameter estimates for the chosen model of informal tornado-related communication on Twitter. The model shows a strong, positive autoregressive term indicating strong autocorrelation in the data. The rate of communication at a particular time t is closely associated with the rate at time $t-1$. Also noteworthy is the strong positive trend term, demonstrating overall increasing rates of communication over time. This is perhaps an indication of increased overall usage of the Twitter platform over time[4]. Next, we consider the diurnal and greater seasonality in the observed communication rates.

[4] We see similar strong positive overall trend in the "control" keywords time series. Recall, the control keywords are chosen at random from Ogden's English word list and are designed to capture overall trends in informal online communication on Twitter.

7.1 Seasonality in Communication

As with many processes of social dynamics, informal communication rates are known to exhibit seasonality (Golder and Macy, 2011). Both time of day and day of week effects are likely to be strong. For example, communication volume should be lower at night then during day time hours because the majority of the population will be sleeping and this not on Twitter at night. The longitudinal scale of our dataset also allows us to explore larger seasonality - day of week and month of the year effects.

We compare seasonality terms for two different models: (1) the model of tornado-related communication dynamics, described in detail above, and (2) a simple autoregressive model that only includes seasonality terms for the control-related content. Recall, *control* keywords are those chosen at random from Odgen's English word list. We compute a times series of factor scores for the multiple control keyword time series using a confirmatory factor model. Using the control-related communication rates as a comparison case allows us to explore which patterns might be specific to hazard communication and which might be more general.



(a) Time of day effects (b) Day of week effects

Figure 2: Seasonality estimates for tornado and control-related communication models

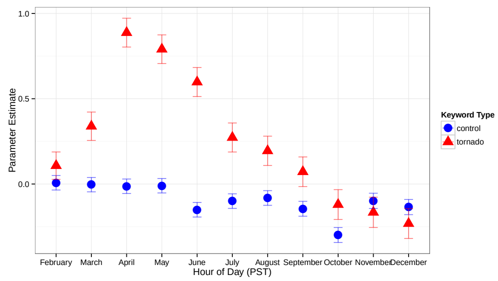
Figure 2(a) shows the parameter estimates for the time of day covariates in the tornado communication dynamics model as well as the control model. Times are converted to Pacific Standard time for ease of interpretation. As seen, both the tornado and control parameter estimates show a clean pattern of increased communication during the day. While the diurnal communication dynamics are similar for tornado and control-related content, and we would expect them to be fairly similar a priori, they are not identical. In fact, tornado-related communication demonstrates a slight delay where the hours of 8:00 am until 11:00 am PST are not quite as elevated as in the control case. Interestingly, this mirrors observed rates of tornado incidence which show that tornadoes occur more frequently between the hours of 3:00 pm and 8:00 pm (local time of the event) [5].

We also observe two peak communication rates for the control data – around 12:00 pm PST and 5:00 pm PST. One explanation of this phenomena would be that it results from the differences in use across time zones in the U.S. Population in the United States in concentrated along the east and west coastlines. Therefore, these parameter estimates are smoothing across two different times zones. Perhaps the first peak is really around 3-5:00 pm EST, following by a similar peak on the west coast. Whereas the control-related content might include a larger geographic distribution of contributors, tornado activity in the Unites States is fairly concentrated is the mid-west, or what has colloquially be termed "Tornado Alley." These difference in diurnal effects could be further explored in future work by considering the relationship between time of day parameter estimates and hazard type.

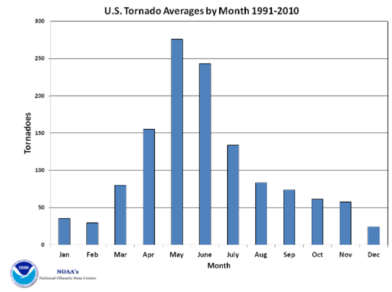
Next, we consider weekly seasonality. Figure 2(b) shows the parameter estimates for both tornado and control-related communication rates for day of week terms. In general, for both cases, we find the communication rates tend to be higher during the week and lower on weekends. Communication rates increase over the beginning on the work week, peaking on Wednesday and then gradually falling Thursday and Friday. Lowest rates of communication occur on Saturday and Sunday (the reference category). For both the time of day and day of week features, it is not surprising that the estimates for the control data are slightly larger because the control model has fewer possible parameters that can be used to explain the variance in the data.

The longitudinal scale of the dataset also allows us to explore monthly seasonality.

[5] http://www.ncdc.noaa.gov/oa/climate/severeweather/tornadoes.html



(a) Month of year effects for tornado and control-related communication models.



(b) Average number of tornado event per month in the U.S.

Figure 3: Monthly seasonality effects for tornado and control-related communication models, compared with average monthly events in the U.S. Monthly estimates for actual events are obtained from http://www.ncdc.noaa.gov/oa/climate/severeweather/tornadoes.html.

Figure 3(a) shows the monthly effect sizes for tornado content compared with control communication. While in the previous two seasonality examples, time of day and day of week effects, the results for tornado and control content were similar, here we find distinct differences. Control-related communication rates are consistent for the first few months of the year, January through May. Summer months are associated with a drop in communication rates which gradually increase to winter and spring level; the one exception here is October, which shows a drop in rates of communication for control related content. If we compare this pattern with tornado-related content, we find that hazard communication shows a much different dynamic. Tornado communication rates increase from January through spring, reaching a peak during April and May, before declining again during summer and fall months. We compare these monthly effects with data from NOAA on the average number of tornadoes events per month, as seen in Figure 3(b). Our estimates of monthly seasonality terms match observed pattern of tornado incidence well.

7.2 Informal Communication and Exogenous Events

Figure 4 shows the warning and event-related parameter estimates for the model of tornado communication on Twitter. Each point is colored according to is significance and direction of association with online communication. Red points have a significant positive association, black points are included in the model but not significant, and blue points have a significant negative association with online, tornado-related communication rates. Recall, these predictors of interest here can be separated into two different categories: (1) those that pertain to the physical event itself, and (2) those that describe the population directly affected by the event. Moreover, we separate these two types of features for tornado warnings and realized tornado events - those where a funnel cloud actually touches down.

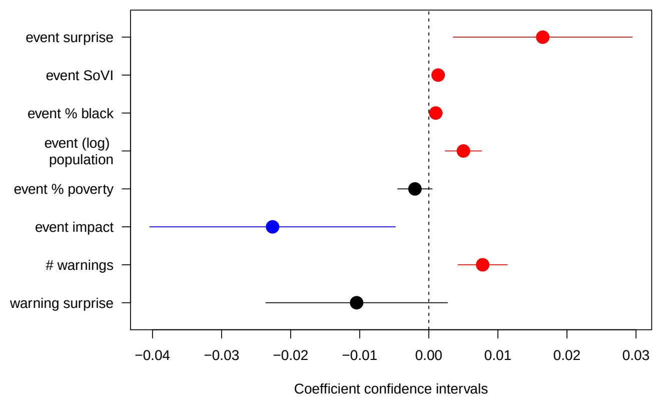


Figure 4: Coefficient estimates for event and population level predictors of interest. Red points represent significant positive estimates and blue points represent significant negative estimates. Intervals show 95% confidence intervals on each parameter estimate.

Tornado events often begin with a tornado watch issued by local weather services. Watches indicate that weather conditions are conducive to the development of tornadoes. A tornado warning is the next step in the process; a tornado warning indicates that a tornado has been sighted, is occurring, or is imminent. Finally, we also obtain data on events that touch down and cause some amount of damage, whether it is minimal or large-scale. The number of warnings/events in a time period and the level of surprise for that warning/event are both predictors of interest in the communication dynamics model. Additionally, for events, we have the measure of impact.

During warning or event periods, information pertaining to these topics is likely more salient to the population, particular those directly affected by the warnings or event. Moreover this information will in many cases be perceived as extremely important and directly relevant to ones behavior. For example, tornado warnings advise one to seek immediate shelter. These features of tornado-related information match Allport and Postman (1947) determinants of rumoring, suggesting we should see increased communication during warning and event periods. Tornado warnings and events may also lead to a state cognitive unclarity. For example, in the case of warnings, people are left to wonder whether events will occur, they are unsure about whether to take protective action. This uncertainty may be even higher in populations that are not used such events, where *surprise* at the warning is high.

The number of warnings has one of the stronger associations with increased communication rates. While the average *surprise* of the warnings is also included in the model it does not have a significant association with the volume of communication. While the number of warnings and subsequent events in a given period are correlated (0.61) they are not identical. Many more warnings are issued than actually turn into tornado events. The surprise score of the event does have a strong, significant positive association with informal communication. In fact, the event surprise has the largest effect size of the positively associated predictors.

While event impact is selected in the model, it has a significant, negative association with communication volume. This is an intriguing result, contradictory to many theories of disaster response and recovery. One potential explanation is that during severe events people are too busy responding to potentially life-threatening situations on the ground to be bothered with posting messages and content to Twitter. However, not all communication comes from those directly affected by the tornado event.

Another plausible explanation for scenarios characterized by high event impact but low levels of communication would be that this occurs during low warnings scenarios. If a tornado warning is issued people have time to take protective action and time to communicate about the event, as the previous results demonstrate. Exposure to the warning information, whether through formal or informal channels, reduces injuries and deaths. On the other hand, if a tornado strikes without warning, populations in its path don't have time to activate informal communication channels, search for information, or take appropriate preemptive protective action. This explanation is in line with work on disaster onset which shows that the length of the warning period is associated with impact (Simmons, 2008).

7.3 Population Characteristics

Each warning and event is specific to a particular location, allowing us to explore differences in informal communication dynamics based on the number of people directly affected by the event and the characteristics of that affected population. Barton (1969), for example, posited that both formal and informal communication volume was a direct function of the number of individuals involved. Indeed, our results suggest the same. The size of the population effected by tornado events has a significant positive association with communication volume. The more people who are directly affected, the more online communication.

Finally, we observe that communication dynamics vary based on social and economic characteristics of the affected population. Although the percentage of the tornado-affected population below the poverty line is included in the model it is non-significant. The percentage of the population that self-reports as Black, or African American alone, has a significant positive association communication volume. In other words, tornado-related communication volume is higher on Twitter during periods where tornado events occur in areas with higher percentages of African Americans. Perhaps this reflects the increased appeal of Twitter in particular to African Americans, who have higher usage rates than other racial and ethnic groups (Duggan and Brenner, 2013).

The last event feature of note is the index of social vulnerability. Recall, this is a aggregate measure of many factors designed to measure the vulnerability of U.S. counties to environmental hazards (Cutter, 2008). Community vulnerability is positively associated with tornado-related communication volume, suggesting that events that occur in high vulnerability areas coincide with more attention in online social media. While this is consistent with much of the disaster literature, it disconfirms the presence of status effects in this case. In fact, many researchers have suggested that disaster contexts remove many social boundaries, allowing for communication, aid, and support across status barriers (Drabek, 2005).

This analysis evaluates the applicability of traditional theories of rumor, social status and vulnerability to online communication during tornado events. Our results demonstrate the many of these theories are consistent with the observed relationship between communication volume and exogenous events, suggesting what social processes might be active during these periods. We discuss come of the implications of these results in the following section.

8 Conclusion

Disaster events result in non-routine environments which individuals must navigate in the face of imminent danger and imperfect information. These exogenous shocks may directly impact a small geographic area but the indirect effects are likely far reaching. One need only consider the global response to events such at the 2010 earthquake in Haiti or the 2010 Tohoku earthquake and tsunami in Japan to see how large-scale crisis events draw the attention of the masses. Even regional events can gain significant attention; more than 20 million messages were posted to Twitter during the week Hurricane Sandy hit the east coast of the United States [6]. In the context of social media, informal communication behaviors during crises is just beginning to be explored.

[6]http://thelede.blogs.nytimes.com/2012/11/02/

This is the first study we are aware of to explore the volume of informal online communication and its relationship to exogenous hazards over time and across multiple events. We consider the case of tornado-related communication over a period of 24 months. Incorporating data from the National Oceanic and Atmospheric Administration and the U.S. Census, we evaluate the relationship between communication volume and event characteristics including impact and features of the affected population. Results suggest strong seasonality patterns in the observed rate of communication; both time of day, day of week, and monthly effects are important. In particular we have shown the theories of information saliency and rumoring apply to tornado-related communication on Twitter. We have also shown that theories of social vulnerability are supported but found no evidence for social status effects.

References

2009. "Large-Scale Community Detection on YouTube : Global Clustering via Local Algorithms." *Intelligence* .

2010. "Microblogging during two natural hazards events: what twitter may contribute to situational awareness." In *Proceedings of the 28th international conference on Human factors in computing systems*, pp. 1079–1088. ACM.

Allport, G.W. and L. Postman. 1947. *The psychology of rumor*. Henry Holt.

Barton, A. 1969. *Communities in Disaster: A Sociological Analysis of Collective Stress Situations*. Garden City, NY: Doubleday and Company.

Bordia, Prashant and Nicholas DiFonzo. 2004. "Problem solving in social interactions on the Internet: Rumor as social cognition." *Social Psychology Quarterly* 67:33–49.

Butts, C. T., J. Sutton, and E. S. Spiro. 2011. "Hazards, Emergency Response, and Online Informal Communication Project Data."

Caplow, T. 1947. "Rumors in war." *Social Forces* 25:298–302.

Choudhury, M De, Michael Gamon, and Scott Counts. 2012. "Happy, nervous or surprised? Classification of human affective states in social media." In *ICWSM*. AAAI.

Cutter, Susan L. 2008. "The Vulnerability of Science and the Science of Vulnerability." *Annals of the Association of American Geographers* pp. 37–41.

Cutter, Susan L, Bryan J Boruff, and W Lynn Shirley. 2003. "Social Vulnerability to Environmental Hazards." *Social Science Quarterly* 84.

Davis, Kingsley and Wilbert E. Moore. 1945. "Some Principles of Stratification." *American Sociological Review* 10:242–249.

Drabek, Thomas E. 1968. *Disaster in Aisle 13: A Case Study of the Coliseum Explosion at the Indiana Fairgrounds, October 31, 1963*. Columbus, OH: The Ohio State University Press.

Drabek, Thomas E. 2005. "Sociology, disasters and emergency management." *Disciplines, Disasters and Emergency Management: The Convergence of Concepts Issues and Trends From the Research Literature* .

Drabek, Thomas E. and David a. McEntire. 2003. "Emergent phenomena and the sociology of disaster: lessons, trends and opportunities from the research literature." *Disaster Prevention and Management* 12:97–112.

Duggan, Maeve and Joanna Brenner. 2013. "The Demographics of Social Media Users 2012."

Dunbar, Robin. 1997. "Grooming, Gossip, and the Evolution of Language." *New aspects of human ethology* pp. 77–89.

Erickson, B.H., TA Nosanchuk, L. Mostacci, and C.F. Dalrymple. 1978. "The flow of crisis information as a probe of work relations." *Canadian Journal of Sociology/Cahiers canadiens de sociologie* pp. 71–87.

Festinger, L. 1950. "Informal social communication." *Psychological review* 57:271–82.

Golder, S. a. and M. W. Macy. 2011. "Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures." *Science* 333:1878–1881.

Granovetter, M.S. Mark. 1974. *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA: Harvard University Press.

Kapferer, JN N. 1989. "A mass poisoning rumor in Europe." *Public Opinion Quarterly* 53:467–481.

Lansdall-Welfare, Thomas, Vasileios Lampos, and Nello Cristianini. 2012. "Nowcasting the mood of the nation." *Significance* 9:26–28.

Latané, Bibb, J. H. Liu, a. Nowak, M. Bonevento, L. Zheng, and B. Latane. 1995. "Distance Matters: Physical Space and Social Impact." *Personality and Social Psychology Bulletin* 21:795–805.

Liu, Hugo. 2008. "Social Network Profiles as Taste Performances." 13:252–275.

Mileti, Dennis et al. 1999. *Disasters by design: A reassessment of natural hazards in the United States*. National Academies Press.

Perry, R W and M R Green. 1982. "The Role of Ethnicity in the Emergency Decision-Making Process*." *Sociological inquiry* 52:306–334.

Rosenbaum, J. E. 2011. "The Complexities of College for All: Beyond Fairy-tale Dreams." *Sociology of Education* 84:113–117.

Rosnow, R L, J L Esposito, and L Gibney. 1988. "Factors influencing rumor spreading: Replication and extension." *Language & Communication* .

Scanlon, Joseph. 2007. "Sampling an unknown universe: problems of researching mass casualty incidents (a history of ECRU's field research)." *Statistics in medicine* 26:1812–1823.

Shibutani, Tamotsu. 1966. *Improvised News: A Sociological Study of Rumor*. Indianapolis, New York: The Bobbs-Merrill Company, Inc.

Simmons, K. M., & Sutter D. 2008. "Tornado warnings, lead times, and tornado casualties: An empirical investigation." *Weather and Forecasting* pp. 246–258.

Starbird, Kate and L. Palen. 2010. "Pass It On ?: Retweeting in Mass Emergency." In *Work*, number December 2004, pp. 1–10.

Starbird, Kate, Leysia Palen, A.L. Hughes, and Sarah Vieweg. 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*, pp. 241–250. ACM.

Sutton, J, L Palen, and I Shklovski. 2008. "Backchannels on the front lines: Emergent uses of social media in the 2007 southern California wildfires." *Crisis* .

Sutton, J N. 2010. "Twittering Tennessee: Distributed Networks and Collaboration Following a Technological Disaster." In *Proceedings of the 7th International ISCRAM Conference–Seattle*, volume 1.

Tilly, Charles. 1998. *Durable Inequality*. University of California Press.

Tumin, M.M. 1953. "Some principles of stratification: A critical analysis." *American Sociological Review* 18:387–394.

Turner, R H. 1994. "Rumor as Intensified Information Seeking: Earthquake Rumors in China and the United States." In *Disasters, Collective Behavior, and Social Organization*, edited by R R Dynes and K J Tierney, pp. 244–256. Newark, NJ: University of Delaware Press.

Vieweg, Sarah, Leysia Palen, Sophia B Liu, Amanda L Hughes, and Jeannette Sutton. 2008. "Collective Intelligence in Disaster: Examination of the Phenomenon in the Aftermath of the 2007 Virginia Tech Shooting." .

Walker, Charles J and Bruce Blaine. 1991. "The Virulence of Dread Rumors: A Field Experiment." *Language & Communication* 11:291–297.

Type	Covariate	Estimate	Std. Error	
Model	AR(1)	0.9037	0.0033	***
	intercept	-1.2974	0.0666	***
	trend	2.5735	0.0747	***
Time of Day	01:00 GMT	-0.0203	0.0091	*
	02:00 GMT	-0.0365	0.0125	**
	03:00 GMT	-0.1034	0.0148	***
	04:00 GMT	-0.2416	0.0166	***
	05:00 GMT	-0.4344	0.0180	***
	06:00 GMT	-0.6044	0.0191	***
	07:00 GMT	-0.7605	0.0199	***
	08:00 GMT	-0.8512	0.0206	***
	09:00 GMT	-0.8811	0.0211	***
	10:00 GMT	-0.8179	0.0214	***
	11:00 GMT	-0.6930	0.0216	***
	12:00 GMT	-0.5391	0.0217	***
	13:00 GMT	-0.4042	0.0216	***
	14:00 GMT	-0.2812	0.0214	***
	15:00 GMT	-0.1753	0.0211	***
	16:00 GMT	-0.0784	0.0206	***
	17:00 GMT	-0.0246	0.0199	
	18:00 GMT	-0.0169	0.0191	
	19:00 GMT	0.0098	0.0180	
	20:00 GMT	0.0431	0.0166	**
	21:00 GMT	0.0754	0.0148	***
	22:00 GMT	0.0541	0.0125	***
	23:00 GMT	0.0318	0.0091	***
Day of Week	Monday	0.0574	0.0215	**
	Tuesday	0.0849	0.0269	**
	Wednesday	0.1149	0.0289	***
	Thursday	0.1080	0.0289	***
	Friday	0.0711	0.0267	**
	Saturday	-0.0324	0.0214	
Month	February	0.1141	0.0817	
	March	0.3585	0.0846	***
	April	0.8960	0.0861	***
	May	0.7965	0.0858	***
	June	0.6100	0.0867	***
	July	0.2765	0.0868	**
	August	0.2055	0.0874	*
	September	0.0944	0.0887	
	October	-0.1283	0.0891	
	November	-0.1518	0.0904	
	December	-0.2178	0.0890	*
Warning	# warnings	0.0078	0.0018	***
	surprise	-0.0113	0.0067	
Event	population	0.0052	0.0013	***
	% poverty	-0.0021	0.0013	
	% Black	0.0010	0.0005	*
	SoVI	0.0013	0.0004	**
	surprise	0.0174	0.0066	**
	impact	-0.2261	0.0907	*

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1

Table 2: Parameter estimates and standard errors for model of communication dynamics of tornado-related conversation.