Understanding Communication Dynamics on Twitter during Natural Disasters: A Case Study of Hurricane Sandy

Nastaran Pourebrahim
Department of Geography, Environment, and Sustainability
UNC Greensboro
PO Box 26170, Greensboro, NC 27402-6170
n_poureb@uncg.edu

Selima Sultana
Department of Geography, Environment, and Sustainability
UNC Greensboro
PO Box 26170, Greensboro, NC 27402-6170
s_sultan@uncg.edu

John Edwards
Social Science Research Center
Mississippi State University
PO Box 5287, Mississippi State, MS 39762
John.Edwards@ssrc.msstate.edu

Amanda Gochanour
Social Science Research Center
Mississippi State University
PO Box 5287, Mississippi State, MS 39762
amanda.gochanour@ssrc.msstate.edu

Somya Mohanty*
Department of Computer Science
UNC Greensboro
PO Box 26170, Greensboro, NC 27402-6170
mohanty_somya@uncg.edu

*Corresponding author

Understanding Communication Dynamics on Twitter during Natural Disasters: A Case Study of Hurricane Sandy

Abstract

This study investigates Twitter usage during Hurricane Sandy following the survey of the general population and exploring communication dynamics on Twitter through different modalities. The results suggest that Twitter is a highly valuable source of disaster-related information particularly during the power outage. With a substantial increase in the number of tweets and unique users during the Hurricane Sandy, a large number of posts contained firsthand information about the hurricane showing the intensity of the event in real-time. More specifically, a number of images of damage and flooding were shared on Twitter through which researchers and emergency managers can retrieve valuable information to help identify storm damages and plan relief efforts. The social media analysis revealed the most important information that can be derived from twitter during disasters so that authorities can successfully utilize such data. The findings provide insights into the choice of keywords and sentiments and identifying the influential actors at different stages of disasters. A number of key influencers and their followers from different domains including political, news, weather, and relief organizations participated in Twitter-based discussions related to Hurricane Sandy. The connectivity of the influencers and their followers on Twitter play a vital role in information sharing and dissemination throughout the hurricane. These connections can provide an effective vehicle for emergency managers towards establishing better bi-directional communication during disasters. However, while government agencies were among the prominent Twitter users during the Hurricane Sandy, they primarily relied on one-way communication rather than engaging with their audiences, a challenge that need to be addressed in future research.

Keywords: Disaster management; Social media; Twitter; Social network analysis; Information diffusion; Hurricane

1. Introduction

Climate change is a major threat of our time, and is expected to intensify the frequency of extreme weather events [1]. Providing emergency information to population living in the vulnerable areas, therefore, has become a measure of disaster resilience and a policy priority [2]. With nearly 2 billion Facebook active users, 6 billion YouTube videos viewed and approximately 700 million tweets posted on Twitter every day, social media platforms have become the major channels for people to communicate and stay informed [3,4]. Given the increasing presence of Social media in everyday life, it can be a major platform for sharing emergency information such as warnings, disaster relief efforts, crisis mapping for escape routes, search and rescue, and connecting community members following a disaster [5–7].

Twitter is one of the prime social media platforms, which has been used not only during emergency situations, but it also changed the way people create, disseminate, and share emergency information [8]. The real-time characteristic of Twitter makes it a suitable crowdsourcing platform for dissemination and collection of information including texts and pictures during disasters and crisis events [9], which enhances the public awareness of a situation instantly. One of the major challenging issues facing emergency officials is the development of warning methods for residents at risk in order for them to take appropriate actions immediately [10]. Twitter use has been growing rapidly especially during disasters by the local officials [11], but the effectiveness of this system is an on-going debate. Hence, it is important to understand its potential as a mean of communication during disaster since it informs people and in turn might influence their actions in preparing for a natural disaster [2]. While the potential role of Twitter during disaster has been discussed, the effectiveness of widespread use of Twitter for receiving and disseminating risk information during such events has been less investigated in academic research [2,12]. Without evidence-based strategies of how these technologies are being used, the implementation of social media tools in disaster management remains challenging [2].

Identifying the appropriate ways to use and analyze social media data is an important task for researchers to draw reliable conclusions so that the full potential of social media during emergencies can be achieved [13]. The importance of content analysis and social network structure during disaster has been identified in the literature [14], but application of Twitter data in academic research is still at its infancy. Given the anticipated increase in frequencies of natural hazards such as hurricanes and their resulting damages in the United States in the coming decades [15], better understanding of social media data is crucial so that it can be utilized by emergency authorities. Examining the structures of the online social networks, the underlying mechanism of online users' behaviors and their shared contents, therefore, is the key to achieve this goal [14].

This study investigates the usage of Twitter during Hurricane Sandy by conducting a general population survey of the affected regions, and analyzing the content / social network structure of messages and users on Twitter during this period. Our objectives are to (1) Identify the sources of information received by and shared with the coastal areas' residents affected by Hurricane Sandy., (2) examine the Twitter users' involvement in discussions about Hurricane Sandy and the content of messages shared before, during, and after the hurricane**,** and (3) examine the social network structure of Twitter users before, during, and after the hurricane. The paper is organized as follows: we first present a review of relevant literature regarding social media data and disaster management in section 2 followed by discussions on study site in section 3, and data collection and methods in section 4. We then discuss the results of our analysis in section 5. We present conclusions, limitations and suggestions for future research in section 6.

2. Social Media Analysis in Disaster

Social media platforms such as Twitter allow public and officials to share texts and photos, which can be a powerful means of communications during disasters. Sharing and transferring local knowledge in all stages of crisis (pre, during, post) and building social capital are essentials for communities' resiliency [16,17]. Social media can facilitate this process since a number of people, disaster-affected communities, and organizations are linked via these online networks [14]. Research shows that local city officials' evaluations of their ability in controlling a crisis and the strength of their responses are positively related to the extent of social media they use [18]. People also expect fast arriving of help after posting a request on a social media site [19]. With the rapid growing of social media as emergency communication channels [14], various scholars have investigated the potential of generated big data using different techniques. For example, The potential use of Twitter during disasters such as floods [20,21], earthquakes [22,23], hurricanes [24], tornados [25], tsunamis [26], wildfires [27], volcanic hazards [28], and droughts [29] has been investigated. The shared information on social media platforms such as Twitter are useful for public and emergency management authorities to understand on-the-ground realities during emergencies [25]. Information exchange behavior of social media users, their various spatial and temporal activity patterns on social media, and the content of shared messages are examples of such useful information. These behaviors and activities have been reported to change based on the crisis life cycle, affected regions, event's type and characteristics [8,30].

Social media contents during disasters have been analyzed using text mining and sentiment analysis. Text mining is becoming a popular method to understand the unstructured text information [14]. A word cloud is commonly used to determine the most frequently used words and illustrate a visual representation of text data generated in social media [14]. However, to understand the true meaning of the text, sentiment analysis has been mainly applied to classify texts into positive and negative [12]. Researchers have become more interested in sentiment analysis during emergency situations using machine learning techniques. Ragini et al., [31], for example, used support vector machine (SVM) to classify the tweets and their positive or negative sentiments of individual's needs during different disasters. In another study, Vo and Collier [32] identified various emotions such as anger, calm, unpleasantness, sadness, anxiety, fear and relief during four of the Japan's earthquakes. The content analysis of social media data coupled with spatial and visual analysis techniques are becoming a major research area to understand potential of such data in disaster management [33].

Another line of study related to social media is exploring online users' behaviors and interactions through social network analysis. Social network analysis is used to study network structures, relationship properties of networks, communication's patterns between users, the role of various actors in the network, and community detection [34,35]. Graph theory is the major approach in social network analysis where a network of nodes and links is created to examine the relationship among social media users [36]. Three main analyses found in literature are influence analysis, link analysis, and community detection [36]. Studies have used various metrics (e.g., degree centrality) and algorithms (e.g., modularity) to explore online network structures, communities and the users' engagement during disasters [6,14]. As the number and expectations of social media users during disasters continue to grow [19], the sources and types of information received by and shared with people and the functioning of social media during extreme whether events need more investigation particularly from a social network perspective. Data generated by social media such as Twitter has been applied in various fields of academic research such as public health, crime analysis and transportation [37–39]. Its application in disaster management research, however, is still at its early stages [12] and our paper intend to fill that gap in the disaster management literature.

3. Context of the Study

Hurricane Sandy started on October 22, 2012, moved from the Caribbean to the U.S. Eastern Seaboard, and finally made landfall near Brigantine, New Jersey, around 8:00 p.m. on October 29, 2012 (Figure 1) [40]. Sandy caused 147 direct deaths, and around 650,000 damaged or destroyed houses, and left approximately 8.5 million customers without power during and after the storm [40], making it one of the deadliest and most destructive hurricanes in the history of the United States [41]. New York, New Jersey and Connecticut, especially in and around the New York City metropolitan received the record levels of storm surges (Figure 2) [40]. Despite heavy power outages and disruptions, people used social media such as Twitter as a critical platform to express the impacts of Sandy on their lives and material goods [42–44]. The hurricane-related tweets that were posted during the Sandy across the entire Twitter platform, and a large volume of geo-tagged Twitter data generated in the affected coastal counties in New York, New Jersey and Connecticut are used for our research.

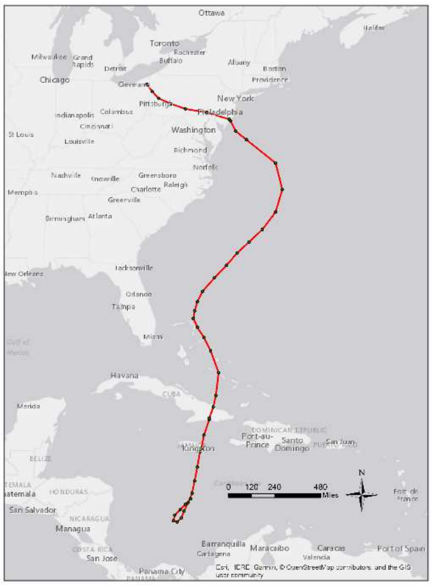

Figure 1. Track of Sandy (source: NHC [45]).

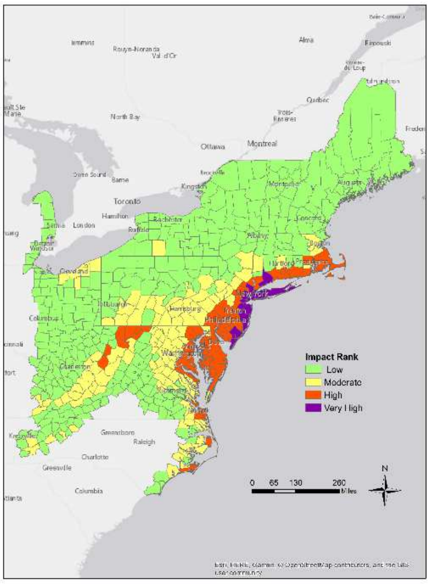

Figure 2. FEMA Hurricane Sandy impact analysis (source: FEMA MOTF [46])

4. Data and Methods

In addition to Twitter data, this study includes survey data. For the first research objective (section 5.1), a two-part survey of respondents who lived in the coastal counties of Connecticut, New Jersey, and New York at the time of Hurricane Sandy was conducted (Figure 3).The first phase of data collection involved the use of a telephone-based survey. Since we were highly interested in the ways respondents used Twitter during Hurricane Sandy, the telephone-based survey was supplemented with a web-based survey of respondents recruited via Twitter. The web-based survey of Twitter users also helped to ensure a large enough sample of Twitter users so that meaningful comparisons between Twitter users and non-Twitter users can be made. The data collection period was started on January 28, 2015 and ended on May 7, 2015.

Figure 3. Survey catchment area.

4.1 Telephone-Based Survey

The first phase of data collection involved a telephone-based survey of a representative sample of residents from 23 counties (Figure 3) in Connecticut (Fairfield, Middlesex, New Haven, New London), New Jersey (Atlantic, Bergen, Cape May, Essex, Hudson, Middlesex, Monmouth, Ocean, Somerset, Union), and New York (Bronx, Kings, Nassau, New York, Queens, Richmond, Rockland, Suffolk, Westchester). Respondents were screened to ensure that they were residents of one of the appropriate states and counties at the time Hurricane Sandy made its landfall. The telephone numbers were purchased from Survey Sampling International. A dual-frame (cellphone & landline), random-digit-dialing (RDD) sample was purchased that included a random sample drawn from the universe of existing telephone numbers assigned to the catchment area. The final sample size for those recruited and surveyed via telephone was 514, with a cooperation rate of 12.4%. We think that the relatively low cooperation rate is likely due to the length of time that passed between Hurricane Sandy and the data collection period (approximately 2 ½ years). The low cooperation rate might have also been influenced by respondents' reluctance due to the large number of research projects that have attempted to survey the local population in the years following Hurricane Sandy. Regarding the sample of respondents recruited via telephone, the demographic characteristics (Table 1) were evenly split according to gender, with the sample comprised of 49.2 percent men and 49.8 percent women. The racial composition of the sample was comprised of 51.6 percent white, 18.9 percent black, 5.1 percent Asian, 12.5 percent "others" and 3.7 percent multiracial. With regard to level of education, the respondents were above average compared to the national population, with about half of the sample holding a bachelor's degree or a graduate degree. The survey includes questions about the source of information received by residents during the hurricane and the types of information they shared. Example of survey questions are represented in Table 2.

Table 1. Demographic characteristics of telephone and web surveys.

	Gender		Racial Composition					Education
	Male	Female	White	Black	Asian	Others	multiracial	Bachelor/Graduate degree
Telephone Survey	49.2%	49.8%	51.6%	18.9%	5.1%	12.5%	3.7%	50%
Web Survey	57%	43%	57.3%,	3.9%	6.8%	3.9%	5.3%	50%

Table 2. Examples of survey questions.

Survey Questions
During Hurricane Sandy, did you get any weather-related information from the following sources (Check all that apply)?
In what ways did you receive weather-related information?
Of the ways you received weather-related information, from which did you get *most* of your information?
During the storm, did the way that you received most of your information change? If yes: a) After it changed, how did you get most of your information? b) Why did the way you received most of your information change during the storm?
At any time during Hurricane Sandy, did your household lose access to:
During Hurricane Sandy, did you tweet or retweet storm-related information? If yes: a) What was the nature of the information that you tweeted or retweeted? b) What was the source of the information that you tweeted or retweeted? c) What form of information did you tweet or retweet?
Do you follow any of these sources of information on Twitter?

4.2 Web-Based Survey

In the second phase of data collection, participants were recruited through Twitter. We used all geo-tagged tweets sent from the catchment area (Figure 3) before, during, and directly following Hurricane Sandy's landfall to identify Twitter users. From this data, approximately 26,000 unique Twitter users were identified, 20,000 of which were randomly selected for inclusion in the web-based survey sample. A Twitter message was sent as a "@ reply" to each Twitter user. When participants clicked a link embedded in the tweet, they were directed to a web-based survey which was virtually identical to the telephone survey. The 20,000 tweets generated 207 click-throughs for a total of 170 completed surveys, resulting in a response rate of slightly less than one percent. Much like the telephone sample, this low response rate is likely attributable to the length of time between the event and the data collection period, and to the respondent's reluctance. Furthermore, Twitter-based survey recruitment is a relatively novel approach and users may have been less willing to consider surveying a legitimate use of Twitter's platform. Regarding the sample of respondents recruited via Twitter, the demographic characteristics are as follows (Table 1): The modal category for gender is men, with men making up 57.0 percent of the sample and women comprising 43.0 percent of the sample. Regarding race, the vast majority of the sample was white 57.3%, followed by 6.8 percent Asian, 5.3 percent multi-racial, 3.9 percent black, and 3.9 percent "others". For level of education, the respondents were above average compared to the national population, with about half of the sample holding a bachelor's degree or a graduate degree.

4.3 Twitter Dataset

A total of 13.7 million Twitter messages were collected from Oct. 22 to Nov. 7, 2012 using Firehose streaming API via GNIP [47]. The raw data were indexed and inserted into a distributed NoSQL (MongoDB) database for storage. This database serves as the central repository of data for all subsequent analyses. We created two datasets based on: 1) Keyword – Twitter messages matching a set of Sandy-related terms comprised of keywords, hashtags, and user names (9.3 million Twitter messages); and 2) Geo-tagged – Twitter messages from New York, New Jersey, and Connecticut (4.4 million Twitter messages). Both datasets (keyword and geo-tagged) were further divided into three temporal phases: 1) Pre-hurricane (10/22/2012 – 10/28/2012), 2) During-hurricane (10/29/2012 – 10/31/2012), and 3) Post-hurricane (11/01/2012 – 11/07/2012).

4.4 Temporal Analysis

Identifying temporal trends during a disaster is an important method to determine the tracking system of tweets over different phases of the disaster [48,49]. The collected datasets contained metadata attributes of the time at which the messages were posted on the Twitter network. The time-stamp attribute has a resolution of milliseconds in relation to the GMT time zone. Analysis on the keyword and the geo-tagged datasets were conducted to illustrate the *peaks* and *valleys* in the data in order to better understand the involvement of Twitter users in discussions about Hurricane Sandy during the three identified phases. The data were aggregated by number of messages per-hour and the number of unique users' per-hour to visualize the results in the temporal zones.

4.5 Klout Analysis

Each user captured by the data, contains the metadata attribute of Klout score. Klout score is a metric to measure influence of users on online social networks [50]. Klout score of a user is measured based on three components including: true reach, which measures how many people a user influences; amplification, which refers to how much the user influences them; and network impact that measures the influence of the user's network [51]. This score for a Twitter user is a numerical value from 1 to 100, with a higher score representing a higher level of influence. It is based on the size of user's social network (friends, followers) and correlates with the reactions to the user's posting by other Twitter users. In this research, the scores pertaining to individual users were aggregated by hour to understand the involvement of influential users in the discussions about Hurricane Sandy. The data were then compared across the temporal zones for both the keyword and the geo-tagged datasets. It should be noted that the Klout scores were collected after the Hurricane Sandy and there might have been changes in the scores during that time. The Klout score service provider was shut down later and there is not a way to verify the information.

4.6 Text Analysis

The contents of the Twitter messages were analyzed using word cloud and word collation/co-existence in order to understand the key-topics that were discussed in the temporal phases of the study. The text contents of the tweets for the different phases of the hurricane from both datasets were extracted. The text was then processed for cleanup (removal of stopwords, hyperlinks, common terms), and word stemming was utilized for abbreviations in the short messages. Word importance weights were calculated using Term Frequency – Inverse Document Frequency – TF-IDF [52] for individual words, where each word's weight is an importance metric calculated across the corpus of all documents (in our case tweets). The values were represented in word clouds for both keyword and geo-tagged datasets.

Word co-existence analysis was done by examining existence of bi-grams of words present in Twitter messages. For all pairs of collated words, a graph was constructed using the word pairs as nodes in the graph with edges/links denoting a co-existence connection. Weights were used for the nodes and edges representing the repeated occurrence of entities and pairs. The resulting network was analyzed and visualized for connectedness (degree), influence (betweenness centrality) and community memberships. Louvain modularity [53] was used for detection of communities of words. The method utilizes recursive grouping of smaller communities by representing them as nodes and evaluating how dense the connections are while maximizing/optimizing for modularity. The high positive value of modularity measure indicates the presence of community structure with dense connected nodes within the partition sets [54]. The network graph was developed using Python, and the analysis of the graph was done using Gephi, an open source software for network analysis and visualization [55].

4.7 Sentiment Analysis

Sentiment analysis is an ongoing field of research determining the positive or negative opinion of people about a text message, a specific entity, or a topic in general [36]. Support vector machine (SVM), a common machine learning method for classification [12,56,57] was used to determine polarity of the messages (positive, negative, or neutral) sent during the three phases of Hurricane Sandy. The process includes various techniques including text cleanup (stop-words removal, URL removal, emoticon filtering), feature-extraction (context features, natural language processing, rare word filter, stemming), and vectorization of text. The model used for the analysis of the present datasets is based on a trained dataset of approximately 4.2 million codified tweets in each category, and is purpose-built for codifying short text messages such as tweets with an accuracy of 84% (10-fold cross-validation with 2.1 million codified tweets). The model also resulted in 81% accuracy in a double-blind verification by human coders. The sentiment model assigns each tweet a value between -2.0 to +2.0, where -2.0 to -1.0 is codified as negative, -0.99 to +0.99 is codified as neutral and +1 to +2 is codified as positive. Both the keyword and geo-tagged datasets are codified with sentiment values for each tweet and average sentiment values are aggregated by hour.

4.8 Social Network Analysis

Social network analysis uses the mathematics of graph theory to understand social phenomenon through the relationship among people, groups, and things [58]. A social network is represented by a graph of nodes and links, where nodes are individual actors and links are social ties, relationships, exchanges, or interactions among actors [59]. The user network analysis was performed based on the metadata of user mentions in Twitter Messages for keyword dataset. Twitter users can mention other users by posting a message using the format of '@username' to reference a particular user or reply to another user's tweet. A directed graph was created for each mention of the user and edges were added from the originating user to the mentioned user. Different network metrics (for each stage of the hurricane) were calculated including average degree, average weighted degree, graph density, network diameter, centrality (in degree, out degree, betweenness, eigenvector), average clustering coefficient, and average path length. A full description of these measures can be found in [60]. Louvain method [53] in Gephi was used to calculate modularity and divide the graph into groups of users who are well connected or clustered together in the graph known as communities. The clusters were visualized by Force Atlas algorithm [61].

5. Results

5.1 Research objective 1: Identify the sources of information received by and shared with the coastal areas' residents affected by Hurricane Sandy.

The telephone and web-based survey instruments were used to understand the use of Twitter as a communication platform during Hurricane Sandy. The following three research questions are addressed from the survey: 1) How did people obtain information during Hurricane Sandy? 2) From whom did people obtain information during Hurricane Sandy and how did these sources differ between Twitter users and non-Twitter users? and 3) What type of information did Twitter users share during Hurricane Sandy?

5.1.1 Telephone and Web-based Survey Results

In both the telephone and web-based surveys, individuals were asked whether or not they sought information through ten different mediums. Of particular interest was the difference between Twitter and non-Twitter users in receiving information. Utilizing Chi-Square test, we found statistically significant differences between Twitter users and non-Twitter users on four of the mediums including text message with X^2 (1, N = 667) = 19.67 and $p < .001$; Internet (non-social media) with X^2 (1, N = 667) = 44.99 and $p < .001$; cellphone weather apps with X^2 (1, N = 667) = 39.141 and $p < .001$; and social media other than Twitter with X^2 (1, N = 667) = 50.85 and $p < .001$. Compared to non-Twitter users, a greater percentage of Twitter users reported receiving weather-related information via the four modalities listed above. Within the full sample (including both Twitter users and non-Twitter users), there exists a difference in ways of receiving information when taking into account the loss of electrical power. Television was the major source of information for 82% of our respondents with power, which is consistence with past findings [2,62]. Internet and radio were the other main sources of receiving storm-related information for 32% and 29% of the sample respectively (Figure 4). Only 2% of respondents received the information through Twitter. Similar results were observed in a study by Feldman et al., [2] confirming the low usage of Twitter (2.3) in communications related to flood risk for 164 residents of Newport Beach, California. However, Twitter was more desired as a medium of future communications among these residents. While Feldman et al., [2] did not consider the impact of power outage on the use of Twitter, our analysis shows how the respondents who lost power during Hurricane Sandy reported having received storm-related information (Figure 5). In our study, individuals who lost access to electrical power, moved away from their reliance on television for receiving weather-related information and toward a reliance on radio, telephone, cellphones (through both text messages and weather apps), and the Internet, including social media. Interestingly, 18% of respondents without power used Twitter as their information source.

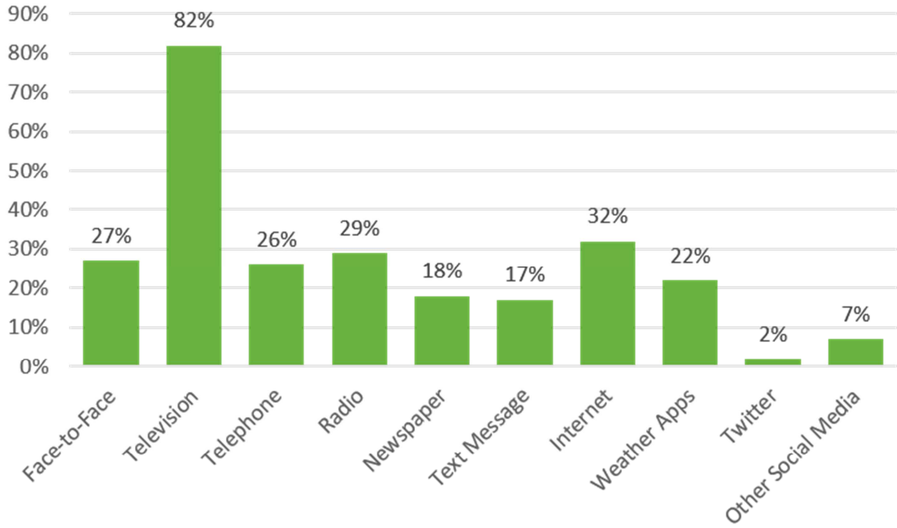

Figure 4. Information sources for respondents with power (*n*=205).

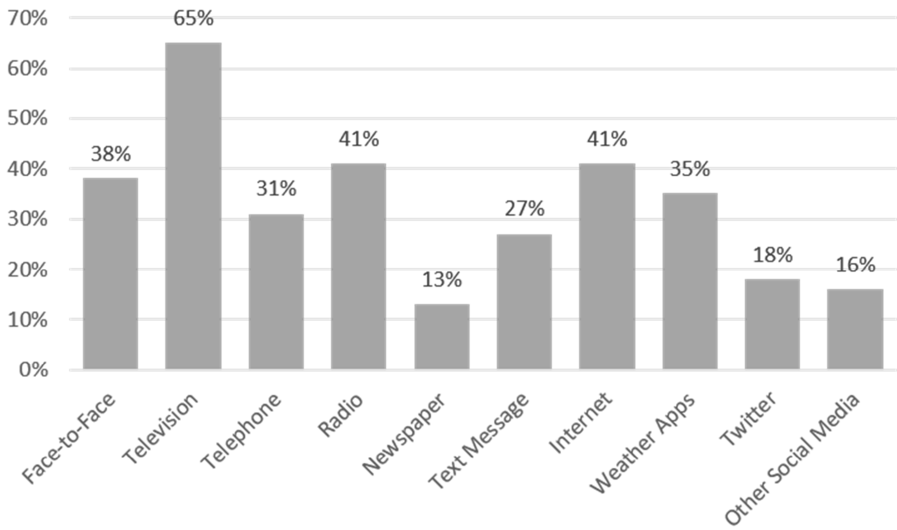

Figure 5. Information sources for respondents without power (*n*=414).

Respondents were also asked whether or not they received information from any of seven different sources including friends, family, household member, local news, national news, federal agencies, and state agencies. Figures 6 and 7 represent the sources of weather-related information for Twitter users and non-Twitter users, respectively.

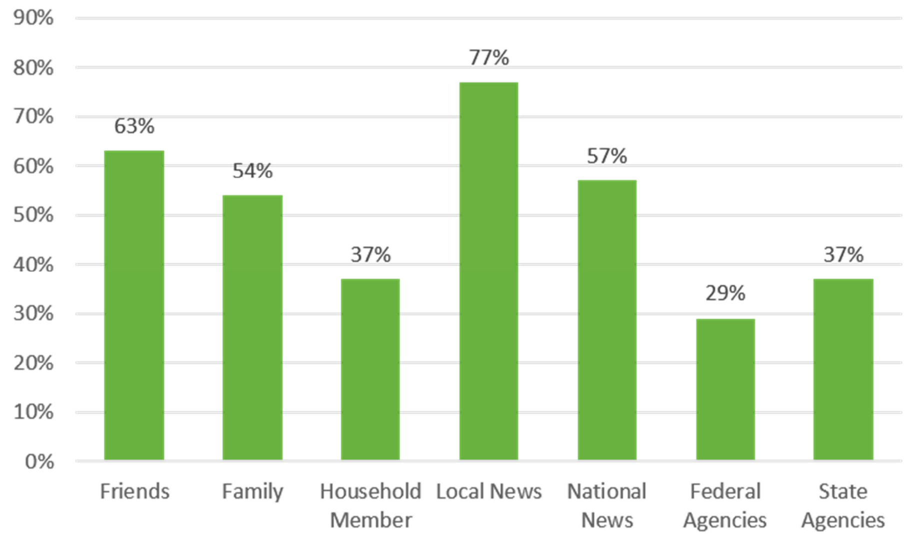

Figure 6. Weather-related information sources for Twitter users (*n*=266).

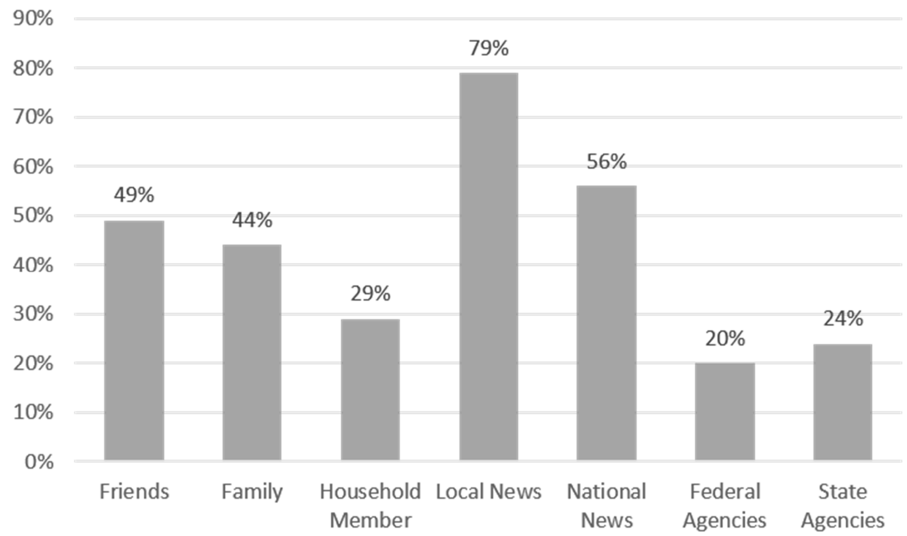

Figure 7. Weather-related information sources for non-Twitter users (*n*=420).

The positive impact of news media engagement on Twitter activities during disruptive events has been identified in literature [63]. While both Twitter users and non-Twitter users in our study relied on local and national news at about the same frequency, Twitter users reported having received information from family, friends, and government agencies at a higher rate than non-Twitter users. This suggests Twitter users depend more on information accessed via their family, friends [64], and government agencies when compared to non-Twitter users. One potential application of Twitter data, therefore, is that government agencies can utilize Twitter during emergency situations and, by doing so, reach at least a subset of the population most at risk during weather-related emergencies. However, our social network analysis shows (see section 5.3.1) the bi-directional communication does not usually happen during emergencies. A secondary analysis relating to this research question examined what sources the Twitter users followed on Twitter (Figure 8). While Twitter users follow a myriad of weather-related sources, the three top sources were: 1) Local Television News, 2) The National Weather Service, and 3) The Weather Channel.

For our third question, approximately 59% of Twitter users in the survey reported that they shared weather-related information via Twitter during Hurricane Sandy. Figure 9 illustrates the types of information these users shared. The data indicate that photographs were the most frequently shared form of information (62%) by Twitter users, followed (in frequency) by personal experiences (56%), and information about storm damage (53%). It is not necessarily the case that the categories of information Twitter users shared are mutually exclusive, meaning respondents may have posted a tweet sharing their personal experiences with the storm and attached an image to that tweet. Regardless, these findings illustrate Twitter's utility as an image-sharing platform.

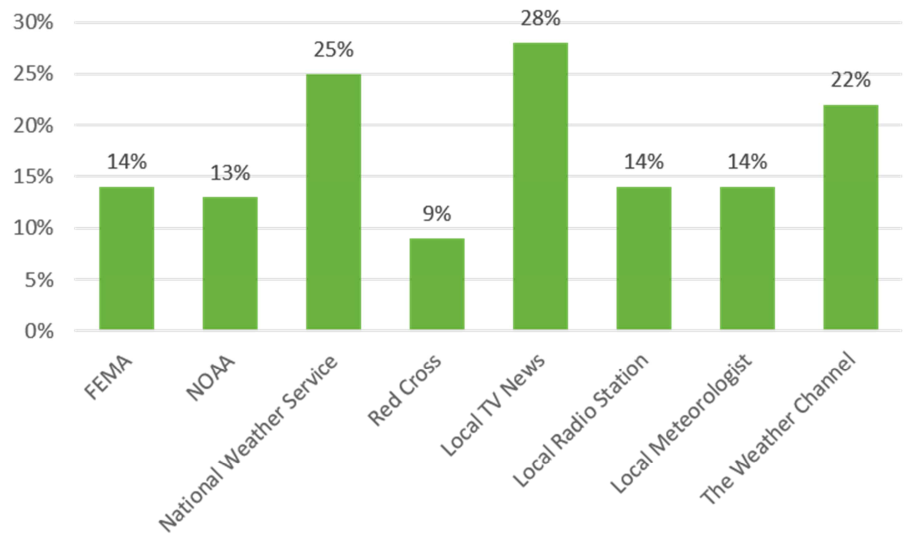

Figure 8. Weather-related Twitter accounts followed by survey respondents (*n*=266).

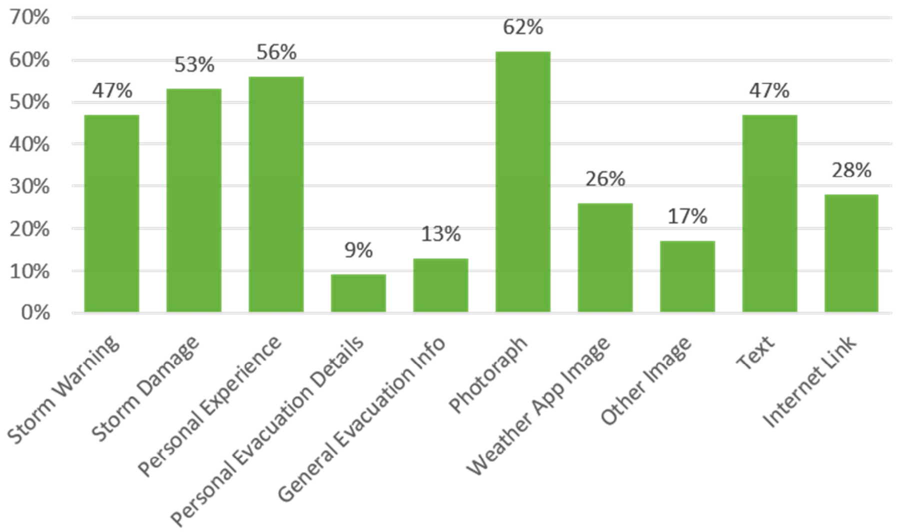

Figure 9. Types of information Twitter users shared during Hurricane Sandy (*n*=156).

5.2 Research objective 2: Examine the Twitter users' involvement in discussions about Hurricane Sandy and the content of messages shared before, during, and after the hurricane.

Twitter data analysis addressed the following three research questions: 1) How did the participation rate of Twitter users in discussions about Hurricane Sandy change during the three phases of analysis? 2) What were the top words posted by Twitter users and how did they change during the three phases? 3) How did the positive/negative sentiments of shared messages change over the three phases?

5.2.1 Temporal Analysis

The temporal analysis of the keyword and the geo-tagged datasets shows that the results are in line with past research that use of social media in a disaster starts very early [65], and reach its peak mostly while it is happening [66]. Analysis of the keyword dataset (Figure 10) shows that there was a substantial increase in the number of messages and unique users contributing to the hurricane discussions from Oct 26, 2012 with the highest peak occurring on Oct 29, 2012 (approximately 237 K unique messages being shared per hour) at 6:00pm EST. At the time Hurricane Sandy made landfall (8:00pm EST in Atlantic City), approximately 223 K unique messages were being shared across the network by 187 K unique users per hour. In the following days, both during and post-Hurricane Sandy, the number of Twitter messages along with the number of users decreased over time. The post-hurricane phase revealed a larger number of messages being shared across the network in comparison to the pre-hurricane period. This is due to large frequency of relief-related tweets in the aftermath of landfall. For example, the increase in the Twitter users and tweets on November 2 was related to NBC's live telethon, 'Hurricane Sandy: Coming Together' encouraging Twitter users to live tweet using the hashtag #SandyHelp [48]. The general patterns observed in this study are in consistent with other Sandy-related studies [48,49] showing the accuracy of our collected data and analysis.

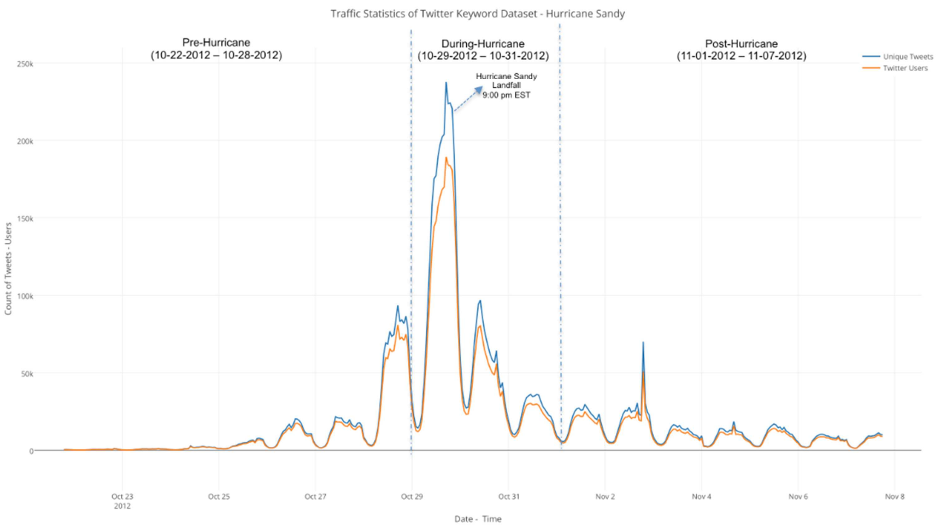

Figure 10. Temporal analysis of keyword dataset aggregated by hour

Figure 11 shows the temporal dynamics of the geo-tagged messages and their corresponding users collected from the New York, New Jersey and Connecticut area. In comparison to the keyword dataset, the geo-tagged traffic shows a large number of peaks in all three phases of the hurricane. During the pre-hurricane phase, the number of tweets shared from the locations decreased on Oct. 28 and then gradually rose to the peak during landfall at 8:00pm EST on Oct. 29. The traffic then gradually decreased over the following days and peaked again on Nov. 7, which was a result of President Obama's re-election. These messages provide valuable information from individuals residing in the hurricane-affected areas. More specifically, the messages contain firsthand information about the hurricane, along with disaster-related images taken in real-time (Figure 12). A large number of posts contained weather-specific photographs showing the intensity of the hurricane in real-time, along with images of damage and flooding through which researchers and emergency managers can retrieve information to help identify storm damage and plan relief efforts. With the growing number of Twitter users particularly during power outages as our study shows, this information can be useful for a better management.

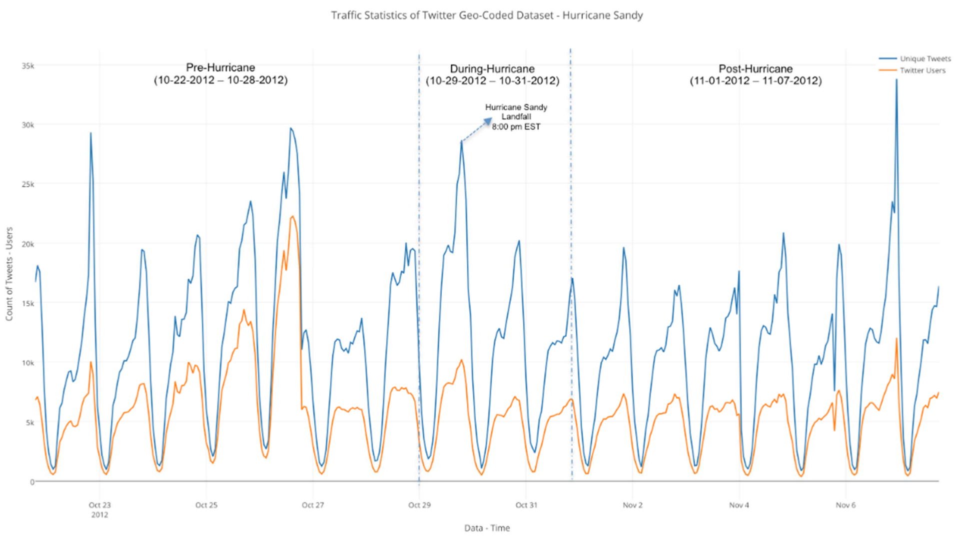

Figure 11. Temporal analysis of geo-tagged tweets aggregated by hour

Figure 12. Photos in New Jersey (a) and New York (b) shared by two Twitter users on October 30

5.2.2 Klout Analysis

The analysis of the aggregated Klout scores shows a decrease in the average score per hour nearing landfall (Figures 13 & 14). This indicates an increase in participation of the general population (lower Klout Scores in comparison to influential users) in the discussions leading up to, and following, Hurricane Sandy. While influential users with high Klout scores also participated in the discussions, the general population were more active in sending messages that were being shared across the network. The analysis also shows that for the geo-tagged dataset, users have a much lower average Klout score than the keyword dataset users. This may indicate that the majority of the general population have their geo-location services enabled on their mobile devices. While users with higher scores have been identified as more effective in spreading the information [50], geo-tagged tweets shared by general public might be more valuable for identifying the most affected people and areas for a better disaster management.

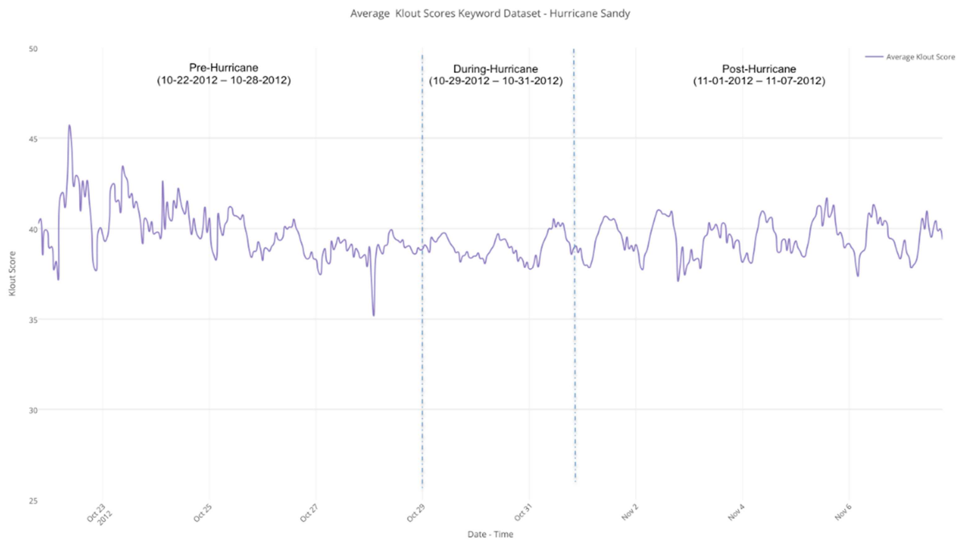

Figure 13. Average Klout scores aggregated by hour in keyword dataset

Figure 14. Average Klout scores aggregated by hour in geo-tagged dataset

5.2.3 Text Analysis

The word cloud analysis shows clear differences among the words used before, during, and after the hurricane, a result confirmed in other disaster-related studies [14]. The level of hazards and risk, and trend of the incident can be identified through word analysis [14]. In the keyword dataset (Figure 15), analysis of the word distributions reveals that the most frequently occurring words in the pre-hurricane phase were: *Sandy, Hurricane, Frankenstorm, Storm, New York, Coming*, and *Tomorrow*. In comparison, the most frequently occurring words during hurricane were: *Sandy, Hurricane, Power, HurricaneSandy, Safe, Stay Safe, East, Prayer*, and *Good*. In the post-hurricane phase, the most frequently occurring words were: *Help, Relief, Sandy, Hurricane, New York, SandyHelp, Aftermath*, and *Power*. The analysis shows the transition of discussion across the different phases of the hurricane from people advising and spreading the news of the hurricane → Twitter users being concerned about the well-being of their friends and followers → relief and rescue efforts in the aftermath of the hurricane. The results are paralleled with another study by Spence et al., [67] examining the content of tweets during Hurricane Sandy. The discussion of power outages also occurred in the during- and post-phases of the hurricane. The geo-tagged dataset presents similar results with one key difference in the post-hurricane phase, where the discussion was more of an everyday social interaction along with some discussions on President Obama's re-election (Figure 16).

Pre-hurricane phase During-hurricane phase Post-hurricane phase

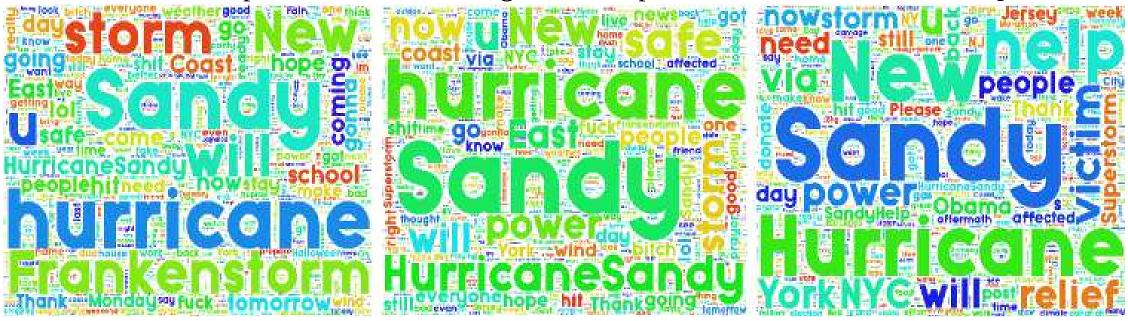

Figure 15. Word clouds in pre, during and post hurricane phases (keyword dataset)

Pre-hurricane phase During-hurricane phase Post-hurricane phase

Figure 16. Word clouds in pre, during and post hurricane phases (geo-tagged dataset)

The word co-exsistence analysis shows that the words *hurricane* and *sandy* were the top ranking words in the pre-phase (Figure 17). They were connected by less occurring words such as *frankenstorm*, *storm*, *hurricanesandy*, *school*, and *monday*. The during-phase (Figure 18) shows a substantial change in the words that were tweeted. Words that formed clusters included: [*stay, safe, strong*], [*friends, family, share*], [*New, York, Jersey, Atlantic*], [*prayers, thoughts*], and [*power, out, home*]. The increasing use of more emotional tweets in this phase might be a good mechanism to release anxiety or stress, return people to a normal state, and provide a sense of connection among those who experience the similar situation [68]. In the post-phase (Figure 19), clusters related to the aftermath [*power, still, out*], relief [*food, water*], [*help, relief, sandyhelp*] and donation [*redcross, donate*] formed the main topics of discussion.

With the sheer volume of data shared during emergencies, one key problem is how to select proper information at different stages. Looking for keywords such as *sandy* or *hurricane* will result in a large dataset that might not be helpful. A closer look at the frequency and coexistence of words can help us to identify specific keywords to search for at different stages of emergencies. For example, the keywords *please*, *need*, *flooding*, and *power* had a higher frequency in the during-phase of Hurricane Sandy (Figure20). The word *victim*, on the other hand, had a higher frequency in the post-phase. These keywords might be useful to locate the event and people in need during disasters. Another useful information are Tweets containing the word *evacuate*. These tweets were mostly started on 28th and increased sharply during 29th when the mandatory evacuation order was announced. These data in combination with geo-tagged information can help in understanding the evacuation behavior of residents. The coexistence analysis also revealed valuable information. The word *power*, for example coexisted with words such as *outages*, *knock*, *lost*, *lose*, *cuts*, *downed*, *dead*, *nuclear*, *plants*, *customers*, *homes*, and *still*, in the during-phase of Hurricane Sandy. These words might be appropriate search keywords for power-related companies to identify affected areas. In the post-phase, the words *food*, *water*, *shelter*, *clothes*, *distribution*, *drive*, *truck*, and their co-existence with words such as *victims* and *need* not only show useful keywords to search for, but also are valuable for disaster-related agencies to identify victims and provide their urgent needs in a timely manner. The co-existence of the words *gas*, *lines*, *stations*, *situation*, *problems*, *shortage*, and *long* shows other useful keywords that can help to identify gas-related problems in specific areas.

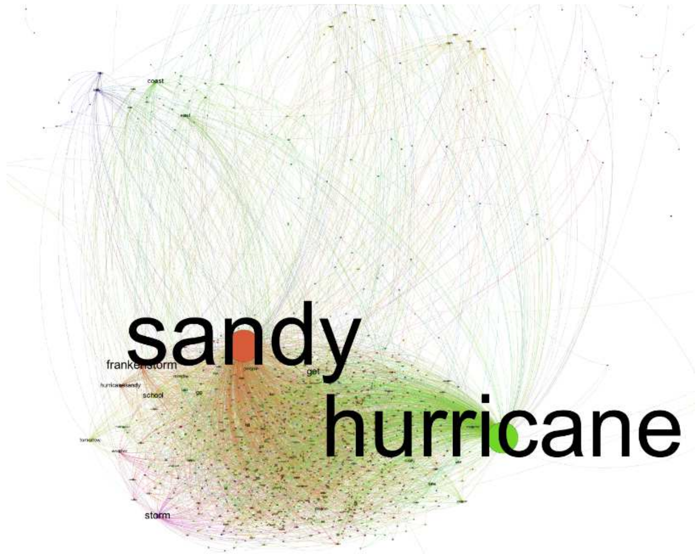

Figure 17. Word co-existence: pre-hurricane phase (keyword dataset)

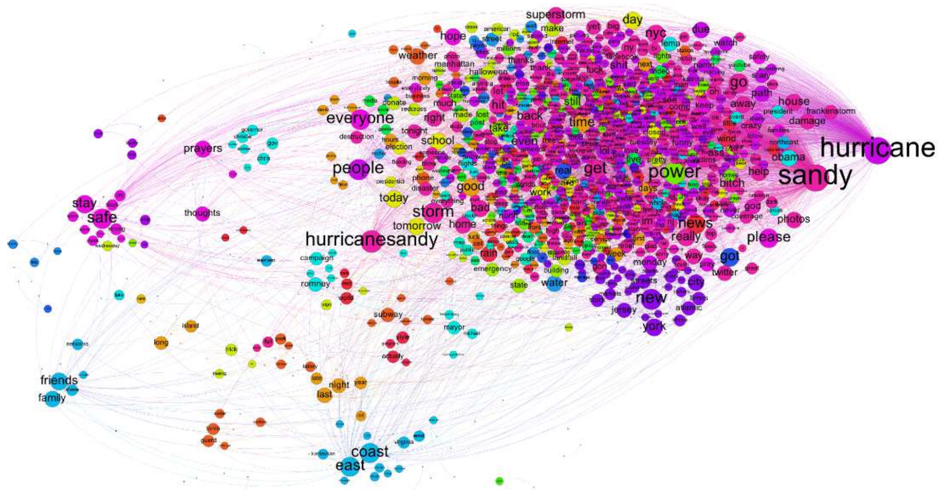

Figure 18. Word co-existence: during-hurricane phase (keyword dataset)

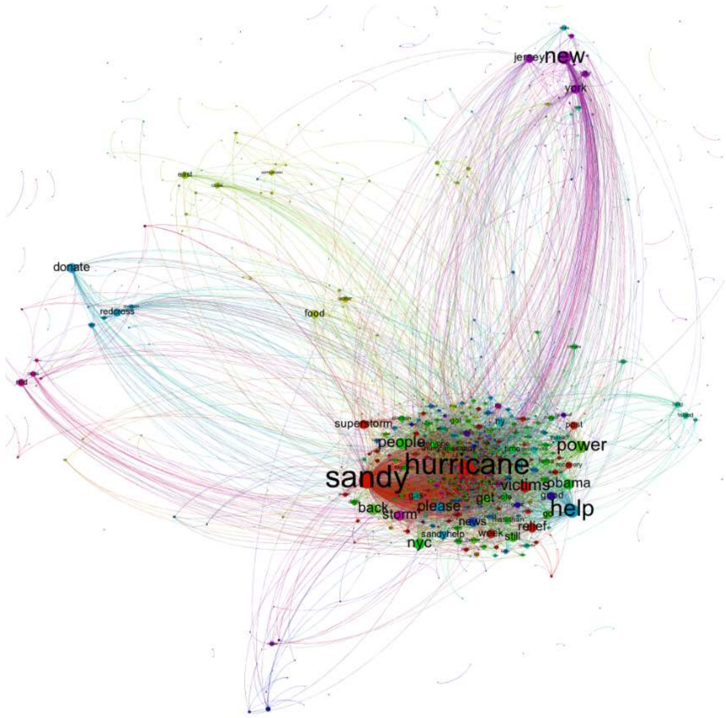

Figure 19. Word co-existence: post-hurricane phase (keyword dataset)

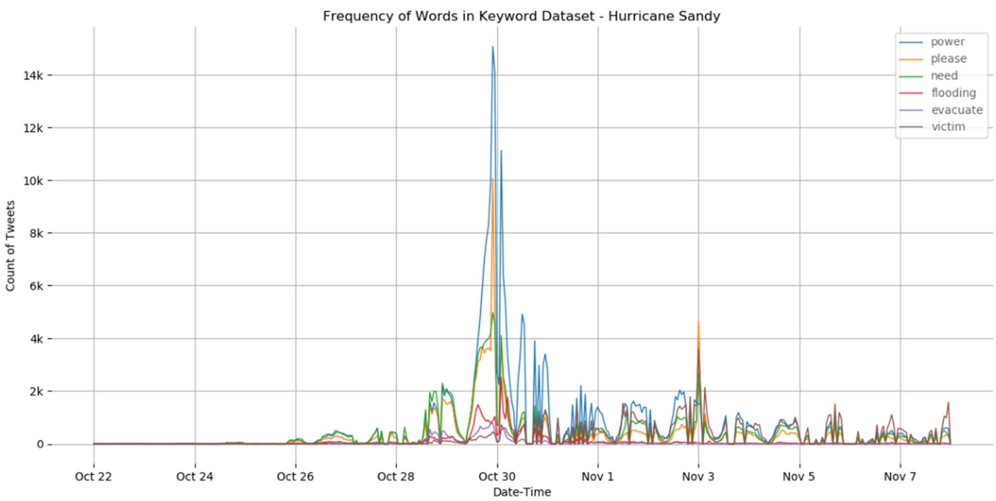

Figure 20. Frequency of words aggregated by hour (keyword dataset)

5.2.4 Sentiment Analysis

Sentiment analysis of different phases of the study (Figures 21 & 22) reveals the messages shared by Twitter users were more negative in the during-hurricane phase of the study. The pre-hurricane phase of the study in both datasets displayed a more positive sentiment, while gradually decreasing toward the mid-point of the during-hurricane phase (lowest sentiment average scores) and then increasing in the post-hurricane period. This shows the dynamics of Twitter users who were posting messages with an increasingly negative attitude to Hurricane Sandy at the peak of the storm. The comparison of the keyword and geo-tagged datasets also reveals that the sentiment of Twitter messages originating from the Hurricane Sandy areas (geo-tagged dataset score range between -0.2 to +0.5) are more negative than the keyword dataset (score range between +0.1 to +0.5) suggesting the influence of distance (from disaster area) on sentiments of tweets. While sentiment analysis is a valuable approach to detect and locate disasters [69], it is not fully used by authorities. This is due to the inability to efficiently sort and categorize the sheer volume of data generated during disasters [66]. Our analysis is a confirmation of the concept that for a better utilization of social media data during natural disasters and this type of analysis, authorities should mostly use the negative tweets in the during-phase of disaster, when more people use social media to communicate their needs. Selecting these negative tweets would be more helpful to detect and locate people at risk in a timely manner rather than focusing on all tweets. Evaluating negative tweets in time and space and integrating them into a system that use various modalities such as text and network analysis remains for future research.

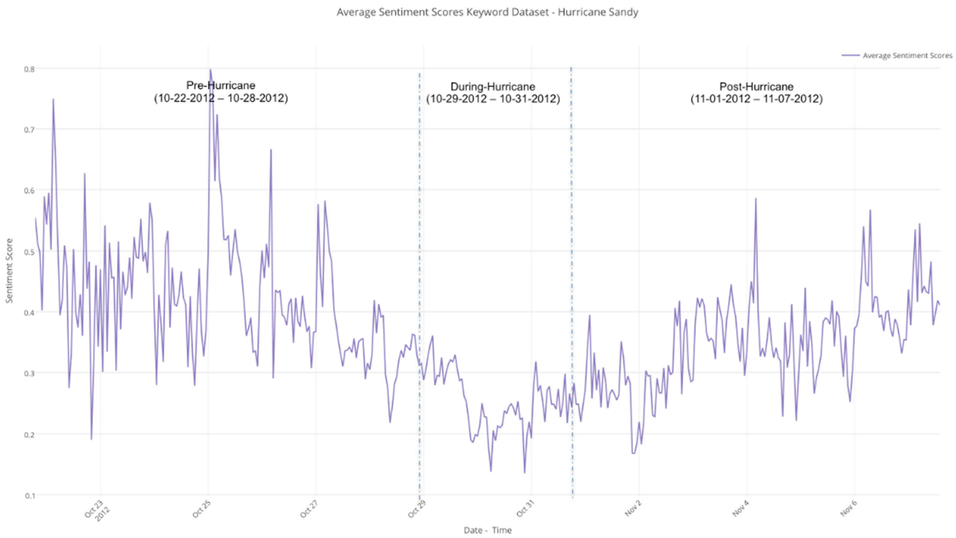

Figure 21. Average sentiment score aggregated by hour before, during and after hurricane (keyword dataset)

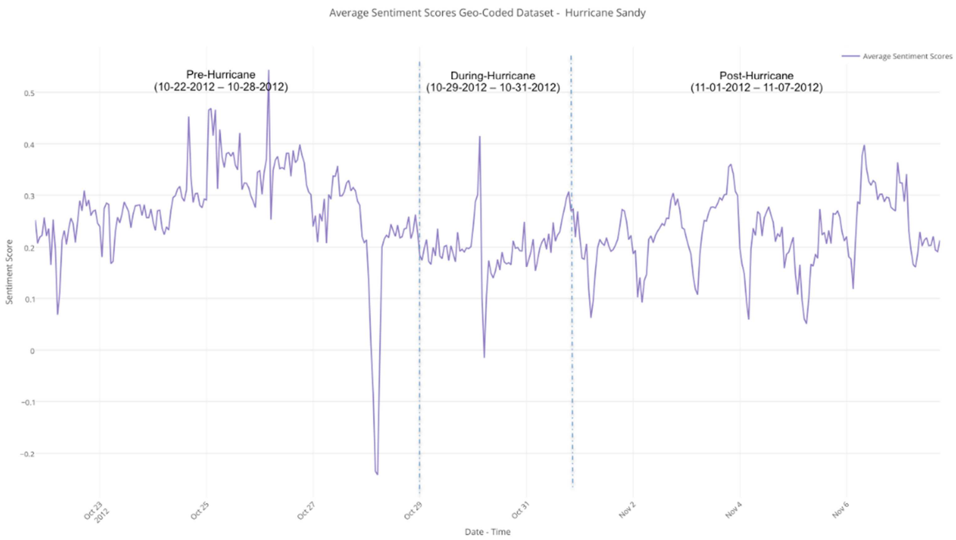

Figure 22. Average sentiment score aggregated by hour before, during and after hurricane (geo-tagged dataset)

5.3 Research objective 3: Examine the social network structure of Twitter users before, during, and after the hurricane.

Network analysis of the users addressed the following three research questions: 1) How did the network structure of Twitter users evolve during the three phases of Hurricane Sandy? 2) How did the users' communities form during the three phases? 3) Who were the key influencers across the networks?

5.3.1 Social Network Analysis

We used users' networks to analyze the interconnectivity of Twitter users during each phase of the hurricane. Network data statistics are shown in Table 3. The highest numbers of nodes and edges were observed in the during-hurricane phase representing a larger number of active users communicating during the hurricane. The larger network diameter in the during-phase also supports the analysis with larger participants increasing the size of the connected network. Average degree and weighted degree are larger in the post-hurricane phase indicating stronger bi-directional communication in the relief efforts. In-degree (the number of incoming edges) and out-degree (the number of outgoing edges) centrality measures represent the prominent and influential users in a network respectively [70]. In our study, news agencies, political figures, weather, and other disaster-related agencies and organizations (e.g., FEMA, Red Cross) were the users with top in-degree centrality (Tables 4). However, the low out-degree measure of these users show the one-way communication between these users and general public. While many people mentioned disaster-related agencies such as FEMA, these users were not much responsive. Similar results were observed for the users with the highest eigenvector centrality (Table 6). The assumption behind the eigenvector centrality is that the centrality of a node is proportional to the sum of the centralities of its neighbors, therefore it describes how well the node in the network is connected to other well-connected nodes [71]. This suggests the central/influential roles of identified users in the overall structure of the network [70] as potential sources of disaster-related information and social engagement. However, the users spreading the information were mostly among public figures as shown by out-degree measures (Table 5), a result observed in other disaster-related studies [14]. These public figures were engaged with approximately 100-200 other users, a much lower number compared to the number of users who mentioned top in-degree users. Interestingly, disaster-related agencies such as FEMA were not among the top out-degree users, while the public expectation is to receive more information from these users. Betweenness centrality measures the extent to which a node in the network lies between other nodes [71]. Our results for the top betweenness centrality (Tables 6) include political and public figures, weather agencies, and other disaster-related agencies and organizations suggesting their role as bridges in the communication network. Users with high betweenness centrality in a network are also called gatekeepers since they control how information flow between communities [71]. Public and political figures can play a significant role during disasters. Authorities can benefit from connecting to these users to spread relevant information and receive more information on the people's need, but our analysis did not find these connections.

Table 3. Network statistics of Twitter users (keyword dataset)

	Pre-hurricane	**During-hurricane**	**Post-hurricane**
Network Type	Directed	Directed	Directed
Nodes	10,788	50,607	32,879
Edges	36,269	234,000	176,396
Average Degree	3.362	4.624	5.365
Average Weighted Degree	4.961	6.269	8.284
Network diameter	17	30	18
Graph Density	0.000	0.000	0.000
Average Path Length	5.472	6.683	5.605
Modularity	0.571	0.502	0.521
Average Clustering Coefficient	0.057	0.045	0.069

Table 4. Users with top in-degree measure and their associated out-degree

Pre-hurricane			**During-hurricane**			**Post-hurricane**		
	In-	Out-		In-	Out-		In-	Out-
MikeBloomberg	599	13	GovChristie	4312	0	GovChristie	5140	27
fema	467	26	MikeBloomberg	2636	20	MikeBloomberg	4365	14
twc_hurricane	458	47	NYGovCuomo	2468	55	RedCross	3731	3
NHC_Atlantic	420	0	CoryBooker	2116	93	CoryBooker	3299	201
GovChristie	396	7	NYCMayorsOffice	1953	20	nytimes	3125	0
NYCMayorsOffice	396	18	twc_hurricane	1933	117	ABC	2961	0
weatherchannel	384	17	RedCross	1929	11	NYGovCuomo	2798	67
AP	376	0	BarackObama	1856	0	FoxNews	2791	0
CoryBooker	367	67	AP	1653	3	CBS	2428	0
JimCantore	359	5	cnnbrk	1499	0	NPR	2328	0

Table 5. Users with top out-degree measure and their associated in-degree

Pre-hurricane			**During-hurricane**			**Post-hurricane**		
	Out-	In-		Out-	In-		Out-	In-
editchick	102	0	rightnowio_feed	203	10	CoryBooker	201	3299
weatherplaza	95	6	SeanPCollins	183	16	farside314	181	7
MarnieTWC	73	25	weeddude	180	101	BlondeVelvet	163	6

JustCouch	71	5	ScottBeale	150	64	blogdiva	158	78
CoryBooker	67	367	JustCouch	146	0	wishuponahero	154	17
blogdiva	64	23	ninatypewriter	143	53	GrandmaJer_ETSY	134	0
sahnetaeter	62	6	DAKGirl	142	11	HealthcareWen	115	40
HumanityRoad	58	30	farside314	137	0	EarlyShares	114	3
trdelancy	58	6	azipaybarah	132	80	ConEdison	104	1770
weeddude	57	14	SustainablDylan	122	0	SINYCliving	100	114

Table 6. Users with top Betweenness and Eigenvector measures

Pre-hurricane		**During-hurricane**		**Post-hurricane**	
Betweenness	**Eigenvector**	**Betweenness**	**Eigenvector**	**Betweenness**	**Eigenvector**
twc_hurricane	MikeBloomberg	FDNY	GovChristie	CoryBooker	GovChristie
CoryBooker	NHC_Atlantic	twc_hurricane	MikeBloomberg	ConEdison	MikeBloomberg
WSJweather	fema	CoryBooker	NYGovCuomo	fema	RedCross
garytx	twc_hurricane	NYGovCuomo	NYCMayors*	MikeBloomberg	CoryBooker
NYGovCuomo	weatherchannel	RedCross	CoryBooker	FDNY	NYGovCuomo
fema	NYCMayors*	AntDeRosa	RedCross	NYGovCuomo	nytimes
RyanMaue	RedCross	rqskye	twc_hurricane	GovChristie	ABC
weatherchannel	JimCantore	fema	BarackObama	blogdiva	FoxNews
JimCantore	NOAA	ConEdison	AP	RedCross	NYCMayors*
USCG	GovChristie	granthansen	fema	JCP_L	BarackObama

* NYCMayorsOffice

We identified and visualized the user community clusters using the Louvian modularity measure. The nodes represent Twitter users and the lines between the nodes represent the connectivity by the attribute of mention. The size of the nodes (and the size of username) represents the number of times a particular user was mentioned. The varying colors of the nodes and the edges represent the various clusters. In the pre-hurricane phase (Figure 23), political Twitter users such as, MikeBloomberg, NYGoverner, NYCMayorsOffice, CoryBooker clustered together. Whereas news agencies (NHC_Atlantic, breakingstorm, wunderground, twc_hurricane, weather_channel, ABCnews) and federal agencies (FEMA, NOAA) each form separate clusters. RedCross and GovChristie have their own influence groups, which are separate from the other clusters. During hurricane (Figure 24), most of the users clustered together, suggesting a dense network of communication with information being shared in a bi-directional model across the platform. GovChristie achieved the highest number of mentions, followed by other political Twitter users (BarakObama, NYCMayorsOffice, NYGovCuomo, MikeBloomberg, CoryBooker). FEMA and RedCross gravitated closer to news agencies such as HuffPost, nytimes, AP, cnnbrk. The formation of a blue circular cluster (on the right side of the data structure) proved to be noise unrelated to the hurricane. This cluster represents a high volume of communication between two United Kingdom musical bands and their producers' discussions about a song titled "Sandy." In the post-phase of the hurricane (Figure 25), users clustered near the center of the graph, suggesting that most of the centroid users where highly connected to other users in the network. RedCross was highly mentioned as a result of its involvement in the relief efforts.

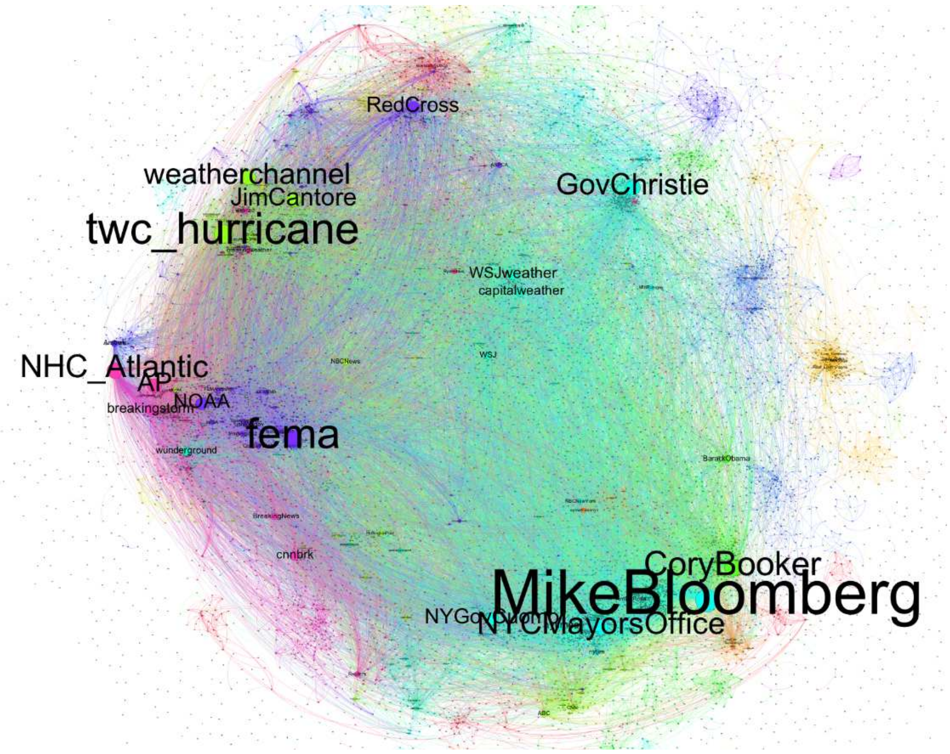

Figure 23. User network analysis: pre-hurricane phase (keyword dataset)

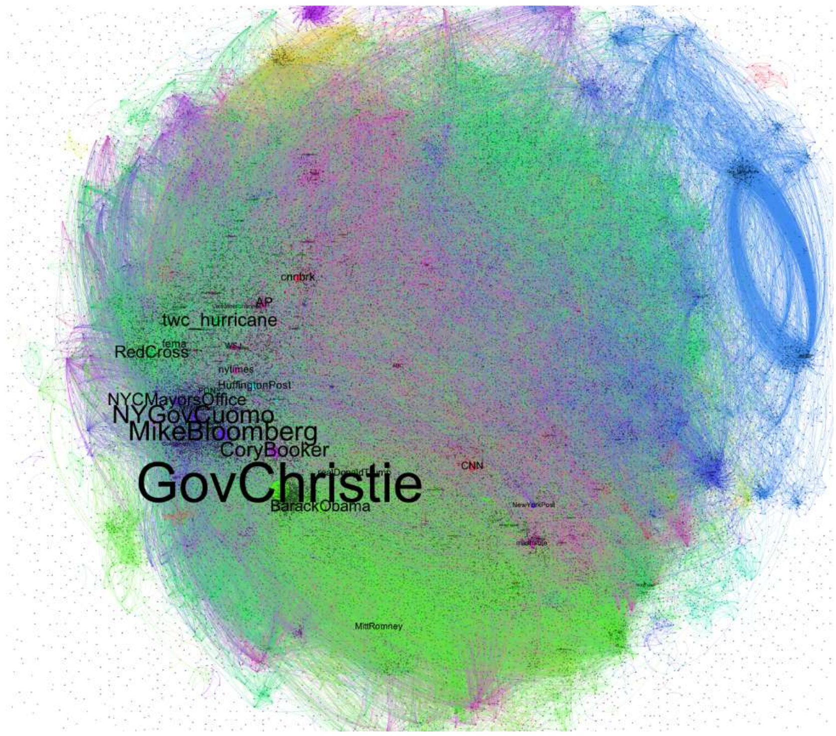

Figure 24. User network analysis: during-hurricane phase (keyword dataset)

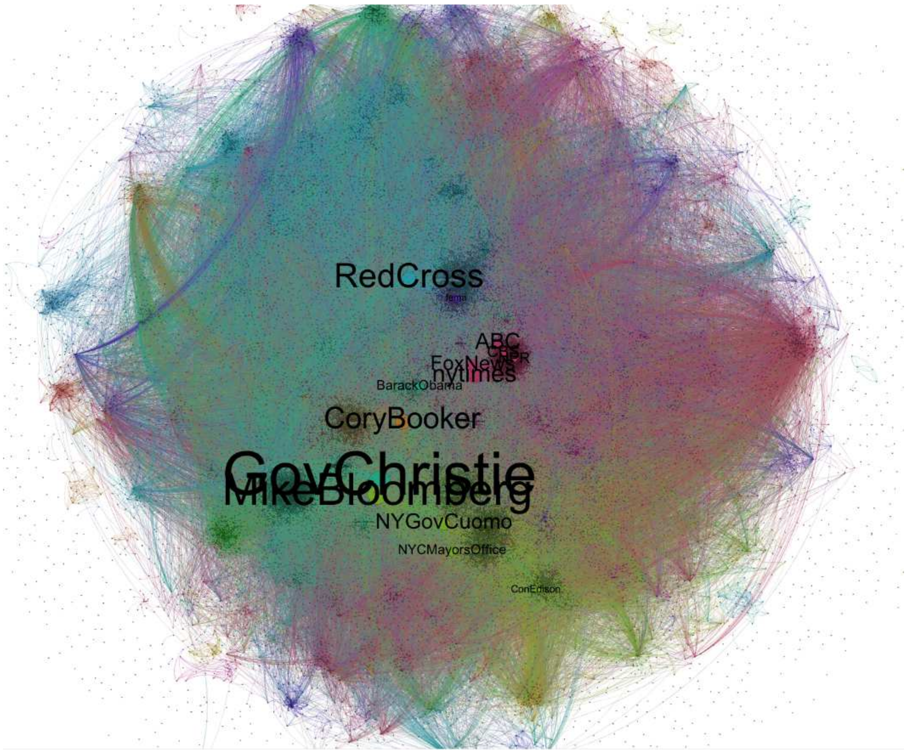

Figure 25. User network analysis: post-hurricane phase (keyword dataset)

6. Discussion and Conclusions

We examined the usage of social media during Hurricane Sandy from both survey and communication perspective. Research has validated the prominent role of social media, particularly Twitter as a source of information [72,73]. Limitations of other media sources in providing useful information during the disaster is one of the reasons for using social media [27]. The results of our study indicate that Twitter can serve as a valuable medium for communication through sharing texts and photos during weather-related emergencies, especially during power outages time. Our findings also indicate that Twitter users generally receive emergency information from various sources at higher rates than non-Twitter users. This is due to the easier access of Twitter users to prominent individuals, organizations, and agencies who share weather-related information as our social network analysis reveals. Temporal analysis of tweets showed that use of Twitter during Hurricane Sandy started very early and continued in the aftermath of disaster. With the change of discussions in different phases of hurricane from people advising and spreading the news, to concerns about the well-being of their friends and followers, and to relief and rescue efforts in the aftermath of the hurricane. The emergency officials can use such messages as a valuable source of information from individuals residing in the hurricane-affected areas. Lower sentiment scores (negative polarity) of tweets originating from the affected areas also revealed the influence of distance on Twitter messages. Authorities should mostly use the negative tweets in the during-phase of disaster to detect people at risk and their immediate needs.

Our major aim was to understand the communication dynamics through different modalities including temporal, text, user, sentiment, and social network analyses and identify the most important information that can be derived from twitter during disasters. Government / relief / emergency agencies can utilize such information to reach and support at least a subset of the population most at risk. While government agencies are among the prominent Twitter users during disasters, they primarily rely on one-way communication rather than engaging with their audiences [74], a result confirmed by our network analysis. A major advantage of social media over traditional media sources is its potential for a bi-directional communication where public could provide the agencies with useful information for disaster management [68]. However, this was not the case during hurricane Sandy.

With the increasing frequency of extreme weather events, Twitter users' networks along with tweets' texts and geolocation information provide real-time data that can equip emergency management tools. Our findings illustrate Twitter's utility as an image-sharing platform that might be useful for developing applications to identify relevant images for disaster response using the location information embedded in the messages. However, the major challenge is the quality assessment of such data [75]. It is important to evaluate the relevancy and credibility of data and select only tweets related to the event. While current research is mostly focused on keyword filtering, developing a new approach integrating various data types is an important step for future research to utilize social media successfully in natural disasters. In our future research, we will develop an automated data-learned model to filter geocoded images shared across the Twitter platform by leveraging a multi-model that analyzes geospatial, image, user and text data.

Twitter is a useful medium through which individuals can communicate during weather-related emergencies; a valuable source for researchers to better understand these communications; and a useful way for agencies to communicate with individuals at risk. However, various barriers (e.g., lack of resources, lack of organizational support) to the use of social media analytic tools in organizations need to be first addressed to facilitate disaster management efforts [76]. With the difficulty of sorting and locating relevant information for public [68], the government agencies should facilitate the process by developing automated emergency management tools that extract and analyze information shared on social media. These tools can also benefit from social network analysis to help authorities in accelerating information diffusion during disasters. Future studies should examine other social media platforms in a multi-case approach to increase the effectiveness of social media in disaster management.

Funding

The work was supported by the Coastal Storm Awareness Program (New York Sea Grant), National Oceanic and Atmospheric Administration - NOAA (Award number: NA130AR4830229) and Mississippi Agriculture and Forestry Experimentation Station, Mississippi State University.

References

[1] W.-H. Nam, M.J. Hayes, M.D. Svoboda, T. Tadesse, D.A. Wilhite, Drought hazard assessment in the context of climate change for South Korea, Agric. Water Manag. 160 (2015) 106–117. doi:10.1016/j.agwat.2015.06.029.

[2] D. Feldman, S. Contreras, B. Karlin, V. Basolo, R. Matthew, B. Sanders, D. Houston, W. Cheung, K. Goodrich, A. Reyes, K. Serrano, J. Schubert, A. Luke, Communicating flood risk: Looking back and forward at traditional and social media outlets, Int. J. Disaster Risk Reduct. 15 (2016) 43–51. doi:10.1016/j.ijdrr.2015.12.004.

[3] Internet Live Stats, (n.d.). www. internetlivestats.com (accessed August 30, 2018).

[4] N. Mohammadi, Q. Wang, J.E. Taylor, Diffusion Dynamics of Energy Saving Practices in Large Heterogeneous Online Networks., PLoS One. 11 (2016) 1–23. doi:10.1371/journal.pone.0164476.

[5] J.B. Houston, J. Hawthorne, M.F. Perreault, E.H. Park, M. Goldstein Hode, M.R. Halliwell, S.E. Turner Mcgowen, R. Davis, S. Vaid, J.A. Mcelderry, S.A. Griffith, Social media and disasters: A functional framework for social media use in disaster planning, response, and research, Disasters. 39 (2015) 1–22. doi:10.1111/disa.12092.

[6] J. Kim, M. Hastak, Social network analysis: Characteristics of online social networks after a disaster, Int. J. Inf. Manage. 38 (2018) 86–96. doi:10.1016/j.ijinfomgt.2017.08.003.

[7] J. Qadir, A. Ali, R. ur Rasool, A. Zwitter, A. Sathiaseelan, J. Crowcroft, Crisis analytics: big data-driven crisis response, J. Int. Humanit. Action. 1 (2016) 12. doi:10.1186/s41018-016-0013-9.

[8] L. Li, Q. Zhang, J. Tian, H. Wang, Characterizing information propagation patterns in emergencies: A case study with Yiliang Earthquake, Int. J. Inf. Manage. 38 (2018) 34–41. doi:10.1016/j.ijinfomgt.2017.08.008.

[9] F. Yuan, R. Liu, Feasibility study of using crowdsourcing to identify critical affected areas for rapid damage assessment: Hurricane Matthew case study, Int. J. Disaster Risk Reduct. 28 (2018) 758–767. doi:10.1016/j.ijdrr.2018.02.003.

[10] J.H. Sorensen, B.V. Sorensen, A. Smith, Z. Williams, Results of An Investigation of the Effectiveness of Using Reverse Telephone Emergency Warning Systems in the October 2007 San Diego Wildfires, 2009. doi:ORNL/TM-2009/154.

[11] Y. Tyshchuk, C. Hui, M. Grabowski, W.A. Wallace, Social Media and Warning Response Impacts in Extreme Events: Results from a Naturally Occurring Experiment, in: 2012 45th Hawaii Int. Conf. Syst. Sci., IEEE, 2012: pp. 818–827. doi:10.1109/HICSS.2012.536.

[12] J.R. Ragini, P.M.R. Anand, V. Bhaskar, Big data analytics for disaster response and recovery through sentiment analysis, Int. J. Inf. Manage. 42 (2018) 13–24. doi:10.1016/j.ijinfomgt.2018.05.004.

[13] B.R. Lindsay, Social Media and Disasters: Current Uses, Future Options and Policy Considerations., 2011. doi:R41987.

[14] J. Kim, J. Bae, M. Hastak, Emergency information diffusion on online social media during storm Cindy in U.S., Int. J. Inf. Manage. 40 (2018) 153–165. doi:10.1016/j.ijinfomgt.2018.02.003.

[15] M. Eshghi, H.R. Schmidtke, An approach for safer navigation under severe hurricane damage, J. Reliab. Intell. Environ. 4 (2018) 161–185. doi:10.1007/s40860-018-0066-1.

[16] L. Grube, V.H. Storr, The capacity for self-governance and post-disaster resiliency, Rev. Austrian Econ. 27 (2014) 301–324. doi:10.1007/s11138-013-0210-3.

[17] A. Joshi, M. Aoki, The role of social capital and public policy in disaster recovery: A case study of Tamil Nadu State, India, Int. J. Disaster Risk Reduct. 7 (2014) 100–108. doi:10.1016/j.ijdrr.2013.09.004.

[18] M.W. Graham, E.J. Avery, S. Park, The role of social media in local government crisis communications, Public Relat. Rev. 41 (2015) 386–394. doi:10.1016/j.pubrev.2015.02.001.

[19] B.F. Liu, J.D. Fraustino, Y. Jin, How Disaster Information Form, Source, Type, and Prior Disaster Exposure Affect Public Outcomes: Jumping on the Social Media Bandwagon?, J. Appl. Commun. Res. 43 (2015) 44–65. doi:10.1080/00909882.2014.982685.

[20] L. Ann St Denis, L. Palen, K.M. Anderson, Mastering Social Media: An Analysis of Jefferson County's Communications during the 2013 Colorado Floods, in: Proc. 11th Int. ISCRAM Conf., University Park, Pennsylvania, 2014: pp. 737–746.

[21] D. Murthy, S.A. Longwell, TWITTER AND DISASTERS, Information, Commun. Soc. 16 (2013) 837–855. doi:10.1080/1369118X.2012.696123.

[22] S. Muralidharan, L. Rasmussen, D. Patterson, J.-H. Shin, Hope for Haiti: An analysis of Facebook and Twitter usage during the earthquake relief efforts, Public Relat. Rev. 37 (2011) 175–177. doi:10.1016/j.pubrev.2011.01.010.

[23] J.-Y. Jung, M. Moro, Multi-level functionality of social media in the aftermath of the Great East Japan Earthquake, Disasters. 38 (2014) s123–s143. doi:10.1111/disa.12071.

[24] A.L. Hughes, L.A.A. St. Denis, L. Palen, K.M. Anderson, Online public communications by police & fire services during the 2012 Hurricane Sandy, in: Proc. 32nd Annu. ACM Conf. Hum. Factors Comput. Syst. - CHI '14, ACM Press, New York, New York, USA, 2014: pp. 1505–1514. doi:10.1145/2556288.2557227.

[25] S. V. Ukkusuri, X. Zhan, A.M. Sadri, Q. Ye, Use of Social Media Data to Explore Crisis Informatics, Transp. Res. Rec. J. Transp. Res. Board. 2459 (2014) 110–118. doi:10.3141/2459-13.

[26] A. Acar, Y. Muraki, Twitter for crisis communication: lessons learned from Japan's tsunami disaster, Int. J. Web Based Communities. 7 (2011) 392. doi:10.1504/IJWBC.2011.041206.

[27] J. Sutton, L. Palen, I. Shklovski, Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires, in: Proc. 5 Th Int. ISCRAM Conf., 2008: pp. 624–632.

[28] A.T. Chatfield, C.G. Reddick, Understanding Risk Communication Gaps through E-Government Website and Twitter Hashtag Content Analyses: The Case of Indonesia's Mt. Sinabung Eruption, J. Homel. Secur. Emerg. Manag. 12 (2015). doi:10.1515/jhsem-2014-0086.
[29] Z. Tang, L. Zhang, F. Xu, H. Vo, Examining the role of social media in California's drought risk management in 2014, Nat. Hazards. 79 (2015) 171–193. doi:10.1007/s11069-015-1835-2.
[30] M. Yago, Z. Li, S.L. Cutter, Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew, PLoS One. 12 (2017) 1–22. doi:10.1371/journal.pone.0181701.
[31] J. Rexiline Ragini, P.M. Rubesh Anand, V. Bhaskar, Mining crisis information: A strategic approach for detection of people at risk through social media analysis, Int. J. Disaster Risk Reduct. 27 (2018) 556–566. doi:10.1016/j.ijdrr.2017.12.002.
[32] B.-K.H. Vo, N. Collier, Twitter Emotion Analysis in Earthquake Situations, Int. J. Comput. Linguist. Appl. 4 (2013) 159–173.
[33] M. Tsou, C. Jung, C. Allen, J. Yang, S.Y. Han, B.H. Spitzberg, J. Dozier, Building a Real-Time Geo-Targeted Event Observation (Geo) Viewer for Disaster Management and Situation Awareness, in: 2017: pp. 85–98. doi:10.1007/978-3-319-57336-6_7.
[34] S.J. Park, Y.S. Lim, H.W. Park, Comparing Twitter and You Tube networks in information diffusion: The case of the "Occupy Wall Street" movement, Technol. Forecast. Soc. Change. 95 (2015) 208–217. doi:10.1016/j.techfore.2015.02.003.
[35] H.T.P. Williams, J.R. McMurray, T. Kurz, F. Hugo Lambert, Network analysis reveals open forums and echo chambers in social media discussions of climate change, Glob. Environ. Chang. 32 (2015) 126–138. doi:10.1016/j.gloenvcha.2015.03.006.
[36] A. Sapountzi, K.E. Psannis, Social networking data analysis tools & challenges, Futur. Gener. Comput. Syst. 86 (2018) 893–913. doi:10.1016/j.future.2016.10.019.
[37] O. Kounadi, A. Ristea, M. Leitner, C. Langford, Population at risk: using areal interpolation and Twitter messages to create population models for burglaries and robberies, Cartogr. Geogr. Inf. Sci. 45 (2018) 205–220. doi:10.1080/15230406.2017.1304243.
[38] A. Karduni, I. Cho, G. Wessel, W. Ribarsky, E. Sauda, W. Dou, Urban space explorer: A visual analytics system for urban planning, IEEE Comput. Graph. Appl. 37 (2017) 50–60. doi:10.1109/MCG.2017.3621223.
[39] C. Allen, M.H. Tsou, A. Aslam, A. Nagel, J.M. Gawron, Applying GIS and machine learning methods to twitter data for multiscale surveillance of influenza, PLoS One. 11 (2016) 1–10. doi:10.1371/journal.pone.0157734.
[40] National Weather Service, Service Assessment: Hurricane/Post-Tropical Cyclone Sandy, October 22-29, 2012, Silver Spring, MD, 2013. https://www.weather.gov/media/publications/assessments/Sandy13.pdf.
[41] H.M. Saleem, Y. Xu, D. Ruths, Effects of Disaster Characteristics on Twitter Event Signature, Procedia Eng. 78 (2014) 165–172. doi:10.1016/j.proeng.2014.07.053.
[42] E. Guskin, P. Hitlin, Hurricane Sandy and Twitter, Pew Res. Cent. (2012). http://www.journalism.org/2012/11/06/hurricane-sandy-and-twitter/ (accessed June 8, 2018).
[43] M.C. Stewart, B. Gail Wilson, The dynamic role of social media during Hurricane #Sandy: An introduction of the STREMII model to weather the storm of the crisis lifecycle, Comput. Human Behav. 54 (2016) 639–646. doi:10.1016/j.chb.2015.07.009.
[44] T. Shelton, A. Poorthuis, M. Graham, M. Zook, Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of "big data," Geoforum. 52 (2014) 167–179. doi:10.1016/j.geoforum.2014.01.006.
[45] National Hurricane Center, NHC GIS Archive - Tropical Cyclone Best Track, (2012). https://www.nhc.noaa.gov/gis/archive_besttrack.php?year=2012 (accessed August 30, 2018).
[46] ArcGIS, FEMA MOTF Hurricane Sandy Impact Analysis, (2014). https://www.arcgis.com/home/item.html?id=307dd522499d4a44a33d7296a5da5ea0 (accessed August 30, 2018).
[47] Gnip APIs, (n.d.). http://support.gnip.com/apis/ (accessed April 1, 2015).
[48] D. Murthy, A.J. Gross, Social media processes in disasters: Implications of emergent technology use, Soc. Sci. Res. 63 (2017) 356–370. doi:10.1016/j.ssresearch.2016.09.015.
[49] A.T. Chatfield, H.J. Scholl, U. Brajawidagda, #Sandy Tweets: Citizens' Co-Production of Time-Critical Information during an Unfolding Catastrophe, in: 2014 47th Hawaii Int. Conf. Syst. Sci., IEEE, 2014: pp. 1947–1957. doi:10.1109/HICSS.2014.247.
[50] A. Rao, N. Spasojevic, Z. Li, T. Dsouza, Klout score: Measuring influence across multiple social networks, in: 2015 IEEE Int. Conf. Big Data (Big Data), IEEE, 2015: pp. 2282–2289.

doi:10.1109/BigData.2015.7364017.
[51] C. Edwards, P.R. Spence, C.J. Gentile, A. Edwards, A. Edwards, How much Klout do you have…A test of system generated cues on source credibility, Comput. Human Behav. 29 (2013) A12–A16. doi:10.1016/j.chb.2012.12.034.
[52] G. Salton, C. Buckley, Text Retrieval, Inf. Process. Manag. 24 (1988) 513–523.
[53] V.D. Blondel, J.L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech. Theory Exp. 2008 (2008) 1–12. doi:10.1088/1742-5468/2008/10/P10008.
[54] A. Croitoru, N. Wayant, A. Crooks, J. Radzikowski, A. Stefanidis, Linking cyber and physical spaces through community detection and clustering in social media feeds, Comput. Environ. Urban Syst. 53 (2015) 47–64. doi:10.1016/j.compenvurbsys.2014.11.002.
[55] S. Heymann, B. Le Grand, Visual Analysis of Complex Networks for Business Intelligence with Gephi, in: 2013 17th Int. Conf. Inf. Vis., IEEE, 2013: pp. 307–312. doi:10.1109/IV.2013.39.
[56] V.K. Neppalli, C. Caragea, A. Squicciarini, A. Tapia, S. Stehle, Sentiment analysis during Hurricane Sandy in emergency response, Int. J. Disaster Risk Reduct. 21 (2017) 213–222. doi:10.1016/j.ijdrr.2016.12.011.
[57] F. Karimi, S. Sultana, A. Shirzadi Babakan, S. Suthaharan, An enhanced support vector machine model for urban expansion prediction, Comput. Environ. Urban Syst. 75 (2019) 61–75. doi:10.1016/j.compenvurbsys.2019.01.001.
[58] D.L. Hansen, B. Shneiderman, M.A. Smith, Social Network Analysis: Measuring, Mapping, and Modeling Collections of Connections, in: Anal. Soc. Media Networks with NodeXL Insights from a Connect. World, Elsevier Science & Technology, 2010: pp. 31–49.
[59] A.T. Chatfield, U. Brajawidagda, Twitter tsunami early warning network: a social network analysis of Twitter information flows, in: 23rd Australas. Conf. Inf. Syst., 2012: pp. 1–10.
[60] M.O. Jackson, Social and Economic Networks, Princeton University Press, Princeton, 2008. doi:10.1017/CBO9781107415324.004.
[61] M. Bastian, S. Heymann, M. Jacomy, Gephi: An Open Source Software for Exploring and Manipulating Networks, in: Proc. Third Int. ICWSM Conf., 2009: pp. 361–362. http://qualitysafety.bmj.com/lookup/doi/10.1136/qshc.2004.010033.
[62] J.A. Burke, P.R. Spence, K.A. Lachlan, Crisis preparation, media use, and information seeking during Hurricane Ike: Lessons learned for emergency communication, J. Emerg. Manag. 8 (2010) 27–37. doi:10.5055/jem.2010.0030.
[63] H.M. Saleem, Y. Xu, D. Ruths, Novel Situational Information in Mass Emergencies: What does Twitter Provide?, Procedia Eng. 78 (2014) 155–164. doi:10.1016/j.proeng.2014.07.052.
[64] A. k. Abbas, O. Bayat, O.N. Ucan, Estimation of Twitter user's nationality based on friends and followers information, Comput. Electr. Eng. 66 (2018) 517–530. doi:10.1016/j.compeleceng.2017.06.033.
[65] A. Crooks, A. Croitoru, A. Stefanidis, J. Radzikowski, #Earthquake: Twitter as a Distributed Sensor System, Trans. GIS. 17 (2013) 124–147. doi:10.1111/j.1467-9671.2012.01359.x.
[66] C.. Caragea, A.. Squicciarini, S.. Stehle, K.. Neppalli, A.. Tapia, Mapping moods: Geo-mapped sentiment analysis during hurricane sandy, in: ISCRAM 2014 Conf. Proc. - 11th Int. Conf. Inf. Syst. Cris. Response Manag., 2014: pp. 642–651. http://www.iscram.org/legacy/ISCRAM2014/papers/p29.pdf.
[67] P.R. Spence, K.A. Lachlan, X. Lin, M. del Greco, Variability in Twitter Content Across the Stages of a Natural Disaster: Implications for Crisis Communication, Commun. Q. 63 (2015) 171–186. doi:10.1080/01463373.2015.1012219.
[68] K.A. Lachlan, P.R. Spence, X. Lin, M. Del Greco, Screaming into the Wind: Examining the Volume and Content of Tweets Associated with Hurricane Sandy, Commun. Stud. 65 (2014) 500–518. doi:10.1080/10510974.2014.956941.
[69] Y. Kryvasheyeu, H. Chen, E. Moro, P. Van Hentenryck, M. Cebrian, Performance of Social Network Sensors during Hurricane Sandy, PLoS One. 10 (2015) e0117288. doi:10.1371/journal.pone.0117288.
[70] R. Hanneman, M. Riddle, Introduction to Social Network Methods, Riverside, CA, 2005. http://faculty.ucr.edu/~hanneman/.
[71] M. Oliveira, J. Gama, An overview of social network analysis, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2 (2012) 99–115. doi:10.1002/widm.1048.
[72] P.M. Landwehr, W. Wei, M. Kowalchuck, K.M. Carley, Using tweets to support disaster planning, warning and response, Saf. Sci. 90 (2016) 33–47. doi:10.1016/j.ssci.2016.04.012.
[73] D. Westerman, P.R. Spence, B. Van Der Heide, A social network as information: The effect of system generated reports of connectedness on credibility on Twitter, Comput. Human Behav. 28 (2012) 199–206. doi:10.1016/j.chb.2011.09.001.

[74] R.D. Waters, J.M. Williams, Squawking, tweeting, cooing, and hooting: analyzing the communication patterns of government agencies on Twitter, J. Public Aff. 11 (2011) 353–363. doi:10.1002/pa.385.
[75] M. Eshghi, A.A. Alesheikh, Introducing a Conceptual Model to Improve the Quality of Storage of Volunteered Geographic Information: In the Field of “Fitness-for-Use” Indicator, J. Geomatics Sci. Technol. 6 (2016) 19–32. http://jgst.issge.ir/article-1-393-en.html.
[76] S. Anson, H. Watson, K. Wadhwa, K. Metz, Analysing social media data for disaster preparedness: Understanding the opportunities and barriers faced by humanitarian actors, Int. J. Disaster Risk Reduct. 21 (2017) 131–139. doi:10.1016/j.ijdrr.2016.11.014.