Evolution of Communities on Twitter and the Role of their Leaders during Emergencies

Yulia Tyshchuk*, Hao Li†, Heng Ji†and William A. Wallace*
*Department of Industrial and Systems Engineering,
Rensselaer Polytechnic Institute
Email: {tyshcy,wallaw}@rpi.edu; Phone: 518-867-1527
†Computer Science Department,
Rensselaer Polytechnic Institute
Email: {haoli.qc,hengjicuny}@gmail.com; Phone: 518-276-2103

***Abstract*—Twitter is presently utilized as a channel of communication and information dissemination. At present, government and non-government emergency management organizations utilize Twitter to disseminate emergency relevant information. However, these organizations have limited ability to evaluate the Twitter communication in order to discover communication patterns, key players, and messages that are being propagated through Twitter regarding the event. More importantly there is a general lack of knowledge of who are the individuals or organizations that disseminate warning information, provide confirmations of an event and associated actions, and urge others to take action. This paper presents results of the analysis of two events - 2011 Japan Tsunami and 2012 Hurricane Sandy. These results provide an insight into understanding human behavior, collectively as part of virtual communities on Twitter and individually as leaders and members of those communities. Specifically, their behavior is evaluated in terms of obtaining and propagating warning information, seeking and obtaining additional information and confirmations, and taking the prescribed action. The analysis will employ a methodology that shows how Natural Language Processing (NLP) and Social Network Analysis (SNA) can be integrated to provide these results.This methodology allows to extract actionable Twitter messages, construct actionable network, find actionable communities and their leaders, and deternine the behaviors of the community members and their leaders. Moreover, the methodology identifies specific roles of the community leaders. Such roles include dispensing unique/new emergency relevant information, providing confirmations to the members of the communities, and urging them to take the prescribed action. The results show that the government agencies had limited participation on Twitter during 2011 Japan Tsunami compared to an extensive participation during 2012 Hurricane Sandy. The behavior of Twitter users during both events was consistent with the issuance of actionable information (i.e. warnings). The findings suggest higher cohesion among the virtual community members during 2011 Japan Tsunami than during 2012 Hurricane Sandy event. However, during both events members displayed an agreement on required protective action (i.e. if some members were propagating messages to take action the other members were taking action). Additionally, higher differentiation of leadership roles was demonstrated during 2012 Hurricane Sandy with stronger presence of official sources in leadership roles.**

I. Introduction

Twitter is an important channel of information dissemination. It is particularly useful when current and relevant information is required. The format of Twitter messages permits people to exchange information about any occurrence. This capability is very useful during emergencies, events that pose a significant threat to one's well-being. In our work we focused on one type of emergencies - natural disasters. Twitter messages, interview data, and electronic alerts concerning the 2011 Japan tsunami and 2012 Hurricane Sandy provided the data for the research reported on in this paper. During emergencies such as a tsunami and hurricane, when the impact of the people and infrastructure is significant, people engage in information milling - obtaining and exchanging information and/or confirmation. The process requires rapid access to the most current information. Twitter has the capability to provide this functionality. Additionally, Twitter provides people with a way to connect with others affected by the same emergency, which can provide emotional support [1].

One of the significant challenges in studying Twitter is a sheer volume of data and lack of ability to efficiently read the data. In this paper Natural Language Processing (NLP) techniques were used to extract three types of actionable events from 2011 Japan tsunami and 2012 Hurricane Sandy datasets: receive the warning, seek information or confirmation, and take prescribed action. NLP techniques were used to associate tweets with following attributes - modality and polarity. These attributes provide further insights into the information being shared on Twitter. Additionally, first story analysis demonstrated the amount of unique/new emergency relevant information that was exchanged among the Twitter users. The analysis was also used to trace the information initiators.

The paper begins with an evaluation of existing methods. The paper then describes a novel methodology that was applied, which incorporated NLP with Social Network Analysis (SNA) techniques. The paper proceeds to describe the data set used in applying the methodology. The results are then described in detail in the following section. The paper concludes with a discussion of contributions and suggestions for future research.

II. Related Work

A. Warning Response Process during Emergencies

During emergencies affected individuals participate in the warning response process, which includes obtaining and sharing information, the evidence of which can be discovered on Twitter during an emergency. In general, the warning response process for an individual has been segmented into six stages [2]: (1) obtaining/hearing the warning; (2) understanding the contents of the warning; (3) trusting the warning; (4) personalizing the warning; (5) seeking information/confirmation; and (5) taking action.

An individual starts the warning response process by receiving notification of the emergency and ends the process by taking action, where doing nothing is a valid action. However, how and when each stage is accomplished may vary across individuals and emergencies [3]. The first stage of the warning response process for individuals is to obtain the warning from one or many sources. The second stage of the warning response process requires assigning a specific meaning to the warning message, which can vary from individual to individual. This meaning can also be different from what intended by the issuing source. The third stage is trusting the warning message, which is influenced by many factors such as the source of the message, contents, and the channel. The fourth stage requires personalization of the warning to one's situation. This requires an individual to assess her or his willingness to assume the necessary personal risk. The fifth stage of warning response process is to seek additional information or attempt to obtain confirmations about the information already obtained [2]. This process is often referred to as warning confirmation process. The final stage of warning response process is taking action. People engage in the action they believe is the best for them, which may be at odds with a prescribed action. Three stages of the warning response process - obtaining/hearing the warning, seeking information/confirmation, and taking action can be inferred from communication between individuals unlike the other three stages, which are cognitive processes.

B. Social Media during Emergencies

Social media has been used by the public as well as governmental and non-governmental organizations during emergencies. Some examples of the use include rapid information dissemination of one's well-being as it was demonstrated by the researchers in [4]. In Haiti, U.S. government was able to utilize social media, such as Wikipedias and workspace sharing media, as a knowledge based system [5]. The researchers in [6] were able to develop a unique annotation, which facilitated the emergence of the digital volunteers. Social media provides a natural environment for facilitating decentralized coordination for onsite field response teams [7]. During 2011 Japan Tsunami, people utilized Twitter for information milling, warning propagation, providing information about recovery efforts, and emotional support [1].

C. Social Network Analysis (SNA) and Twitter

Social network analysis facilitates the determination of the communication patterns among users. In [1], the researchers showed that social network analysis is a useful tool in identifying information sources. It was demonstrated that there are various techniques rooted in social network analysis to study emergent communities on Twitter [1]. The Twitter communication networks were analyzed to find the structural phenomena related to directed closure and its role in link formation [8]. In [9], researchers studied the Twitter hashtag adoption based on the structural properties of the network. The research showed that Twitter communication networks that drive the daily interactions among people are sparse and are based on existing friends and followers [10].

D. Open-domain Event Discovery

Traditional event extraction work focused on supervised learning for pre-defined event types in formal genres such as newswire [11], [12], [13]. However, these methods are not appropriate for social media, which covers a wide range of diverse topics and lacks labeled data. Early work of event discovery exploited the word distribution differences across instances. For example, Yang et al [14] detected events by clustering documents based on the semantic distance between documents, while Kleinberg et al. [15] used word distributions to discover events by grouping words together. Some recent work attempted to rapidly and automatically adapt an event extraction system to new event types. For example, Li et al. [16] automatically acquired verb clusters from parallel corpora and discovered novel events based on named entity recognition, semantic role labeling and active learning.

Unlike formal genres, social media stream is characterized by short messages with heavily colloquial speech. To handle such data stream, Weng and Lee [17] tackled event discovery task for Twitter by detecting important word tokens and clustering them to represent novel events. They analyzed word-specific signals in the time domain. The advantage was that signals for individual words were built by applying wavelet analysis on the frequency-based raw signals of the words, hence important words were identified based on corresponding signal auto-correlations. The researchers in [18] developed a graphical model to extract event records from Twitter by learning a latent set of records and a record-message alignment simultaneously. However, their method requires a seed set of example records as the source of supervision so it is not appropriate for our use. The researchers in [19] trained a supervised model to extract event tuples from tweets. However their approach is highly restricted to their annotated event types and was not able to capture events in our domain (e.g., evacuation event).

To conclude, our event extraction approach is most related to the research explored in [16] and [17]. Given some event clusters as seeds, we obtained new relevant keywords to expand each event keyword cluster and use these clusters to represent events. In addition, we utilized semantic attributes to declaratively discriminate specific and affirmed events from others. To the best of our knowledge, this is the first work to incorporate semantic attributes into novel event discovery in an open domain.

E. First Story Detection

The traditional approach for first story detection uses a term vector to represent each document (e.g., an news article) [20], [21]. Each new document is then compared with the previous ones, and if its similarity with the closest document is below a threshold, it is declared to be a new story. However, this approach is not feasible for large data sets (e.g., tweets) because of its high computational cost. A computationally better approach for first story detection task utilizes locality-sensitive hashing (LSH) with a variance reduction strategy [22]. This method can achieve similar performance while gain more than an order of magnitude speedup compared with the system previously described in [20]. Experiments using this method were conducted on large streaming Twitter data sets and achieved reasonable results. In this paper, the above-described approach is used for first story detection in tweets. Given a large amount of tweets sorted in the timeline, we apply LSH to group similar tweets together and identify all the tweets that discuss a new bit of information. In addition, we also link later tweets to the previous ones if they are talking about the similar bit of information in order to generate information clusters.

III. Methodology

A. Overview

An overview of the approach taken in this paper is illustrated in Fig. 1. First, data was collected via streaming Twitter API during the time of an emergency. Then the data was processed using the Support Vector Machines (SVMs) based on topic/off topic binary classifier to extract tweets related to the emergency. Note that the on/off topic classification was conducted on 2010 Japan tsunami event only. The 2012 Hurricane Sandy data set was collected using hashtags "#Sandy" and "#Hurricane", therefore, all tweets were on-topic. Next, a selected set of search terms was used to annotate the tweets with actionable events - 'propagate the warning', 'seek information or confirmation', and 'take prescribed action'. To overcome the unstructured format of the tweets' text an appropriate set of NLP techniques was used. The annotation was further enriched through assignment of attributes for each tweet - polarity and modality. This was accomplished via SVMs based event attribute classification. Subsequently, the first story analysis was conducted using Locality Sensitive Hashing algorithm to detect the information clusters as well as the tweets that first introduced the information on Twitter.

The timelines were either constructed utilizing data collected from on-site interviews and publicly available information on the Internet or based on the 24 hour time slices. The timelines were used to construct communication networks for each time slice. A random walk algorithm was employed to discover communities in Twitter communication networks by time slice. SNA was used to identify the leaders of these communities. The knowledge obtained from NLP about the tweet content - actions, attributes, first story identification, and story ranking, enabled us to make inferences about the behaviors of community members and roles of their leaders.

B. NLP Approach

1) Terminology: We defined the following terminology for a series of NLP approaches.

On-topic/Off-topic Tweets: We defined the tweets that were related to the topic of our interest as "on-topic" and the rest as "off-topic". In our case study, all tweets related to Japan tsunami and Hurricane Sandy were on-topic. An on-topic tweet example is as follows: *RT @CBCAlerts: 7.2 magnitude earthquake hits Northern Japan. Tsunami alert has been issued.#Japan #Quake* while an off-topic tweet example is as follows: *I have an early wake up, but 2 hour long Skype sessions w/ distant friends are worth the minimal hours of sleep. #buddies #friendsarefamily*.

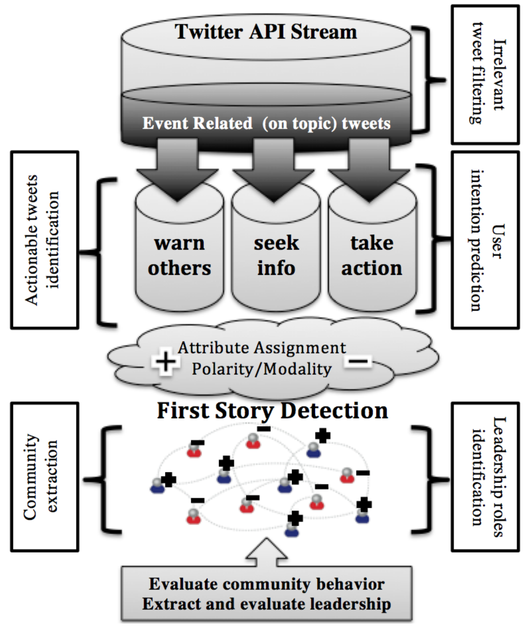

Fig. 1. Overview of the Methodology Picture

Actionable Events: Events that belong to the following categories: receive the warning; seek information or confirmation; and take prescribed action. The categories were selected from the six stages of warning response process previously described in Section II.
Event Attributes: Event attributes were used to measure user intention to participate in an actionable event. Two semantic attributes were adapted from Automatic Content Extraction 2005 Evaluation (ACE2005) [23] to describe each actionable event: (1) modality, where an event was "asserted" when the author or speaker made reference to it as though it were a real occurrence; and (2) polarity, where an event was "positive" when it was explicitly indicated that the event occurred.
Actionable Tweets: Tweets that belong to an actionable event (receive the warning, seek confirmation, and take prescribed action).
First Story Tweets: Tweets that mention for the first time a seminal event and a seminal event is a particular event that occurs at a specific time and space, e.g., an tsunami occured in Sendai, Japan on March 11th, 2011.

2) On-topic Tweet Detection: According to the hashtag definition from Twitter, the hashtag symbol, #, together with a relevant keyword or a phrase in the tweet is used to categorize the tweets and allow them to be displayed more easily in Twitter Search. Also, popular hashtagged words are often characterized as trending topics.

Inspired by the hashtag definition, we developed a novel annotation scheme based on the assumption that tweets with the same hashtag are on the same topic. First, we extracted hashtags with high frequency [1] that indicate trending topics. Then we manually annotated each trending hashtag as either on-topic or off-topic hashtag. After annotating hashtags, we propagated the on-topic/off-topic label of each hashtag to all tweets with each hashtag. We trained an on-topic/off-

[1] we treat hashtags appear more than 50 times as high frequency ones

topic tweet classifier, based on Support Vector Machines (SVMs) [24], using the following features: (1) unigrams (all unique unigrams of a tweet); (2) userID (the ID of the user who posted the tweet); (3) replyID (the ID of the user to whom the tweet is replying); and (4) mentionID (the ID of users wmentioned in the tweet d). All hashtags were removed from tweets during training and testing process, so the trained classifier was able to process all of the genetic tweets without any hashtags.

3) Actionable Event Extraction: After filtering out off topic tweets, we developed a bootstrapping framework to predict actionable events. To expand the key word seeds, we followed the cross-lingual event trigger clustering approach described in [16] to discover words with similar meanings. The algorithm exploited the idea that if two words - w_1 and w_2 on the source side of bi-lingual parallel corpora were aligned to the same word on the target side with high confidence, they should have similar meanings. For each English key word seed, the search was to find other English words that shared the same frequently aligned Chinese terms and vice versa. The word alignment information between each bi-lingual sentence pair was obtained by running Giza++ [25]. To eliminate the noise introduced by automatic alignment, we filtered out stop words and those English-Chinese word alignment pairs with frequency (in parallel corpora) less than a threshold[2]. Finally, we used each expanded keyword set as keywords to retrieve actionable events.

4) Event Attribute Labeling: In addition to identifying actionable events, we also labeled semantic attributes including modality and polarity for each event. We learned a separate SVMs based classifier for each attribute from ACE2005 training data[3]. The learnt classifier was applied to predict modality and polarity values for each actionable event. Because the training data set of ACE2005 includes news articles and our target domain is tweets, we explored the following genre-independent features to bridge the genre gap between news and tweets: (1) lexical features including unique words, lowercase words, lemmatized words and part-of-speech tags; (2) N-gram features, where an n-gram n_g (n=1, 2, 3) was selected as an indicative context feature if it matched one of the following two conditions - (i) n_g appeared only in one class, and with frequency higher than a threshold; and (ii) the probability that n_g occurring in one class was higher than a threshold; where both thresholds were optimized from a small development set including 30 events; and (3) dictionary features, such as expression, consideration, subjective, intention, condition, and negation, were used.

5) First Story Detection (FSD) and Event Clustering: The Locality Sensitive Hashing (LSH) method was used to remove the curse of dimensionality and applied to the FSD problem [22]. LSH was first proposed by Indyk and Motwani [26]. The underlying foundation was that if two documents are close together, then after a "projection" operation these two documents would remain close together. In other words, similar documents have a higher probability to be mapped into the same bucket thus the collision probability will be higher for documents that are close to each other. Given a LSH setting of k bits and L hashtables, two documents x and y are collide if and only if:

$$h_{ij}(x) = h_{ij}(y), i \in [1...L], j \in [1...k] \tag{1}$$

and the hash function $h_{ij}(x)$ is defined as:

$$h_{ij}(x) = sgn(u_{ij}^T x) \tag{2}$$

where u_{ij} are randomly generated vectors with components selected randomly from a Gaussian Distribution, e.g., $N(0, 1)$.

Algorithm 1 shows the pseudocode of LSH approach for First Story Detection and event clustering. All the tweets are sorted in chronological order. Novelty score is then assigned to document d by $Score(d)$, given a threshold $t \in [0, 1]$[4], if $Score(d) \geq t$ then d is a first story, otherwise cluster d with its most similar document that chronologically appears before it. To calculate distance between two documents we adapt the standard Cosine Similarity between two vectors:

$$\begin{aligned} distance(d, d^{'}) = cos(\theta) &= \frac{A \cdot B}{||A|| ||B||} \\ &= \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \sqrt{\sum_{i=1}^{n} (B_i)^2}} \end{aligned} \tag{3}$$

The advantage of LSH is that it only needs to find the nearest neighbor from the set of documents that were mapped to the same bucket instead of all the previous tweets. Compared with the brute force search, the computation cost of score function dropped from $O(|D_t|)$ ($|D_t|$ is the number of tweets have the time stamp before the current tweet's) to $O(1)$.

Algorithm 1: LSH-based FSD

```
foreach document d in corpus do
    add d to LSH;
    S ← set of points that collide with d in LSH;
    dis_min d ← 1;
    foreach d' in S do
        c = distance(d, d');
        if c < dis_min(d) then
            dis_min(d) ← c;
        end
    end
    score(d) = 1 − dis_min(d);
end
```

C. SNA Methodology

1) Network Construction: The communication network of Twitter data was constructed using the communication directional identifiers - @ for directed and mention tweets and RT for the re-tweets. Two relationships were incorporated into the communication network – the directed/mention and the re-tweet relationships. For directed/mention relationship an edge existed if one user tweeted and/or mentioned another user. The user doing the tweeting was at the head of the edge and the user who was mentioned or the tweet was directed to was at the tail of the relationship. For re-tweet relationship the edge existed if a user re-tweeted another user's tweet. The

[2] we set the frequency threshold as 4

[3] http://www.itl.nist.gov/iad/mig/tests/ace/2005/

[4] we set t as 0.2 in our experiments

user who was doing the re-tweeting was at the tail of the edge and user sending the original message was at the head of the relationship. The network was constructed for each of the time slices of the event timeline previously discussed. This allowed for investigation of the evolution and the dynamics of the network. The research evaluated actionable behaviors on Twitter, therefore, only actionable tweets were utilized to construct the network. The constructed network is referred to as Twitter communication network in the following sections.

2) Attribute Setup: The NLP analysis assigned specific attributes to each actionable tweet - modality and polarity. These attributes as well as a type of action (i.e. 'receive the warning', 'seek and/or obtain the confirmation', and 'take the prescribed action') were initially assigned as edge attributes in the Twitter communication network. However, when the Twitter network was constructed multiple and self-loop edges were discovered. Multiple edges represent multiple tweets between two users. The self-loop edges represent edges from the user to itself. The presence of such edges precluded the use of community finding algorithms. In order to address this problem the graph was simplified and edge attributes were automatically collapsed into the node attributes to preserve all of the extracted information. Each node's attribute was the sum of all respective tweet attributes sent or received by the user. These attributes helped define individuals behaviors. For example, if the user (i.e., node) has the following attributes - 'take the prescribed action' with positive modality and polarity, the person is taking the prescribed action. On the other hand, if the user (i.e., node) has the following attributes - 'take the prescribed action' and negative modality and polarity, someone else other than the person tweeting is not taking the prescribed action. The NLP attribute assignment defines individual behaviors as well as collective behaviors of Twitter users who are part of the same community.

3) Community Finding: Currently, most of the algorithms can not handle the directedness of the edges when detecting the communities [27]. In order to overcome this issue, the graphs are often converted into undirected graph for the purposes of community detection [28]. When Twitter users communicate among each other and direct their messages to other users the evidence of communication (tweets) is displayed in the profiles of both users. This dichotomy allowed us to justify the modification of the graph from directed to undirected graph for community detection purposes. The community finding approach utilized in the research was a random walk community detection algorithm. The foundation of the approach lies with the assumption that there are only a few edges that leave communities. Therefore, the algorithm uses a number of random walks on the network and then uses those walks to merge the separate communities in a bottom up manner [29]. This particular algorithm is most appropriate to find communities in the large sparse networks, which commonly occur in the Twitter data.

The social science literature informs the research on the properties of cohesive groups. It suggests that the people in the same community tend to have similar and redundant information. Moreover, there is an ease of information transfer in cohesive groups [30], [31]. In this research, this concept was evaluated in the context of Twitter communication network during emergencies. In order to ascertain if this theory of group behavior applies to the communications and behaviors on Twitter the correlation between the community members based on behaviors derived from the Twitter users' behavioral attributes was evaluated. The size of the communities found in the data enabled us to determine how many people obtained similar information and shared similar intents. The ten largest communities for each time slice were evaluated by examining the similarity (correlation) of behavior among the community members to discover the prevalent behavior.

4) Centrality and Prestige: Once the communities were identified the task was to find the community leaders. Each community was taken separately and a community leaders were identified as the most central/prestigious actors. The centrality/prestige measures that were utilized in this research were outDegree, inDegree, betweenness, and eigenvalue centrality (power). An outDegree centrality measure is simply the number of messages sent by a Twitter user to other users in the network. An outDegree measure is associated with faster information diffusion as it reaches more people. In [1], the researchers showed that people with high outDegree engage in information propagation. An inDegree measure represents a number of incoming messages sent to a Twitter user by other users. Another measure of betweenness represents a level of control one user has over the communication between other users. The users with high betweenness values serve as information gatekeepers [1], the betweenness of a node is the number of the shortest paths between any two nodes in the network that have to pass through this node [32]. A power measure represents the node's connectedness to other central nodes [33].

Each centrality measure is associated with a different kind of behavior, users scored high on each of those measures can represent different types of leadership. Therefore, three types of leaders are defined - the diffuser, the gatekeeper, and the information broker. The diffuser leader is a leader which "diffuses" the information through the network. This type of leader is associated with an outDegree measure as it measures the number of tweets (edges) a node sends out. Another type of leader is a gatekeeper. A gatekeeper is a node that controls an information flow in the network. Measures associated with the role of a gatekeeper are betweenness [34], [35] and power [36]. There are two types of gatekeepers that emerge when betweenness and power measures are combined - critical gatekeeper and unique access gatekeeper [36]. A critical gatekeeper is associated with high betweenness and low power values whereas a unique access gatekeeper is tied to low betweenness and high power values [36]. We defined the final type of the leader as information broker, who has access to valuable information and brokers it to other nodes in the network upon request. An information broker is associated with high inDegree and high power measures. A high power measure suggests access to other central actors and information they able to provide. A high inDegree measure suggests high frequency of inquiry from other users in the community. The frequency of inquiry for information can be inferred from the 'action' attribute - 'seek and obtain confirmation'.

Once the community leaders were identified their behavior was evaluated based on the type of actionable tweets they sent out. That behavior was then compared to the overall behavior of the community members. For example, when a leader of the community sent out a warning to evacuate, which was accompanied by action attribute - 'propagate the warning' and polarity - 'true', the expected result was for the community to follow the lead and send out the tweets with action attributes - 'propagate the warning' and/or 'take a prescribed action' and polarity - 'true'.

IV. Data Description

The methodology presented in this work is generalizable to all emergencies. In order to facilitate the understanding of the methodology and its generalizability two events were chosen: (1) the 2011 Japan Tsunami and (2) 2012 Hurricane Sandy. Two events were different in its impact as well as the duration of their impact. The tsunami occurred on March 11th, 2011 and impacted the entire Pacific Coastline. There were over 15,000 people whose lives were lost due to the tsunami including one in Klamath River, CA, USA. It also produced between $12 and $16 millions of dollars worth of damage in California [37]. In Hawaii, the governor had made a disaster declaration [38]. Throughout the event the tsunami has triggered multiple warnings issued by the Tsunami Warning Centers and evacuation orders issued by the local emergency management organizations. The event spanned over the 24 hours. The 2012 Hurricane Sandy had formed on October 22nd, 2012 and dissipated on October 31st, 2012. The event had affected 24 states along the eastern seaboard and had prompted disaster declarations in eleven states along the U.S. East Coast and New England. Hurricane Sandy had caused a significant impact with at least 286 people dead and $65 billion dollars worth of damage in U.S. alone [39].

Two types of data were collected for both events - qualitative and quantitative data. For 2011 Japan Tsunami the qualitative data was collected via semi-structured interviews with the members of emergency community who were involved during the event - members of Tsunami Warning Centers, emergency managers at Hawaii Civil Defense and Del Norte County Emergency Management Services, and members of local broadcast media. The "After Action Reports" were collected during the interviews, which allowed the construction of the detailed timeline of the event summarized in Table I. Additional information, which was obtained from searching publicly available information, further enriched the knowledge about the event and details about human behavior during the event. For 2012 Hurricane Sandy the qualitative data was obtained via semi-structured interviews with New York State Department of Homeland Security and Emergency Services Public Information Officers. Additional data was made available via public resources provided by state governments and Federal Emergency Management Agency. The summarized version of the timeline for 2012 Hurricane Sandy is described in Table II [40].

The qualitative data for both events included Twitter data. For 2011 Japan Tsunami the data was obtained from Information Sciences Institute through collaborative work and for 2012 Hurricane Sandy the data was collected in-house. Twitter data was collected via streaming Twitter API. The data included all of the tweets sent or received during the time of the events. In addition to the tweet messages, it also included user names, time stamps, and directed communication identifiers such as for directed messages and RT for re-tweets. The data was stored locally and can be accessed upon request.

TABLE I. 2011 Japan Tsunami Timeline

Time Slice	Time(UTC)	Events
1	5:46:28AM – 5:55:02AM	PTWC registers an earthquake 231 mi. from Tokyo, Japan of magnitude 7.9 and issues first bulletin – tsunami watch for HI
2	5:55:02AM – 6:41:22AM	PTWC issues second bulletins (international & HI); EOC's activated in HI
3	6:41:22AM – 7:31:00AM	PTWC issues third bulletin; tsunami warning is issued in HI
4	7:31:00AM – 9:01:00AM	Evacuation is ordered in HI, boat evacuations in HI and AK
5	9:01:00AM – 12:30:00AM	Evacuation travel is completed in HI, U of HI is closed; CA issues evacuation orders; tsunami arrives in King Cove, AK
6	12:30:00AM – 13:36:00AM	Tsunami arrives in HI: Hanalei, Kahului, Hilo
7	13:46:00AM – 17:31:00AM	Tsunami warning is downgraded to advisory in HI; all ports and evacuation zone are closed in HI; tsunami arrives in Crescent City, CA
8	17:31:00AM – 21:26:00AM	All clear is issued in HI
9	21:26:00AM – 6:36:00AM	Final all clear is issued by PTWC

TABLE II. 2012 Hurricane Sandy Timeline

Date	Events
October 22	Tropical Storm Sandy had officially formed
October 23	Possible Tropical Storm Watch for Florida Keys
October 24	Tropical Storm Watch for east coast of Florida
October 25	TFederal Emergency Management Agency (FEMA) elevates the enhanced watch for Washington D.C. FEMA deploys Incident Management Assistance Teams to CT, DE, NY, NJ, MA, NH, PA, and VT. Tropical Storm Watch was issued for NC and SC. The state and federal response coordination efforts continued.
October 26	NY, MD, D.C., PA, NC declare a state of emergency
October 27	FEMA activated the National Response Coordination Center. Non-government coordination (i.e. Red Cross) had begun its coordination.
October 28	Emergency declarations signed for CT, D.C., MD, MA, NJ, and NY. The USGS issued landslide alerts for several areas. New York City had made public transportation closings in preparation to the event.
October 29	Pre-disaster declarations signed for DE, RI, and PA. Hurricane Sandy downgraded to post-tropical storm and made landfall in southern NJ.
October 30	Major disasters declared for CT, NJ, and NY. Coordinated search, rescue, and recover efforts began.
October 31	Continued coordinated search, rescue, and recover efforts.

V. Results

A. Natural Language Processing

For 2010 Japan tsunami data set, we were able to annotate 800 hashtags in a very short time period (1.5 hours) and gathered a large number of human annotated tweets (311,735). As a result, 37 hashtags were annotated as on-topic and the rest were annotated as off-topic and thus 26,554 on-topic tweets and 285,181 off-topic tweets were gathered respectively. To balance the training and testing data, we randomly sampled the same amount of off-topic tweets as on-topic tweets to conduct the experiments. 42,486 tweets were randomly selected for training, and the remaining 10,622 tweets were used for blind test. The accuracy for on- topic classification for 2010 Japan tsunami was 81.93%. The accuracy results for both datasets, 2010 Japan tsunami and 2012 Hurricane Sandy, for polarity and modality were 96.8% and 78.4% respectively.

The actionable tweets were aggregated per time period to evaluate the results and compare analyzed data and Twitter user behavior with the timeline of the events. Table III represents the results for 2010 Japan Tsunami. There is a spike in the

TABLE III. 2010 JAPAN TSUNAMI ATTRIBUTES PER TIME SLICE

Time Slice	Warn	Confirm	Action	- / + Asserted	- / + Polarity
2	58	none	none	24 / 34	none / 58
3	328	2	4	202 / 132	4 / 330
4	6984	588	484	4592 / 3464	481 / 7575
5	2043	360	224	1566 / 1061	231 / 2396
6	1021	312	204	828 / 709	182 / 1355
7	1589	519	274	1299 / 1083	230 / 2152
8	1093	529	122	849 / 895	163 / 1581
9	2026	1498	216	1743 / 1997	470 / 3270

TABLE IV. 2012 HURRICANE SANDY ATTRIBUTES PER TIME SLICE

Day	Warn	Confirm	Action	- / + Asserted	- / + Polarity
Oct 25	2009	1792	283	2220 / 1864	447 / 3637
Oct 26	12731	8856	2686	15181 / 9092	2925 / 21348
Oct 27	16167	10761	5759	20743 / 11944	4689 / 27998
Oct 28	47873	50390	37215	83989 / 51489	18527 / 116951
Oct 29	80721	71092	35992	105720 / 82085	21504 / 166301
Oct 30	70027	60482	25952	89872 / 66589	20412 / 136049
Oct 31	26360	30002	9935	41343 / 24954	8191 / 58106

TABLE V. 2010 JAPAN TSUNAMI COMMUNITIES RESULTS

TS	# of com.	Range	# of com. (>4)	Percentage of $>$ 4 com.
2	10	{2:11}	1	10%
3	62	{2:41}	10	16%
4	1324	{2:248}	126	10%
5	705	{2:110}	39	6%
6	538	{2:51}	19	4%
7	729	{2:51}	33	5%
8	525	{2:51}	25	5%
9	878	{2:61}	33	4%

TABLE VI. 2012 HURRICANE SANDY COMMUNITIES RESULTS

Day	# of com.	Range	# of com. (>4)	Percentage of $>$ 4 com.
1	842	{1:167}	79	9%
2	62	{1:912}	419	11%
3	1324	{1:1481}	546	11%
4	705	{1:11289}	2412	14%
5	538	{1:7428}	3293	12%
6	729	{1:6040}	2531	11%
7	525	{1:2440}	1208	11%

volume of tweets during the time slice 4. This is natural as that's when most of the tsunami warnings were issued and evacuations were ordered along the affected coastline. Moreover, it is evident that the 'receive the warning' tweets are prevalent in earlier time slices and then gradually drops off as the event concludes. This is a natural progression and corresponds to the event timeline. The 'take prescribed action' tweets peak in time slices five, six, and seven after the evacuation orders have been issued. Finally, the confirmation tweets increase in the later time slices after the warnings and evacuation orders were issued. Additionally, during the later time slices people were confirming the well-being of their friends and relatives affected by the event. Similar results can be seen in 2012 Hurricane Sandy in Table IV. The volume of 'receive the warning' tweets rises leading up to and peaks on the day the landfall in southern New Jersey (October 28). The volume of 'seek and obtain confirmation' and 'take the prescribed action' tweets rise leading up to and peaking on the day prior to the landfall. The warnings issued by the government emergency organizations for the northeastern states required impacted population to take action on October 29th. The peaks occurring on Twitter on October 29th for 'seek and obtain confirmation' and 'take the prescribed action' show that users on Twitter followed the patterns of the evolution of the event. The analysis shows that the evolution of behaviors extracted from the NLP action assignments to the tweets correspond to a warning response process cycle and the overall evolution of both events.

B. Twitter Network Communities

First the community results for the 2011 Japan Tsunami are evaluated. Table V shows the results produced by the random walk algorithm. Note that the time slice (TS) one was omitted from the results there were no communities discovered during that time slice. The range in the table represents the size range of the communities - i.e. for time slice 2 the size of the smallest community was 2 and the size of the largest community was 11. A higher percentage of communities of size larger than four ('Percentage of $>$ 4 com.') occur during time slices two, three, and four. This result is expected as the users are exchanging warning information recently issued and confirming prescribed action.

When the communities and its members were examined more closely there was significant correlation found in the behaviors of community members. Over all time slices, every community had 80 percent or greater of its members that had exactly the same behavior - i.e., the same actionable event, modality, and polarity. For those communities, where there was a difference among the members' behaviors, the difference was in actionable events, and not in modality or polarity. The members usually split into two groups within the community, based on the actionable event - warning group, those who received and propagated the warning, and take action group, those who expressed intent to take the prescribed action. The finding suggests that people of a community tend to exhibit similar behaviors. It is important for all members of the community to share similar polarity for their behavior. For example, if the leader sends out a message urging people to evacuate - action 'propagate the warning' and polarity - 'positive', the expected result for the rest of the community is to respond with either action of 'propagate the warning' or 'take prescribed action' with the same polarity. When the polarity was evaluated among the members of the communities only 5 per cent or less of all communities exhibited difference in polarity among its members. Additionally, the tweets with confirmation actionable event rarely occurred in the large communities and were more typical of communities of size less than 4. Moreover, when the communities were traced from time slice to time slice there was little overlap discovered between its members. This suggests that the communities formed on Twitter serve a purpose in each time slice such as propagate the warning, obtain information or confirmation, and exhibit an intent to take the prescribed action. Once the action is completed there is no longer a need to participate on Twitter.

The 2012 Hurricane Sandy event spanned over nine days from its formation on October 22nd to its completion on October 31st . This timespan allows for higher participation in information exchange on Twitter. Table VI shows the results produced by the random walk algorithm for each of the seven days of collected data (October 25 - October 31).

Unlike 2011 Japan Tsunami the anticipated impact of 2012 Hurricane Sandy varied and spanned over the entire east coast of the United States. The vast area of impact required

TABLE VII. TIME SLICE FOUR COMMUNITY RESULTS

com	com size	Action	Modality	Polarity
2	32	receive	non-asserted	positive
4	114	receive	asserted	positive
5	3	receive	non-asserted	positive
6	77	receive	non-asserted	positive
10	325	receive	non-asserted	positive
11	42	receive	non-asserted	positive
15	128	receive	non-asserted	positive
16	94	receive	asserted	positive
17	47	receive	asserted	positive
25	80	receive	asserted	positive

TABLE VIII. TIME SLICE FOUR COMMUNITY LEADERSHIP RESULTS

Leaders	Action	Modality	Polarity	Tense
abc7	receive	non-asserted	positive	past
BreakingNews	receive	non-asserted	positive	present
fema	receive	non-asserted	positive	present
infoBMKG	receive	asserted	positive	past
BBCBreaking	receive	non-asserted	positive	past
CNN	receive	asserted	positive	past
BBCWorld	receive	non-asserted	positive	past
DamnItsTrue	receive	asserted	positive	present
thejakartaglobe	receive	non-asserted	positive	past
cnbrk	receive	non-asserted	positive	past

different actions to be taken by the impacted population. For example, the levels of different actions ranged for areas as small as individual cities, such as New York City. New York City was divided into three possible zones of impact but the evacuation order was issued only for the Zone A. The diversity of the prescribed action resulted in diversity of the behaviors among the members of the same Twitter communities as some members were required to evacuate and others weren't. Members of the same community were receiving and propagating warnings as well as confirming if action for their local area was required. The members of the same community were different in their action attributes, however, the polarity for each action was the same among the members of the same community. This finding is consistent with 2010 Japan tsunami finding on polarity. The first story results suggested that each community exchanged on average 14% of unique new information. The larger communities possessed the least amount of unique new information. The information in such communities was issued by selected individual members and then diffused to the rest of the members of the communities. A new finding, in contrast to 2011 Japan tsunami findings, suggests that tweets with confirmation actionable event were no longer specific to the communities of size less than four. This finding can be explained by the fact that there was more time allotted for people to seek confirmations prior to the hurricane impact.

C. Community Leaders

First the leaders of the communities discovered in 2011 Japan tsunami were evaluated. Specifically, only the communities of size larger than four were examined. It was discovered that the roles of diffuser and gatekeeper were assumed by the same nodes. Additionally, it was confirmed that the action of 'seek information or confirmation' is a characteristic of communities of size smaller than four, therefore, the information broker role was taken by a selected set of users in those communities. As shown in Table VII and VIII, ten largest communities for time slice four, when the critical warning information was issued, were selected for analysis, and diffuser

TABLE IX. LEADERSHIP RESULTS: DAY PRIOR TO THE 2012 HURRICANE SANDY LANDFALL IN SOUTHERN NEW JERSEY

Comm ID	Diffuser/Gatekeeper	Broker
1	MikeBloomberg	twchurricane
2	NHCAtlantic	13News
9	HuriicaneSandy	HurriicaneSandy
16	ASPCA	AMoDELSLIFE
33	rickygervais	rickygervais
36	JamesYammouni	yumyumyumniall
37	googlemaps	googlemaps
38	BBCBreaking	BBCBreaking
39	KagroX	KagroX
40	jimmyfallon	jimmyfallon

TABLE X. LEADERSHIP RESULTS: DAY OF THE 2012 HURRICANE SANDY LANDFALL IN SOUTHERN NEW JERSEY

Comm ID	Diffuser/Gatekeeper	Broker
2	NHCAtlantic	WSJweather
19	fema	BarackObama
29	DMVFollowers	DMVFollowers
34	livestream	mbarilla
42	nytimes	BuzzFeed
91	rickygervais	rickygervais
115	MikeBloomberg	MikeBloomberg
147	ASPCA	LindaFB
163	CP24	DopeHNIC
226	TheIlluminati	GDominico

and gatekeeper roles were combined and defined as community leaders.

The community leaders were the members of traditional media, and primarily focused on the diffusing the information - action attribute of 'propagate the warning', and the other community members were following the leaders by either taking the prescribed action or propagating the warning. When the leaders were issuing information to evacuate, actionable event - 'propagate the warning' and polarity - 'true', the rest of the community followed one of two actions - 'propagate the warning' or 'take the prescribed action', with the same polarity. When the lack of overlap between the communities across the timeline was discovered, a significant finding was the presence of the leaders in all time slices. As the members of communities participated in the communication only during a particular time slice, the leaders continued their participation throughout the event. This evidence suggests that Twitter users were gravitating towards the leaders who were sources of information and at the same time in control of the information, i.e. diffusers and gatekeepers.

Next the leaders of the communities were evaluated for the 2012 Hurricane Sandy event. Only the communities of size larger than four were examined. Two days were selected for demonstration of the results are October 28th, the day prior to the landfall in southern New Jersey, and October 29th the day of the landfall. The finding that the single leader serving as diffuser and gatekeeper is consistent for both 2011 Japan tsunami and the 2012 Hurricane Sandy events. In contrast to 2011 Japan tsunami, the broker type leader, i.e. the leader who was high in InDegree value and was high in confirmation actionable tweets, was now present in the communities of size larger than four. This type of leader provided confirmations to other members of communities on Twitter. The list leaders, which emerged in the day prior to the landfall in southern New Jersey and during the landfall for top ten communities can be seen in the Tables IX and X.

TABLE XI. COMMUNITY RESULTS: DAY PRIOR TO THE 2012 HURRICANE SANDY LANDFALL IN SOUTHERN NEW JERSEY

Comm	-/+ Rec	-/+ Seek	-/+ Act
1	365/1366	102/687	641/2100
2	22/692	7/27	3/68
9	2/7	0/250	0/1
16	16/140	8/84	13/3211
33	2/29	20/384	2/8
36	5/65	0/4	0/0
37	0/7	0/6	3/118
38	0/2	0/4	0/903
39	2/7	6/25	107/19
40	2/37	1/7	0/5

TABLE XII. LEADERSHIP RESULTS: DAY PRIOR TO THE 2012 HURRICANE SANDY LANDFALL IN SOUTHERN NEW JERSEY

Leader	-/+ Rec	-/+ Seek	-/+ Act
MikeBloomberg	5/554	4/11	813/1525
NHCAtlantic	22/1161	4/4	1/1
HuriicaneSandy	0/0	0/1081	0/0
ASPCA	3/75	0/0	11/3064
rickygervais	0/0	7/4709	0/0
JamesYammouni	4/918	0/0	0/1
googlemaps	0/5	0/0	0/1080
BBCBreaking	0/0	0/2	0/896
KagroX	14/0	2/8	611/2
jimmyfallon	2/516	1/5	0/3

TABLE XIII. COMMUNITY RESULTS: DAY OF THE 2012 HURRICANE SANDY LANDFALL IN SOUTHERN NEW JERSEY

Comm	-/+ Rec	-/+ Seek	-/+ Act
2	726/14586	332/2208	1424/11234
19	91/1431	50/433	229/6151
29	12/1719	14/233	2/3309
34	9/1238	8/10	4/24
42	124/5503	12/202	24/388
91	15/104	43/7461	1/28
115	13/788	12/131	154/1320
147	28/272	1/20	366/1799
163	34/2194	1/15	0/18
226	14/1605	6/7	0/1

TABLE XIV. LEADERSHIP RESULTS: DAY OF THE 2012 HURRICANE SANDY LANDFALL IN SOUTHERN NEW JERSEY

Leader	-/+ Rec	-/+ Seek	-/+ Act
NHCAtlantic	10/1037	1/0	0/0
fema	0/149	2/4	2/2734
DMVFollowers	0/12	0/53	0/1696
livestream	0/1470	0/0	0/0
nytimes	6/1382	0/19	0/151
rickygervais	2/4	2/3762	0/2
MikeBloomberg	1/247	0/8	29/503
ASPCA	1/112	0/0	184/838
CP24	10/1088	2/0	0/3
TheIlluminati	3/802	1/0	0/0

As previously discussed, the behaviors of the members of the communities varied due to the variabilities of warnings, however, the peaks and valleys in the distributions of the aggregated actions of the community members followed the peaks and valleys of the distribution of leaders' actions. This finding can be demonstrated in Tables XI and XII for the day prior to the landfall in southern New Jersey and Tables XIII and XIV for the day of the landfall.

The Tables show 'Rec' for 'receive the warning', 'Seek' for 'seek confirmation or information', 'Act' for 'take the prescribed action', and (+) for positive polarity and (-) for negative polarity. These results suggest that community members followed the actions of their respective leaders. The first story analysis has been used to evaluate the role of the leaders in the communities and assessing the uniqueness of the information they had shared through out the event. The number of first stories were aggregated per each leader to identify the percentage of the unique information shared by each leader. The result of the analysis suggests that in the days leading up to the landfall in southern New Jersey the leaders of the communities were sharing unique information with their respective communities. During the landfall and the day after the landfall, the information being shared by the leaders was no longer unique and consisted of previously transmitted information. Moreover, the most unique information was being shared by the official sources such as MikeBloomberg and NYCMayorsOffice. This finding suggests that Twitter users who were part of the communities led by the official sources obtain first hand information quicker than the rest of the users on Twitter.

VI. CONCLUSION AND FUTURE RESEARCH

Two different events were evaluated. Events differ in impact areas, time span, and magnitude of impact. The 2011 Japan Tsunami event spanned over just one day with very limited time to respond, whereas, the 2012 Hurricane Sandy spanned over nine days with much more time to prepare and respond. During 2011 Japan tsunami the governmental emergency management organizations made limited use of Twitter. However, the traditional media outlets utilized Twitter extensively to disseminate warnings. In contrast, during 2012 Hurricane Sandy local as well as state and federal governmental emergency management organizations made an extensive use of social media providing a vast majority of unique information to the Twitter users.

To overcome a lack of knowledge of who are the individuals or organizations that disseminate warning information, provide confirmations of an event and associated actions, and urge others to take action, a methodology that combines natural language processing and social network analyses was successfully applied to two data sets collected from Twitter during 2011 Japan tsunami and 2012 Hurricane Sandy. The methodology employed was as follows: (1) assign actionable events to each on-topic tweet using NLP; (2) construct a communication network of tweets associated with actionable events; (3) use the network to discover communities with SNA; (4) extract the leaders of the communities and identify their roles with SNA; and (5) evaluate the behavior of the community members and their leaders using NLP.

The analysis was able to demonstrate that the behavior of the Twitter users was consistent with the issuance of actionable information based on warnings. It was also discovered that members of the same community demonstrate similar behaviors when faced with very limited time to respond and diverse behaviors when faced with longer time to respond. Additionally, the diversity of the levels of impact and prescribed actions also facilitated diverse behaviors among the members of the same communities during 2012 Hurricane Sandy. During 2011 Japan tsunami the leaders of the communities were typically the traditional media who were propagating the warnings and urging the other community members to take the prescribed action. However, during 2012 Hurricane Sandy the leaders of the communities ranged from celebrities, specialized organizations (e.g. various weather reporting agencies), and local, state, and federal emergency management organizations. Moreover, it was discovered that the leaders maintained their role throughout the entire event, while the rest of the community members were present during a selected time period. The communities formed around the information sources - i.e. the leaders. The leaders of the communities during 2012 Hurricane Sandy were able to introduce unique information into the communities, moreover, it was the local official organizations who introduced the majority of the unique information. The uniqueness of the information shared by the leaders peaked prior to the hurricane landfall in southern New Jersey and declined during and the day after the event.

The key contributions of the research consist of the insight into the human behavior on Twitter during two major extreme events. The paper showed how extreme events with different characteristics can prompt different human behavior on Twitter. The research explored collective human behavior and demonstrated that events that allow more time to respond and impact larger territories can result in weaker cohesion in virtual communities on Twitter. The research also conveyed stronger adoption of Twitter by official emergency response organization during 2012 Hurricane Sandy, a year and a half after 2011 Japan Tsunami. The official sources are not only adopting the new technology offered by Twitter, but also become leading information sources on Twitter as evident from leadership and first story detection analyses for 2012 Hurricane Sandy. In future research, the authors will attempt to include additional event attributes - i.e. location, to better understand the impact of emergencies on communities. In addition, this will allow us to study the co-evolution of the behavior of the community and its leaders and the structure of the network throughout an emergency. It will also provide the means to investigate the flow of actionable information and its distortion over time.

Acknowledgment

This material is based upon work sponsored by the Army Research Lab under Cooperative Agreement number No. W911NF-09-2-0053 (NS-CTA), U.S. NSF under the grant number CMMI V 1162409, U.S. NSF CAREER Award under Grant IIS-0953149, U.S. DARPA Award No. FA8750-13-2-0041 in the “Deep Exploration and Filtering of Text” (DEFT) Program, IBM Faculty award and RPI faculty start-up grant. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory, DARPA, the National Science Foundation or the U.S. Government.

References

[1] Y. Tyshchuk and W. A. Wallace, “Actionable information during extreme events - case study: Warnings and 2011 tohoku earthquake,” in *SocialCom/PASSAT*, 2012, pp. 338–347.

[2] D. Mileti and J. Sorensen, “Communiction of emergency public warnings: A social science perspective and state-of-the-art assessement,” in *State-of-the-Art Assessement. Report prepared for Federal Emergency Management Agency. Oak Ridge, TN: Oak Ridge National Laboratory*, 1990.

[3] M. Lindell and R. Perry, “The protective action decision model: Theoretical modifications and additional evidence,” in *Risk Analysis*, 2012, pp. 32 (4): 616 – 632.

[4] A. Hughes and L. Palen, “Twitter adoption and use in mass convergence and emergency events,” in *In Proceedings of the 6th International Conference on Information Systems for Crisis Response and Management (ISCRAM). Gothenburg, Sweden*, 2009.

[5] D. Yates and S. Paquette, “Emergency knowledge management and social media technologies: A case study of the 2010 haitian earthquake,” *International Journal of Information Management*, p. 31 (1): 6?3, 2011.

[6] K. Starbird and L. Palen, “”voluntweeters”: self-organizing by digital volunteers in times of crisis,” in *CHI*, 2011, pp. 1071–1080.

[7] A. Sarcevic, L. Palen, J. White, K. Starbird, M. Bagdouri, and K. M. Anderson, “”beacons of hope” in decentralized coordination: learning from on-the-ground medical twitterers during the 2010 haiti earthquake,” in *CSCW*, 2012, pp. 47–56.

[8] D. M. Romero and J. M. Kleinberg, “The directed closure process in hybrid social-information networks, with an analysis of link formation on twitter,” in *ICWSM*, 2010.

[9] D. M. Romero, B. Meeder, and J. M. Kleinberg, “Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter,” in *WWW*, 2011, pp. 695–704.

[10] B. A. Huberman, D. M. Romero, and F. Wu, “Social networks that matter: Twitter under the microscope,” *First Monday*, vol. 14, no. 1, 2009.

[11] H. Ji and R. Grishman, “Refining event extraction through cross-document inference,” in *ACL*, 2008, pp. 254–262.

[12] H. Li, H. Ji, H. Deng, and J. Han, “Exploiting background information networks to enhance bilingual event extraction through topic modeling,” in *Proc. of International Conference on Advances in Information Mining and Management*, 2011.

[13] Q. Li, H. Ji, and L. Huang, “Joint event extraction via structured prediction with global features,” in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. Sofia, Bulgaria: Association for Computational Linguistics, August 2013, pp. 73–82.

[14] Y. Yang, T. Pierce, and J. G. Carbonell, “A study of retrospective and on-line event detection,” in *SIGIR*, 1998, pp. 28–36.

[15] J. M. Kleinberg, “Bursty and hierarchical structure in streams,” *Data Min. Knowl. Discov.*, vol. 7, no. 4, pp. 373–397, 2003.

[16] H. Li, X. Li, H. Ji, and Y. Marton, “Domain-independent novel event discovery and semi-automatic event annotation,” in *PACLIC*, 2010, pp. 233–242.

[17] J. Weng and B.-S. Lee, “Event detection in twitter,” in *ICWSM*, 2011.

[18] E. Benson, A. Haghighi, and R. Barzilay, “Event discovery in social media feeds,” in *ACL*, 2011, pp. 389–398.

[19] A. Ritter, Mausam, O. Etzioni, and S. Clark, “Open domain event extraction from twitter,” in *KDD*, 2012, pp. 1104–1112.

[20] J. Allan, V. Lavrenko, D. Malin, and R. Swan, “Detections, bounds, and timelines: Umass and tdt-3,” in *In Proceedings of Topic Detection and Tracking Workshop*, 2000, pp. 167–174.

[21] J. Allan, V. Lavrenko, and H. Jin, “First story detection in tdt is hard,” in *CIKM*, 2000, pp. 374–381.

[22] S. Petrovic, M. Osborne, and V. Lavrenko, “Streaming first story detection with application to twitter,” in *HLT-NAACL*, 2010, pp. 181–189.

[23] LDC, “Ace (automatic content extraction) english annotation guidelines for events (http://projects.ldc.upenn.edu/ace/docs/english-events-guidelines_v5.4.3.pdf),” 2005.

[24] C.-C. Chang and C.-J. Lin, “Libsvm: A library for support vector machines,” *ACM TIST*, vol. 2, no. 3, p. 27, 2011.

[25] F. J. Och and H. Ney, “A systematic comparison of various statistical alignment models,” *Computational Linguistics*, vol. 29, no. 1, pp. 19–51, 2003.

[26] P. Indyk and R. Motwani, “Approximate nearest neighbors: Towards removing the curse of dimensionality,” in *STOC*, 1998, pp. 604–613.

[27] A. Lancichinetti, F. Radicchi, J. J. Ramasco, and S. Fortunato, “Finding statistically significant communities in networks,” *PloS one*, vol. 6, no. 4, p. e18961, 2011.

[28] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3, pp. 75–174, 2010.

[29] P. Pons and M. Latapy, "Computing communities in large networks using random walks," *J. Graph Algorithms Appl.*, vol. 10, no. 2, pp. 191–218, 2006.

[30] R. Burt, N. Lin, and K. Cook, "Structural holes versus network closure as social capital," in *In Social Captial: Theory and Research. Aldine Transaction*, 2011.

[31] R. Reagans and B. McEvily, "Network structure and knowledge transfer: The effects of cohesion and range," in *Administrative Science Quarterly*, 2003, p. 48 (2)(June): 240?67.

[32] S. Wasserman and K. Faust, "Social network analysis: Methods and applications," *Cambridge University Press*, 1994.

[33] P. Bonacich, "Power and centrality: A family of measures," *American Journal of Sociology*, vol. 92, pp. 1170–1182, 1987.

[34] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1979.

[35] ——, "The gatekeeper, pair-dependency and structural centrality," *Quality and Quantity*, vol. 14, no. 4, pp. 585–592, 1980.

[36] D. Conway, "Social network analysis in r."

[37] L. Ewing, "The tohoku tsunami of march 11, 2011: A preliminary report on effects to the california coast and planning implications," in *California Coastal Commission Report. San Francisco, CA: State of California – Natural Resources Agency*, 2011.

[38] C. Blair, "Update: Hawaii tsunami damage in "tens of millions" of dollars," in *Honolulu Civil Beat*, March 14, 2011.

[39] "Billion-dollar weather/climate disasters," in *National Climatic Data Center and National Oceanic and Atmospheric Administration*, January 12, 2014.

[40] "Hurricane sandy: Timeline," in *Federal Emergency Management Agency*, January 12, 2014.