Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy

Article *in* Natural Hazards · October 2017

DOI: 10.1007/s11069-017-2960-x

2 authors:

Bairong Wang
Shanghai Maritime University

Jun Zhuang
University at Buffalo, The State University of New York

Nat Hazards
DOI 10.1007/s11069-017-2960-x

ORIGINAL PAPER

Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy

Bairong Wang[1] · Jun Zhuang[1]

Received: 10 November 2016 / Accepted: 3 June 2017

Abstract Social media has been widely used for crisis communication during disasters, and its use during extreme events has drawn attention from both researchers and practitioners. Although crisis information coverage and distribution speed are important issues, both have not been studied extensively in the literature. This paper fills this gap by studying information distribution and coverage of social media during disasters. To this end, we searched and analyzed 986,579 tweets posted during Hurricane Sandy (October 22 to November 6, 2012). To learn about responses from official agents, we sampled 163 governmental organizations (GO), 31 non-governmental organizations (NGO) and 276 news agent accounts and their tweets for analysis. Specifically, five social media key performance indicators (KPIs) are studied in this paper, including impression, like, mention, re-tweet, and response time, and other variables such as hashtag, tweet frequency, and information type. We also test whether the five KPIs and other variables are different among different user types. Results show that total impression, re-tweet rate, hashtag, and tweet frequency are significantly ($P < 0.05$) different among different user types. Specifically, although news agent users generate a larger number of total impressions and tweet more frequently than GO and NGO users, their re-tweet rates and number of hashtags are lower than the GO and NGO users. Re-tweet rate based on mentioned users (5%) is significantly higher ($P = 0.00$) than that based on regular followers (0.01%). Nearly 89% of total impressions are generated from regular followers, with impressions from re-tweeting being a minority. This paper provides some new insights into how social media was used for crisis communication during disasters.

Keywords Crisis communication · Social media · Hurricane Sandy · Content analysis

✉ Jun Zhuang
jzhuang@buffalo.edu

[1] Department of Industrial and Systems Engineering, University at Buffalo, Buffalo, NY, USA

Published online: 20 June 2017

1 Introduction

Natural disasters that originated from extreme weather events have been in an increasing trend in recent years (Sena et al. 2014). Timely communication is vital during natural disasters. Social media, in particular, has become popular channels for crisis communication and plays complementary roles to traditional media (Takahashi et al. 2015). It has gained high attention as well as development from an information "black channel" [i.e., a secret, unofficial, or irregular means of communication (McCarthy and boyd 2005)] in 2005 to a formal communication channel adopted by the U.S. government agencies during the 2010 Haiti earthquake (Yates and Paquette 2011). As one of the most widely used social media, Twitter has been used for information distribution by official agents during many disasters, such as the 2010 earthquake in Haiti and the 2012 Hurricane Sandy in the U.S. Considering distribution effectiveness would be essential for official agents to deal with highly dynamic, uncertain and extreme situations (Robinson and Brown 2005; Zhang et al. 2011). Many questions still remain unanswered. Have their tweets reached most of the social media users? How long does it take for the information to reach the users? Those research questions have not received sufficient attention (Rimstad et al. 2014). Therefore, a detailed analysis of information distribution would improve the crisis communication during disasters, which is the research gap that this paper is motivated to fill.

We provide a case study on Hurricane Sandy, the largest hurricane in the 2012 Atlantic hurricane season and the second costliest natural disaster in the U.S. history (Blake et al. 2012). It was formed on October 22, 2012, and dissipated on November 2, 2012. Statistics from the National Hurricane Center (Blake et al. 2012) show that Hurricane Sandy caused 147 direct casualties and brought over $50 billion damage. Numerous governmental and non-governmental agents worked together and responded to the disaster. In addition to physical supports, timely and updated information serves as an intangible help, which can reduce uncertainty and fear, and contribute to timely decisions for more efficient evacuations and reliefs (National Governors 1979; McCarthy and boyd 2005).

The rest of this paper is organized as follows: Sect. 2 summarizes related literature; Sect. 3 introduces the variables studied in this paper including social media key performance indicators (KPIs); Sect. 4 introduces research method and data; Sect. 5 presents analysis results; and Sect. 6 concludes.

2 Literature review

Uprising popularity of social media use in disasters has changed the way how information is generated, shared, and disseminated (White et al. 2009; Alexander 2014). It is believed that social media gains advantages over traditional mass media for disaster communication in terms of its greater capacity and interactive two-way communication (Jaeger et al. 2007; Fraustino et al. 2012). Although rumors and credibility issues exist in social media applications (Kostka et al. 2008; Castillo et al. 2011; Friggeri et al. 2014), both practitioners and researchers are optimistic regarding better application of social media in disaster risk communication (Lindsay 2011; Williams et al. 2012). Therefore, a great amount of work has been conducted to investigate the use of social media during disasters (Houston et al. 2015; Abedin et al. 2014; Lundgren and McMakin 2013), including crisis communication (Bruns and Burgess 2014), information credibility (Gupta and Kumaraguru 2012; Spence et al. 2015), situation awareness (Vieweg et al. 2010; Yuan et al. 2013), and information system (Okada and Ogura 2014; di Tada and Large 2010). These works contribute to better social media application in disaster management, where a larger demand for updated information is produced given the diminished communication capacity and increased threats.

However, existing research mainly focuses on how different factors (e.g., number of followers, user's position in social network) impact information dissemination within social media networks (Huberman et al. 2009; Shuai et al. 2012), and how credibility of information or re-tweet number impact distribution behavior during disasters (Ha and Ahn 2011; Berger and Milkman 2012; Li and Sakamoto 2015). Research on information coverage and distribution efficiency on social media use is limited. In addition, given the potential rumor issues of social media during disasters, official response agents are expected to update public users with correct information to combat rumors. Consequently, the research presented here fills the research gap by a case study of Hurricane Sandy to study how official agents disseminate crisis information and how well their information reaches public users on social media. We mainly adopt social media KPIs from the literature of communication and marketing to measure distribution performances of official users, which have not been used by similar case studies in disaster management research (Takahashi et al. 2015; Acar and Muraki 2011; Al-Saggaf and Simmons 2015).

3 Social media key performance indicators (KPIs)

Theoretical framework of this study is the KPIs, which refer to a set of organizational performance measures that are most critical for an organization's current and future success (Parmenter 2015). The same theoretical framework has also been used in marketing to measure attractiveness and profitability of the social media activities (Clifton 2012). In this paper, five KPIs are analyzed, including impression, like, mention, re-tweet, and response time, to measure how tweets from official agents affect their audience, and their effectiveness and efficiency in terms of information sharing and engagement.

3.1 Impression

Impression refers to the number of times a user is served a tweet in the timeline or search results (Twitter Support Center 2017) and is used to measure delivery traffic of the tweet information among social networks. Impression will be generated regardless of which users the information reaches. There are two types of impressions defined in this study: (1) first impression, which is generated by direct followers of the original tweet authors, and (2) second impression, which is generated by the followers of re-tweeters, in which the generation depends on distribution behavior. Note that impression in this paper is defined as the estimated number of impressions a tweet could generate, which is different from the "impression" from Twitter Analytics, which refers to the actual number of impressions that a tweet generates. Since the data of Twitter user logging or reading history is not available to the public, there is no way to get the exact number of impressions. Therefore, an impression means the tweet information is delivered to this user regardless of whether the user reads it or not. An estimated number of impressions is used in this study.

3.2 Like

Like refers to the number of times a user liked the specific tweet (Twitter Support Center 2017). The number of likes that a tweet gets will be used to measure how other users value this information. People are more likely to like a tweet when the information is useful or interesting. Figure 1 shows an example of a tweet unit, which was posted by user @HumanityRoad at 7:48 AM, October 25, 2012, and received 1 like.

3.3 Mention

A mention refers to a tweet containing @username (another user's name) anywhere in the body of the tweet (Twitter Help Center 2017b), where the name of the mentioned users will be tagged by "@" to differentiate it from general text. An example of a mention occurs in Fig. 1, where users @noaa and @NOAASatellites were mentioned in one tweet posted by user @HumanityRoad. Therefore, both @noaa and @NOAASatellites would be notified that they got mentioned in another user's tweet when they first logged onto the Twitter service after the mention. The more mentions a tweet contains, the more people are engaged. To note that, although two users (@noaa and @NOAASatellite) are mentioned at the same time in the example tweet, they are mentioned in different manners. The first user @noaa was mentioned due to re-tweet behavior of user @HumanityRoad and characterized by the following characters of "RT." We define this type of mention as "re-tweet mention." The second user @NOAASatellites was mentioned because the user @noaa shared this tweet with user @NOAASatellites. We categorize this type of mention as "sharing mention."

3.4 Re-tweet

A re-tweet is a re-posting of another tweet and helps users quickly share that tweet with all of their followers (Twitter Help Center 2017a). Re-tweeting can distribute tweet information to a

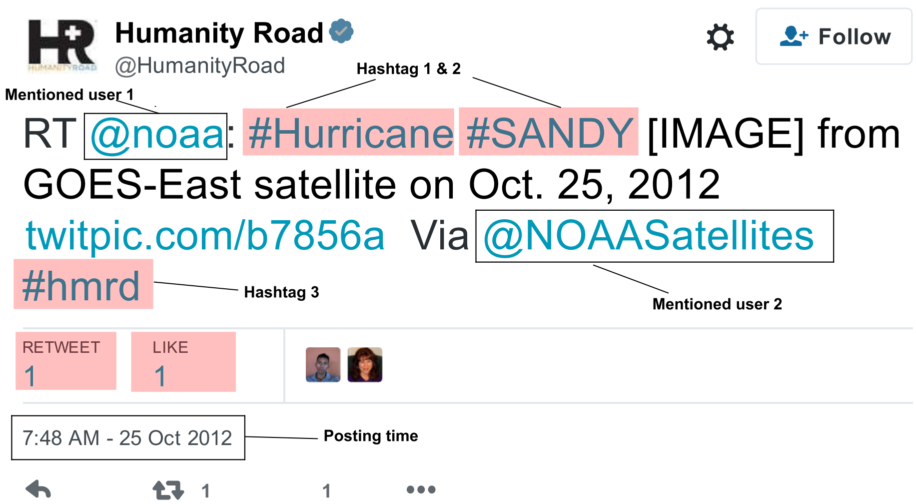

Fig. 1 Example tweet from Humanity Road (Humanity Road 2012)

larger number of users and hence improve the impression of the original tweet. Only valuable tweets would be re-tweeted. There are two types of re-tweet in Twitter service. The first type is re-tweeting without any modification or comments and marked by "RT" to indicate that it was originally generated from other users (Twitter Help Center 2017a), such as the tweet in Fig. 1, which was re-tweeted from user @noaa by user @HumanityRoad. The second type of re-tweet occurs with comments or any modifications of re-tweeters and is indicated by "MT" in the beginning of a tweet. Only the first system re-tweet is studied in this research for response time, impression, and mention analysis.

3.5 Response time

We define response time as the length of time between posting of the original tweet and it's re-tweet(s), and use such time to measure response efficiency as well as the speed of information dissemination.

4 Research method

Content analysis (CA) is originated from communication and media research to derive quantitative information from a body of open-ended qualitative data by analyzing the frequency or occurrence of ideas, or codes (Berg et al. 2004; Neuendorf 2002, 1990). In this study, we apply CA to categorize tweet information types to identify major social media uses during disasters. CA is good at pattern discovery because it provides reliable results by a coding scheme that at least two coders participate in the analysis process. Specifically, official users were first sampled, and then, all of their tweets were analyzed using CA.

4.1 Search and sampling

We analyze tweets posted during Hurricane Sandy from October 22 to November 6, 2012. We applied keywords "Hurricane Sandy" to search tweet and limited the language to English. The search result is summarized in Table 1. "Verified tweet" is the number of tweets posted by verified users. The search yields 16,250 verified users and 50,182 verified tweets in total. Figure 2 illustrates the number of tweets, number of users, and the average number per user tweets. However, the average number of tweets per user posted first increases from October 22 to October 29 (with a peak of 1.39) and then decreases afterward to 1.19 on November 6.

With a random number generator, we tag every verified tweets a number randomly drawn from a uniform probability distribution between 0 and 1 and select only tweets with tag numbers less than 0.1. As a result of this, 1600 users were sampled from the 16,250 verified users for analysis and classified into the following four user types: (1) normal users; (2) governmental organizations (GO); (3) non-governmental organizations (NGO); and (4) news agents. This study focuses on crisis information distribution behavior of official agents. Therefore, only three types of users (GO, NGO, and news agent) and their corresponding tweets are analyzed. GO users include military agents, fire departments, and police departments. For NGO users, we study those related to disaster response work, such as disaster response agency, charitable organizations. We exclude those NGOs, such as business Web site account. We got 163 GO users, 31 NGO users, and 276 news agent users.

Table 1 Number of searched tweets and users in each day

Day (2012)	Individual user	Verified user	Verified tweet	No. of tweet
22-Oct	61	5	5	62
23-Oct	609	15	20	685
24-Oct	4473	396	609	5962
25-Oct	11,863	1032	1742	16,158
26-Oct	23,198	1700	2983	31,216
27-Oct	27,005	1195	2032	34,669
28-Oct	103,591	3160	5820	132,783
29-Oct	249,506	8175	16,578	346,983
30-Oct	118,859	4316	6894	157,124
31-Oct	51,233	2411	3430	66,725
01-Nov	38,767	1978	2686	49,281
02-Nov	39,862	2108	2842	50,442
03-Nov	21,303	860	1117	26,914
04-Nov	18,682	701	897	22,870
05-Nov	21,815	1229	1594	26,723
06-Nov	15,076	770	933	17,982
Total	561,224	16,250	50,182	986,579

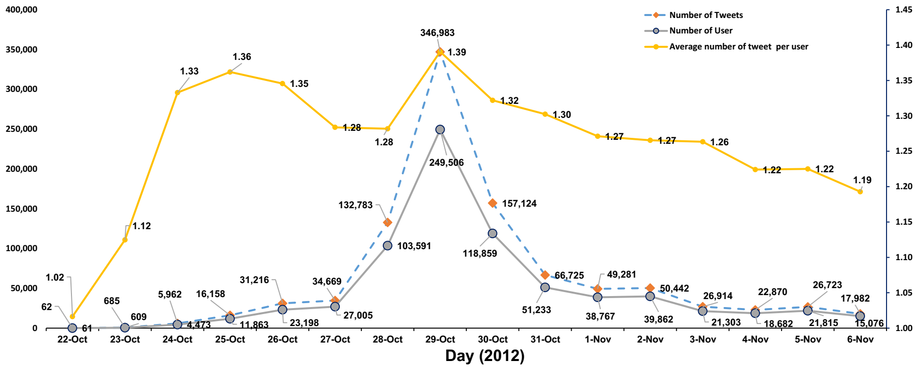

Fig. 2 Tweet frequency, number of tweets and users from Oct 22-Nov 6, 2012

4.2 Content analysis

We apply both manifest and latent CA to the sampled tweets. Manifest CA was applied to get manifest features of tweet data, such as the number of re-tweets and response time (Neuendorf 2002). Latent or manual CA was used to learn social media uses of official users (Neuendorf 2002). The coding process is applied by two coders at the same time, which enables a more accurate discovery of latent, quantifiable patterns in open-ended data (Rourke and Anderson 2004). In our study, both user types and tweet information types are coded. Two coders were trained by a sample with 100 tweets and 200 users. The first reliability test came to 0.73 for tweet information types and 0.91 for user types. After further training, reliability for tweet information was improved to a satisfactory level of 0.87.

Analyzing unit determines the way the data set is analyzed and categorized by the coding schema (De Wever et al. 2006). In this study, we take both user and tweet as analyzing units. For each unit, there are several variables to illustrate their features and characteristics.

User For each user, we analyze their information sharing behaviors during Hurricane Sandy. Variables for user analysis include user type, agent type, number of followers, and tweeting frequency.

Tweet Each tweet was analyzed on the basis of 5 social media KPIs (like, mention, re-tweets, response efficiency, and impression) as well as two other descriptive variables (hashtag and information type). Response efficiency and exposure KPIs require detailed information of the original user's followers and even followers of re-tweeters if the tweet gets re-tweeted.

4.3 Variables

Both manifest and latent variables are used in our content analysis for a thorough and multi-aspect understanding of the information distribution during Hurricane Sandy. Manifest variables are those that can be derived directly (Neuendorf 2002) and describe the surface structure and information (Berg et al. 2004), while latent variables require a deep analysis of the structural meaning of a message and further translation of the data (Berg et al. 2004). For each of the analysis units, we list both the manifest and latent variables analyzed in our study, shown in Table 2.

Table 2 Analyzed variables for users and tweets

Units	Categories	Variables	Definitions
User	Manifest	User type	Indicating the type of official users
		Agent type	Type of an agent based on functions, such as fire department
		Follower	No. of followers
		Frequency	No. of tweets posted by a user
Tweet	Manifest	Hashtag	Hashtags of tweets
		Re-tweets	Re-tweet number a tweet gets
		Like	No. of favorites/likes a tweet received
		URL	URLs in tweet text
		Originality	Whether a tweet is self-composed or re-tweeted from other users
		Impression	No. of times that a tweet information is delivered into the social media network
		Mention	The No. of users mentioned in this tweet
		Response efficiency	Time gap between posting time and response time of users
	Latent	Information type	Type of information tweet conveys, indicating social media user

5 Results

Results are presented on the basis of analyzing units and variables.

5.1 User

As discussed in Sect. 4.1, only GO, NGO, and news agent users are analyzed in this study. Variables for users include: (1) User and agent type; (2) Followers; and (3) Frequency.

- *User and agent type* Three types of users are analyzed in this study, namely GO, NGO, and news agents. For GO and NGO users, they are further classified by their functions. Results for GO and NGO user type classification are summarized in Tables 3 and 4, respectively. Due to the same function of news agent accounts, summary of news agent type is not made in this paper. There are 16 types of agents in GO users and 4 types for NGO users.

  Among GO users, information/weather report type takes the biggest portion (15.34%), followed by emergency management, and citizen service/health/environment types (14.72%). Users from local governments take about 10%, including state, county, and town governments. Science/research/survey agents mainly updated the latest situation about Hurricane Sandy with their own analysis results, such as National Oceanic and Atmospheric Administration (NOAA). For NGO users, about 61% of users are humanity agents, such as @BritishRedCross and @FoodForThePoor. Note that there were at least four animal protection agents (namely @bestfriends, @ASPCA, @peta2, and @HumaneSociety) during Hurricane Sandy that provide tweet tips how to arrange pets.
- *Followers* Figure 3 presents follower histograms for each user type, with the bin sizes being 20,000 for all user types. Modeling results show that the number of followers of the sampled users is power-law-distributed in the format of $f(x) = a * x^b$ and the

Table 3 Agent type classification in GO user

Agent type	Count	(%)	Encoding #
Information/weather	25	15.34	# 01
Emergency management/service	24	14.72	# 02
Citizen service/health/environment	24	14.72	# 03
Science/research/survey	20	12.27	# 04
Government	16	9.82	# 05
Police department	13	7.98	# 06
Army/military agents	13	7.98	# 07
Foreign affairs	10	6.13	# 08
Economy/trade	5	3.07	# 09
Agriculture/forest service	4	2.45	# 10
Fire department	3	1.84	# 11
Transportation	2	1.23	# 12
Law enforcement	1	0.61	# 13
Lawmaking agent	1	0.61	# 14
Communication	1	0.61	# 15
Inter-government	1	0.61	# 16
Total	163	100.00	

Table 4 Agent type classification in NGO user

Agent type	Count	(%)	Encoding #
Humanitarian	19	61.29	# 01
Disaster response/relief	6	19.35	# 02
Animal protection	4	12.90	# 03
Private charitable foundation	2	6.45	# 04
Total	31	100.00	

adjusted R-square is 0.83. The average number of followers is about 255,508 for GO user, 292,872 for NGO user, and 732,801 for news agent user. The ANOVA results show that user type has no significant ($P = 0.18$) impact on the number of followers that a user can have. However, medians of the follower for NGO and news agent users are 61,592 and 75,214, which are much higher than that of GO users, which is 15,993 (Table 5).

- *Frequency* Before presenting results for frequency variable, we first define active users as those who tweet at least once during disasters. Among three types of users, only 7.36% of GO users, 9.68% of NGO users, and 28.62% of news agent users are active, as shown in Table 6. The most active agents for GO, NGO, and news agent users are @NYSDHSES (New York State Division of Homeland Security & Emergency Services) with 62 tweets, @HumanityRoad (Humanity Road) with 56 tweets and @HuffingtonPost_hurricane (The Huffington Post) with 170 tweets, respectively. ANOVA results show that users from different types tweet significantly ($P = 0.00$) different. Specifically, the GO and NGO users are less active in updating latest information than news agent users. Results in Table 6 show that GO and NGO users tweet 0.35 and 0.44 tweets per day on average, which is significantly ($P = 0.00$) lower than the average number of tweets for news agent 1.00.

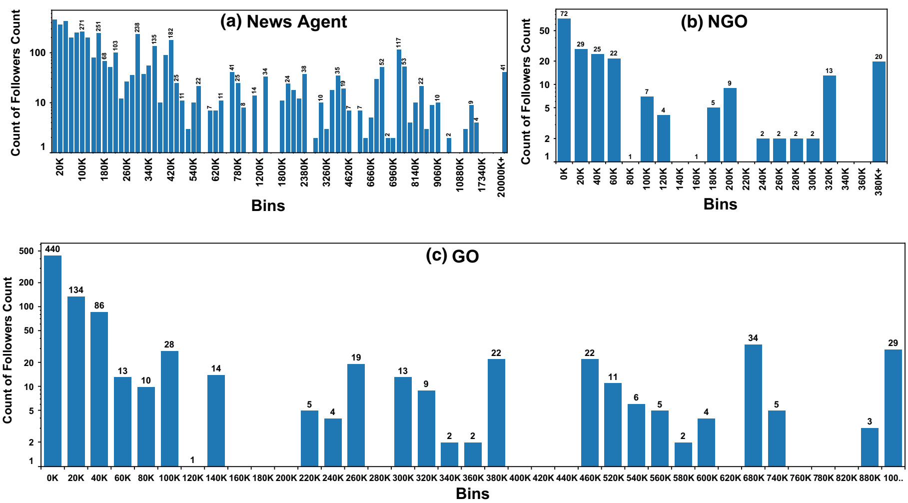

Fig. 3 Histograms of follower numbers for **a** news agents, **b** NGO and **c** GO users

Table 5 Followers number and its distribution modeling results in each user type

User type	Count	Median	Average
GO	163	15,993	255,508
NGO	31	61,592	292,872
News agent	276	75,214	732,801

Table 6 Tweet frequency for GO, NGO and news agent users

User type	User count	Average tweet	Active user (%)	Tweet per day	SD
GO	163	5.66	7.36	0.35	7.70
NGO	31	6.97	9.68	0.44	10.36
News agent	276	15.90	28.62	1.00	21.22

5.2 Tweets analysis

Only tweets from sampled official users were analyzed, in which we received 5528 tweets in total, 923 (16.70%), 216 (3.91%), and 4389 (79.40%) of which were from GO, NGO, and news agent users, respectively, as shown in Table 7.

Hashtag, URL and originality Hashtag is a symbol "#" used before a relevant keyword or phrase (no spaces) in tweet to categorize tweets topics and help hashtagged topics become more popular in Twitter Search. For all 5528 tweets from three types of users, the most frequently used keywords for hashtag are "Sandy" for 2057 times, "hurricanesandy" for 462 times, and "hurricane" for 360 times. To learn hashtag frequency of different types of users, we exclude non-original tweets from hashtag frequency analysis since re-tweets contain hashtags generated by other users. Original tweets numbers for each type of user are summarized in Table 7. Histograms of hashtags for different users are summarized in Fig. 4, showing that GO and NGO users used more words for hashtag than news agent users. There are at most 7 hashtags for all tweets, and the percentages of tweets with less

Table 7 Hashtags, URLs and originality for Tweets

User type	Tweets count	Hashtagged tweets	Hashtagged rate	Original tweets	Original rate	Linked tweets	Linked rate
GO	923	525	0.66	791	0.86	725	0.79
NGO	216	151	0.85	178	0.82	177	0.82
News agent	4389	1984	0.48	4166	0.95	3739	0.85
Total:	5528	2660					

than 3 hashtags in GO, NGO, and news agents are 88.62, 79.78, and 96.16%, respectively. ANOVA analysis for hashtagged original tweets shows that the hashtag number among different user types is significantly different ($P = 0.00$). Specifically, on average, news agents hashtag (with 0.69 words) much less than the GO (with 1.21 words) and NGO (with 1.67 words) users.

Originality is used to indicate whether the tweets are self-composed or re-tweeted from others. Table 7 shows that more than 82% of the tweets are self-composed for each type of user with a higher rate of 0.95 for news agent users. Due to the limits of 140 characters in Twitter service, users prefer to add URL in tweet context for further information. Results show that more than 79% of tweets in each type of user contain URL links. In this sense, URL contributes greatly to communication effectiveness by combining the advantages from traditional media with larger capacity for details and the advantages from Twitter with nearly real-time information posting.

Like Twitter offers "like" function to users for them to mark their favorite tweets. It is also used to measure user engagement. Results for like analysis, in Table 8, show that tweets from GO, NGO, and news agent users receive 4.37, 2.15, and 2.38 likes, respectively, for their tweets on average. ANOVA results show that the number of likes for a tweet is significantly different among different user types ($P = 0.00$). Specifically, tweets from GO users get more likes than those from NGO and news agent users.

Re-tweet and mention Before presenting our results for re-tweet and mention, we introduce the re-tweet data set first. In this study, Twitter API is used to get re-tweet data. However, one limit of the API is that we can get at most 100 system re-tweets for each tweet. Out of 5528 tweets, 4255 were system re-tweeted in our data set and we have 4133 of the 4255 tweets that have at least one of their re-tweets traced. This adds up to 60,283 re-tweets in total. For tweets with less than 100 re-tweets (3900 out of a total of 4133), the percentage of obtained re-tweets to original re-tweet ranges from 25.0% to 100.0%, with an average and median of 88.6% and 94.1%, respectively. For tweets with more than 100 re-tweets (233 out of a total of 4133), the percentage of obtained re-tweets to original re-tweet ranges from 1.2% to 87.7%, with an average and median of 42.3% and 39.1%, respectively. Information can be disseminated more widely by re-tweeting. For example, the larger number of followers that a user has, the more frequently a tweet from that user could be re-tweeted. Mention is another way to distribute information by reminding mentioned user(s) to read the tweet information after being mentioned. Mention is different from re-tweet in that only mentioned user(s) can read tweet information, while re-tweeting enables all followers of the re-tweeter to read the original tweet. We are interested in whether a mentioned user will be more likely to re-tweet the information than regular followers. Therefore, we compare re-tweet rates based on followers and mentioned users. They are denoted by RT (F) and RT (M), respectively. For mentions, only self-composed tweets of

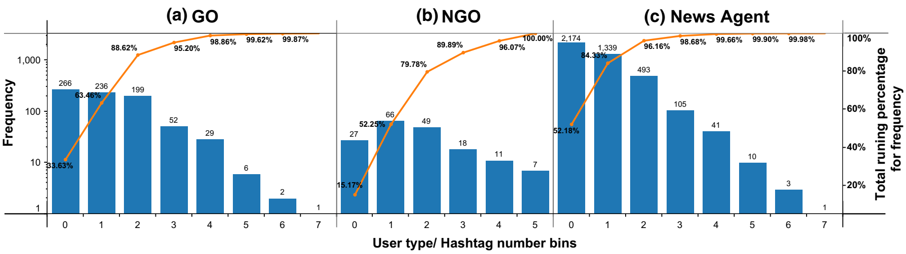

Fig. 4 Hashtag number histograms for GO, NGO and news agent users

Table 8 Likes of tweets from GO, NGO and news agent users

Index	User		
	GO	NGO	News agent
Mean	4.37	2.15	2.38
Median	0	0	0
Mode	0	0	0
SD	21.72	5.54	11.74
Minimum	0	0	0
Maximum	386	55	508
Like sum	4037	465	10,439
Tweet count	923	216	4389

sampled users are analyzed for mention and mention-based re-tweet analysis in order to exclude the mentions from original tweets. Results in Table 9 show that for each tweet, it will be re-tweeted 23.46 times on average by regular followers. The re-tweet rate per tweet is about 0.01%. On average, there are 0.19 mentioned users in one tweet, 0.07 mentioned users on average that would re-tweet. Re-tweet rate based on mentioned users is 5%. ANOVA results show that re-tweet rate based on mentioned users is significantly higher $(P = 0.00)$ than that based on regular follower. Therefore, mention is effective in improving information impression by "borrowing" attention power from other users, especially from those with a large number of followers.

ANOVA test is applied to check whether re-tweet rates based on followers and mentioned users are significantly different among three types (GO, NGO, and news agent) of users . The number of self-composed tweets with mentioned users and at least one re-tweet obtained for GO, NGO, and news agents users is 115, 48, and 471, respectively, as shown in Table 10. Results show that both re-tweet rates (based on regular followers and mentioned users) are significantly different among different user types. The P values are 0.01 for follower-based re-tweet and 0.00 for mention-based re-tweet rate. Follower-based re-tweet rates of news agent users are significantly lower than those of GO and NGO users. GO and NGO users gain advantages to have their tweets re-tweeted compared with news agent users. One possible reason for this is that GO and NGO users may enjoy much more authorities and therefore gain more credibility than general news agents.

Table 9 Re-tweet rates based on mentioned users and regular followers

Index	Followers	RT (F)	RT (F) rate	Mentions	RT (M)	RT (M) rate
Mean	5.38E+05	23.46	0	0.19	0.07	0.05
Median	4.77E+05	3	0	0	0	0
Mode	NA	0	0	0	0	0
SD	2.72E+05	132.21	0	0.52	0.25	0.21
Minimum	363	0	0	0	0	0
Maximum	4.05E+07	6740	0.01	7	1	1
Sum	2.53E+08	1.30E+05	0.62	962	42	33.67
Count	470	5528	5528	5135	634	634

Table 10 Re-tweet rates based on both mentioned users and followers

User type	Regular follower		Mentioned users	
	Tweet count	RT rate	Tweet count	RT rate
GO	923	3.6E−04	115	0.02
NGO	216	2.3E−04	48	0.13
News agent	4389	0.5E−04	471	0.05

Response efficiency For the response efficiency variable, we focus on the time gap between posting time and re-tweet time to learn how fast the tweet information is distributed among Twitter users. Re-tweeting is an effective way to improve information impression, which is also a valuable feature of social media network. However, one question proposed asks how about the distribution speed. For tweet information with short value period, the distribution speed does matter greatly. Figure 5 shows the fitting curve of response time for three types of users based on “hour” unit. More than half of re-tweets happened within an hour, which is quite efficient for information distribution. For those happened long after the original posting time, their distribution may be not useful if the information is out of date. Modeling results show that response time is distributed as general power law function with the form of $f(x) = a * x^b$. Fitting parameters are shown in Table 11. We get the adjusted R-squares as 1.00 for all three types of users, in which we concluded that response time for tweet information could be power-law-distributed.

In addition, we sampled users among the three types of users who have more than 500 re-tweets to model their response time distribution and got 29 qualified users in total. It is reported that all of the users’ response time can be modeled as power law distribution with the same curve fitting method used in response efficiency analysis. Fitting results for 10 of the 29 users are shown in Table 12, and all of the coefficients are within 95% confidence bounds. Response time histograms for 4 individual users, including @NASA, @DYSDHSE, @Huffingtonpost and @Breakingstorm, are illustrated in Fig. 6. We set the upper limit and size for response time bin as 72 and 1 h, respectively. As shown in Fig. 6, most of the re-tweets happened within 1 h, but there could be some re-tweets occurring quite long (144 h) after the original tweet’s posting time. ANOVA results show that user type response efficiency is significantly different among three types of users ($P = 0.01$) and that the average response time of news agent user tweets is 32 h. Those for GO and NGO users are 12 and 16 h.

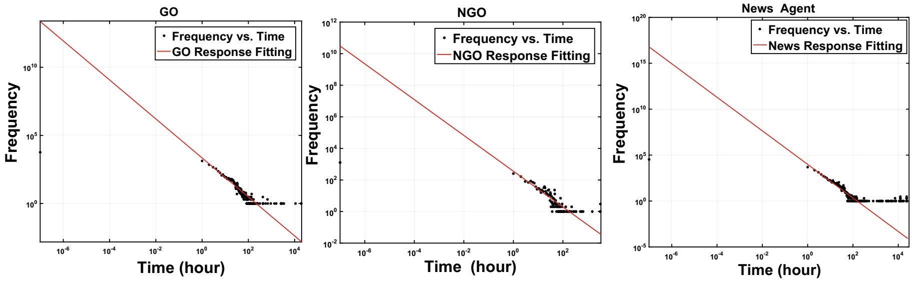

Fig. 5 Curve fitting results for GO, NGO and news agent tweets

Table 11 Response time curve fit results for GO, NGO and news agent users

Coefficients	GO	NGO	News
a	2223	359.30	9813
b	−1.43	−1.13	−1.83
Goodness of fit			
SSE	5.60E+04	3640	4.90E+05
R-square	1.00	1.00	1.00
Adjusted R-square	1.00	1.00	1.00
RMSE	18.83	5.73	47.44

Table 12 Response time curve fit results for 10 individual users

Index	User				
	ABC	AP	Breakingnews	Breakingstorm	Humansociety
Coefficients					
a	5.30	3.32	3.12	1142	22.75
b	−0.24	−0.25	−0.21	−1.90	−0.63
Goodness of fit					
SSE	8055	4308	1035	1.29E+04	381.20
R-square	1.00	1.00	1.00	1.00	0.99
Adjusted R-square	1.00	1.00	1.00	1.00	0.99
RMSE	14.02	11.79	5.15	13.87	2.42
Index	User				
	cnnbrk	Fox29phily	Huffingtonpost	NYSDHSES	WhiteHouse
Coefficients					
a	1040	3.97	6.72	3.52	94.60
b	−1.58	−0.21	−0.26	−0.19	−0.93
Goodness of fit					
SSE	1.29E+04	1476	2.07E+04	23.5	1034
R-square	0.98	1.00	1.00	1.00	1.00
Adjusted R-square	0.98	1.00	1.00	0.99	0.99
RMSE	10.01	10.27	18.57	6.36	3.12

Impression There are two types of impressions in this study: (1) first impression, and (2) second impression. First impression refers to the impression generated from tweet user's followers, and second impression comes from re-tweeter's followers. Hence, first impression equals to the number of followers that a user has. Second impression depends on both the number of re-tweets and the number of followers of the re-tweeters. Figure 7 shows the histograms of second impression for different types of users and for all tweets. More than 46% of second impressions are larger than 1000. ANOVA results show that second impression is not significantly $(P = 0.60)$ different among three types of users. Figure 8 shows the histograms of total impression for different users and all tweets, where GO, NGO, and news agent users have an average of 679, 136, 209, 195, and 1,422,048

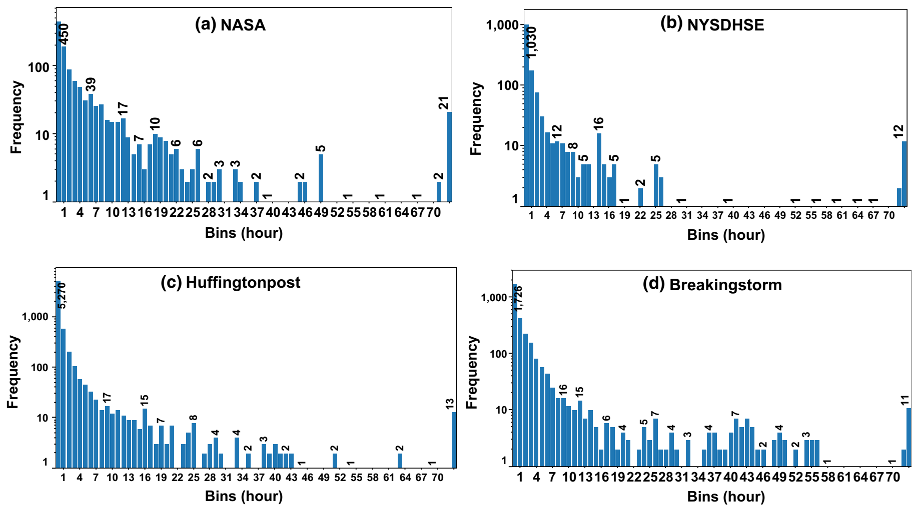

Fig. 6 Histograms of response time for 4 individual users

impressions, respectively, as shown in Table 13. ANOVA results show that there is a significant ($P = 0.00$) difference in total impressions among three types of users. News agent users are most powerful in terms of information distribution, followed by the GO and NGO users. Among all of the impressions, first impression takes about 89%, which means that information distribution depends greatly on followers. The data also show that one retweet can generate 7637 second impressions on average. For the GO users, their information distribution power is less than news agent users, but GO users are regarded as more credible information sources (Shan et al. 2014). However, GO users or NGO users can leverage the distribution power of news agent to reach more public members during disasters.

Information type We coded all tweet text in our data set to learn how social media is used by official agents. Results show that social media was used to distribute 8 types of information, categories and descriptions as well as coding results for three types of users as listed in Table 14. Nearly 70% of tweets are used to inform the public about event updates, information tips, and response report. But for different types of users, their majority uses of social media are different, shown in Fig. 9. For GO users, their tweets focus on information tips and response tips in order to make the public members become more informed about how to respond to disasters and where to get the latest information. For the NGO users, the majority of their tweets are about how people respond to the disasters, then followed by response tips and disaster updates. For news agent users, tweets are mostly used to distribute updated disaster information, and direct people where to get the latest information about Hurricane Sandy, which usually come from their own news accounts.

An interesting finding for news agent users is that they also proactively post tweets to collect information from the public members to check how they respond to disasters and also collect latest pictures/news from the public, which matches the idea of Web 2.0 for two-way communication (Huang et al. 2010). We also have two "eyewitness" news agent users, @ABC11_WTVD and @ABC7NY, collecting news from the public to discover local and interesting news. An example tweet comes from @CBS12, saying that "Send your Hurricane #Sandy storm photos to pics@cbs12.com." But for GO and NGO users,

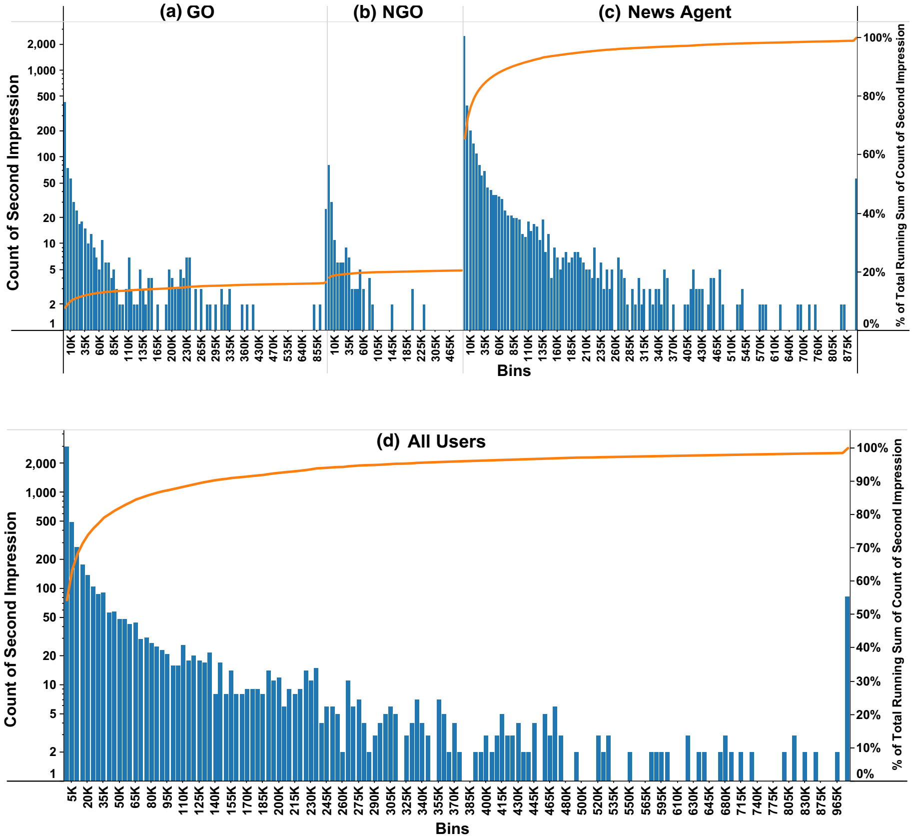

Fig. 7 Second impression histograms

there is no such collection, because they use tweets traditionally for one-way information broadcasting, which is consistent with the findings from (Muralidharan et al. 2011).

6 Conclusion and discussion

Results in this study contribute to the literature on information distribution of official users during disasters, such as re-tweet and social media use. The remainder of this section introduces a summary of analysis results, discussion of both theoretical and practical implications as well as limitations and future research directions.

6.1 Summary

Three types of official users are analyzed in this study, including 163 GO users, 31 NGO users, 276 news agent users and their corresponding tweets. Results for user analysis show that the number of followers is not significantly $(P = 0.18)$ different among user types, but significantly $(P = 0.00)$ different for tweet frequency. On average, each of GO, NGO, and news agent users tweet 5.66, 6.97, and 15.90 times, respectively,

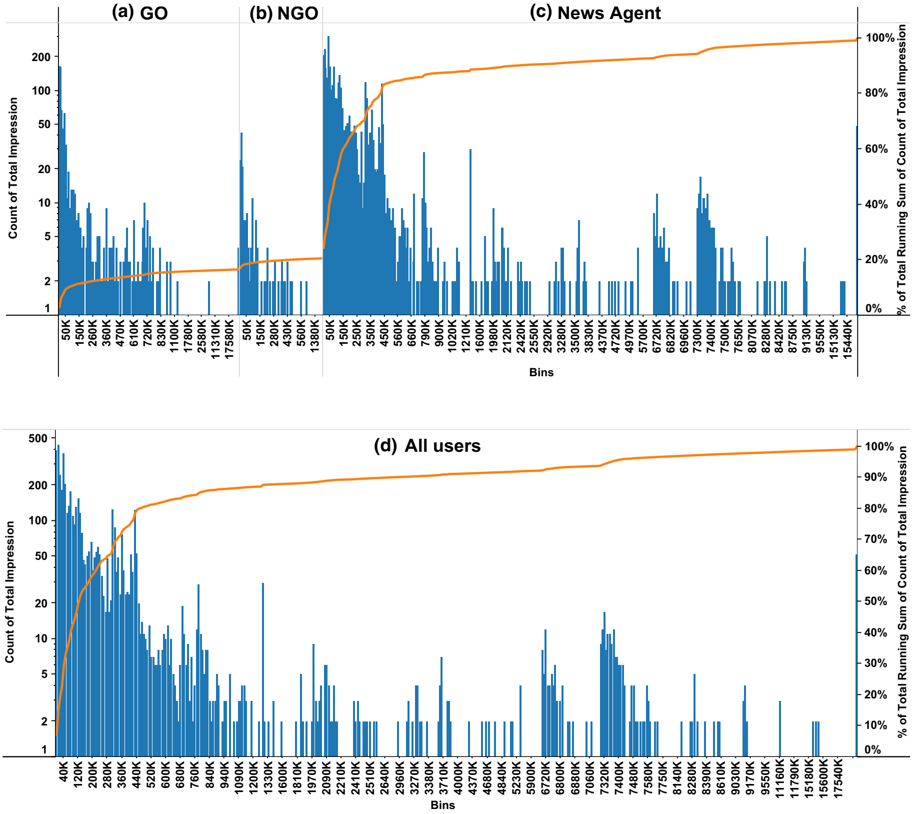

Fig. 8 Total impression histograms

Table 13 Total impression statistics for GO, NGO and news agent users

User type	Count	Sum	Average	Variance
GO	923	6.27E+08	679,136	7.50E+12
NGO	216	4.52E+07	209,195	1.80E+11
News Agent	4389	6.24E+09	1,422,048	2.20E+13

over the 16-day period. The number of followers for the sampled users is power-law-distributed, which is consistent with the literature from social network (Barabási 2009). We categorize GO and NGO users by their functions and get 16 types of agents in GO users, and 4 for NGO users. Among GO users, the number of agents from information/weather report type ranks first and the second is from emergency management type. For NGO users, 61% of users are humanitarian agents, such as "Catholic Charities USA," who are active in disaster response and information distribution.

Results for tweet analysis include hashtag, like, URL, originality, re-tweet, mention, response efficiency, and impression. For hashtag, we find that the news agent users use less

Table 14 Tweet information type during disasters

Information categories	Description	(%)
Disaster/events update	This type of information is used to report updates about the disaster, such as location, wind power	33.93
Information source tips/ recommendations	This use refers recommend the public members where to get the latest disaster information	23.77
Response/preparing report	This use is to report how officials response to or prepare disasters	15.72
Situation report	This type of information is to report people's own situation or how people are affected by this disaster, such as death, power lose	12.40
Response/preparing tips	This use is to offer tips about how to prepare and respond to Hurricane Sandy	9.90
Discussions	Include presidential elections, origin of disaster names, etc.	2.81
Information collection	This use refers to collect how people response to disasters and their situation information	0.83
Expressing wishes and memorializing	This use refers to express emotions and feelings, memorializing victims, etc.	0.32
Ads	This type of tweet is used to advertise products to people, such as insurance, generator	0.13
Requesting help	Request help from other, such as distributing information	0.13
Help tips	This use is to direct people how to offer hands during disasters	0.06

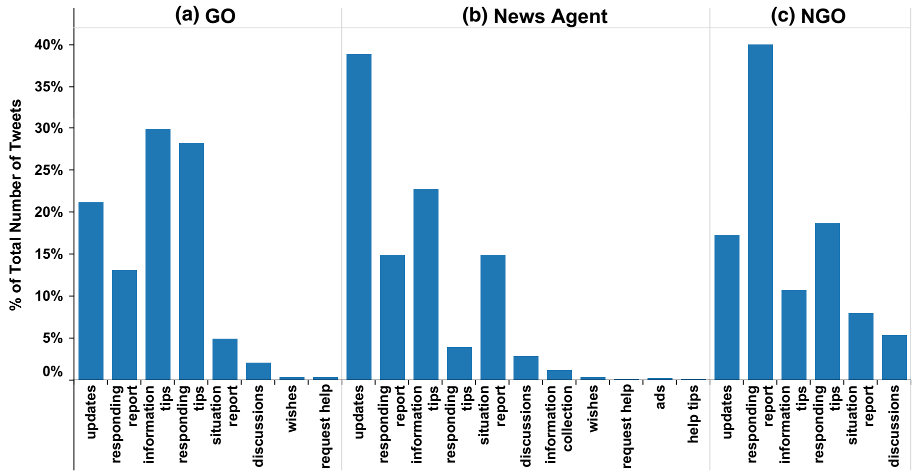

Fig. 9 Tweet information type for GO, NGO and news agent users

words than GO and NGO users ($P = 0.00$) do for hashtag, "#Sandy," "#hurricane Sandy," and "#hurricane" are the three most frequently used keywords for hashtag. For like, ANOVA results show that the like number of tweets received from each types of user is significantly different ($P = 0.00$). For re-tweet and mention, results show that re-tweet rate based on the mentioned users is 0.05, which is significantly higher than the rate of 0.0001 based on regular followers. In addition, re-tweet rates from regular followers are 0.00036, 0.00025, and 0.00005 for GO, NGO, and news agent users, which is significantly

$(P = 0.00)$ different. Re-tweet rate from mentioned users is also significantly $(P = 0.01)$ different among three users types. On average, re-tweet rates based on mentioned users from GO, NGO, and news agent users are 0.02, 0.13, and 0.05, respectively.

For response efficiency, this study takes re-tweet as the response action of users. We find that time gap between re-tweet time and original time is power-law-distributed in three types of users with adjusted R-square values being 1.00. In addition, we model the response time distribution for individual users who have at least 500 re-tweets and get 29 qualified users in total. We find that all of these users' response time can be well modeled with a power law distribution. Nearly 67% of re-tweets occur within 1 h, which is efficient for information distribution. However, the average time is longer than 27 h, meaning that there are some users who re-tweet pretty long after the original tweet posting. For impression variable, our results show that one re-tweet can contribute 7637 impressions on average. Interestingly, we find that the second impression consists of only 11% of the total impressions, which means that the information distribution still heavily depends on first-level followers, instead of the distribution behavior. ANOVA results show that neither the first nor the second impression number is significantly different among three types of users. But tweets from news agent users can gain a significantly larger number of total impressions than tweets from GO and NGO users $(P = 0.00)$.

6.2 Theoretical implications

Our results for follower distribution among three types of users further support the theory that the number of followers are power-law-distributed (Barabási 2009). Besides, our results show that re-tweet time is also power-law-distributed, which has not been studied in the literature. Results show that 67% of re-tweets occur in 1 h, which means that some re-tweets occur in long tail part of power law distribution. The information spreading may not be effective if the original tweet information expires by the time the re-tweet occurs. In this case, re-tweeting behavior may not be efficient if out-of-date information is distributed.

6.3 Practical implications

Results presented here are important for disaster responding officials to spread accurate and latest emergency information to social media users as wide as possible. They could be used by responding agents to improve the performances of crisis information distribution. Specifically, the finding of power law distribution of re-tweet time can be used by responding officials to estimate the number of re-tweets within a period of a time. In addition, with the finding of average number of impressions generated by one re-tweet, a responding agent is able to estimate the number of impressions of their tweets within a period of time to measure the performances of their information distribution on Twitter. However, re-tweet time analysis also indicates some of the re-tweets are not effective at all if the information is out of date by the time of the re-tweet. Therefore, instead of only posting tweets for information sharing, responding agents should also pay attention to the effectiveness of their information distributions. Results also show that the GO and NGO users failed to provide sufficient disaster information during Hurricane Sandy, which is consistent with Barbier et al. (2012). Hence, the responding officials are suggested to post more frequently and timely to avoid losing existing or potential followers (Muralidharan et al. 2011). By information sharing and cooperating among responding agents, attention from different official users could be collected and utilized together to generate a larger number of impressions within a short period of time.

References

Abedin B, Babar A, Abbasi A (2014) Characterization of the use of social media in natural disasters: a systematic review. In: 2014 IEEE fourth international conference on big data and cloud computing (BdCloud), Sydney, NSW, pp 449–454

Acar A, Muraki Y (2011) Twitter for crisis communication: lessons learned from Japan's tsunami disaster. Int J Web Based Communities 7(3):392–402

Alexander DE (2014) Social media in disaster risk reduction and crisis management. Sci Eng Ethics 20(3):717–733

Al-Saggaf Y, Simmons P (2015) Social media in Saudi Arabia: exploring its use during two natural disasters. Technol Forecast Soc Chang 95:3–15

Barabási A-L (2009) Scale-free networks: a decade and beyond. Science 325(5939):412–413

Barbier G, Zafarani R, Gao H, Fung G, Liu H (2012) Maximizing benefits from crowdsourced data. Comput Math Organ Theory 18(3):257–279

Berg BL, Lune H, Lune H (2004) Qualitative research methods for the social sciences, 5th edn. Pearson, Boston

Berger J, Milkman KL (2012) What makes online content viral? J Mark Res 49(2):192–205

Blake ES, Kimberlain TB, Berg RJ, John CP, Beven II JL (2012) Tropical cyclone report: Hurricane Sandy (AL182012), 22–29, Technical report. National Hurricane Center, Miami, FL, USA, p 2013

Bruns A, Burgess J (2014) Crisis communication in natural disasters : the queensland floods and christchurch earthquakes. In: Weller K, Bruns A, Burgess J, Mahrt M, Puschmann C (eds) Twitter and society, digital formations. Peter Lang, New York, pp 373–384

Castillo C, Mendoza M, Poblete B (2011) Information credibility on Twitter. In: Proceedings of the 20th international conference on world wide web, WWW'11, New York, NY, USA, ACM, pp 675–684

Clifton B (2012) Advanced web metrics with Google analytics, 3rd edn. Wiley, Hoboken, NJ

De Wever B, Schellens T, Valcke M, Van Keer H (2006) Content analysis schemes to analyze transcripts of online asynchronous discussion groups: a review. Comput Educ 46(1):6–28

di Tada N, Large T (2010) Information system to assist survivors of disasters. In: 2010 4th IEEE international conference on digital ecosystems and technologies (DEST), pp 354–359

Fraustino JD, Liu B, Jin Y (2012) Social media use during disasters: a review of the knowledge base and gaps. Technical report, National Consortium for the Study of Terrorism and Responses to Terrorism [START]

Friggeri A, Adamic LA, Eckles D, Cheng J (2014) Rumor cascades. In: Proceedings of the 8th international conference on weblogs and social media, pp 101–110

Gupta A, Kumaraguru P (2012) Credibility ranking of tweets during high impact events. In: Proceedings of the 1st workshop on privacy and security in online social media, PSOSM'12, New York, NY, USA, ACM, pp 2:2–2:8

Ha S, Ahn JH (2011) Why are you sharing others tweets?: The impact of argument quality and source credibility on information sharing behavior. In: ICIS 2011 Proceedings

Houston JB, Hawthorne J, Perreault MF, Park EH, Hode MG, Halliwell MR, Turner McGowen SE, Davis R, Vaid S, McElderry JA, Griffith SA (2015) Social media and disasters: a functional framework for social media use in disaster planning, response, and research. Disasters 39(1):1–22

Huang C-M, Chan E, Hyder AA (2010) Web 2.0 and internet social networking: a new tool for disaster management?—lessons from Taiwan. BMC Med Inf Decis Mak 10(1):1–5

Huberman BA, Romero DM, Wu F (2009) Social networks that matter: twitter under the microscope. First Monday 4(1). http://rstmonday.org/ojs/index.php/fm/article/view/2317/2063

Humanity Road (2012) [image] from goes-east satellite on Oct. 25. https://twitter.com/HumanityRoad/status/261479343308537856

Jaeger PT, Shneiderman B, Fleischmann KR, Preece J, Yan Q, Philip Fei W (2007) Community response grids: e-government, social networks, and effective emergency management. Telecommun Policy 31(1011):592–604

Kostka J, Oswald YA, Wattenhofer R (2008) Word of mouth: rumor dissemination in social networks. In: Shvartsman AA, Felber P (eds) Structural information and communication complexity. Springer, Berlin, pp 185–196

Li H, Sakamoto Y (2015) Re-tweet count matters: social influences on sharing of disaster-related tweets. J Homel Secur Emerg Manag 12(3):737–761

Lindsay BR (2011) Social media and disasters: current uses, future options, and policy considerations. Technical report, Washington, DC, USA

Lundgren RE, McMakin AH (2013) Risk communication: a handbook for communicating environmental, safety, and health risks, 5th edn. IEEE, Piscataway, NJ

McCarthy JF, Boyd DM (2005) Digital backchannels in shared physical spaces: experiences at an academic conference. In: CHI'05 extended abstracts on human factors in computing systems, CHI EA'05, New York, NY, USA, ACM, pp 1641–1644
Muralidharan S, Rasmussen L, Patterson D, Shin J-H (2011) Hope for Haiti: an analysis of Facebook and Twitter usage during the earthquake relief efforts. Public Relat Rev 37(2):175–177
(1979) Comprehensive emergency management: a governor's guide. Department of Defense, Defense Civil Preparedness Agency, Washington
Neuendorf KA (1990) Qualitative evaluation and research methods, 2nd edn. SAGE, Newbury Park
Neuendorf KA (2002) The content analysis guidebook. Sage, Thousand Oaks
Okada A, Ogura K (2014) Japanese disaster management system: recent developments in information flow and chains of command. J Conting Crisis Manag 22(1):58–62
Parmenter D (2015) Key performance indicators: developing, implementing, and using winning KPIs, 3rd edn. Wiley, Hoboken, NJ
Rimstad R, Njå O, Rake EL, Braut GS (2014) Incident command and information flows in a large-scale emergency operation. J Conting Crisis Manag 22(1):29–38
Robinson CD, Brown DE (2005) First responder information flow simulation: a tool for technology assessment. In: Proceedings of the 37th conference on winter simulation, winter simulation conference, WSC '05, pp 919–925
Rourke L, Anderson T (2004) Validity in quantitative content analysis. Educ Technol Res Develop 52(1):5–18
Sena A, Corvalan C, Ebi K (2014) Climate change, extreme weather and climate events, and health impacts. In: Freedman B (ed) Global environmental change. Springer Netherlands, Dordrecht, pp 605–613
Shan L, Regan A, De Brun A, Barnett J, van der Sanden MCA, Wall P, McConnon A (2014) Food crisis coverage by social and traditional media: a case study of the 2008 Irish dioxin crisis. Public Underst Sci 23(8):911–928
Shuai X, Ding Y, Busemeyer J (2012) Multiple spreaders affect the indirect influence on Twitter. In: Proceedings of the 21st international conference on world wide web, WWW'12 Companion, New York, NY, USA, ACM, pp 597–598
Spence PR, Lachlan KA, Lin X, del Greco M (2015) Variability in twitter content across the stages of a natural disaster: implications for crisis communication. Commun Q 63(2):171–186
Takahashi B, Tandoc EC Jr, Carmichael C (2015) Communicating on Twitter during a disaster: an analysis of tweets during typhoon Haiyan in the Philippines. Comput Human Behav 50:392–398
Twitter Help Center (2017a) FAQs about retweets (RT). https://support.twitter.com/articles/77606
Twitter Help Center (2017b) What are replies and mentions? https://support.twitter.com/articles/14023
Twitter Support Center (2017) Tweet activity dashboard. https://support.twitter.com/articles/20171990
Vieweg S, Hughes AL, Starbird K, Palen L (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In: Proceedings of the SIGCHI conference on human hactors in computing systems, ACM, pp 1079–1088
White C, Plotnick L, Kushma J, Hiltz SR, Turoff M (2009) An online social network for emergency management. Int J Emerg Manag 6(3–4):369–382
Williams R, Williams G, Burton D (2012) The use of social media for disaster recovery. University of Missouri Extension, Joplin
Yates D, Paquette S (2011) Emergency knowledge management and social media technologies: a case study of the 2010 Haitian earthquake. Int J Inf Manag 31(1):6–13
Yuan W, Guan D, Huh E-N, Lee S (2013) Harness human sensor networks for situational awareness in disaster reliefs: a survey. IETE Tech Rev 30(3):240–247
Zhang G, Yang Y, Mao X (2011) Disaster recovery evaluation PROC model framework based on information flow. In: IEEE 2011 International conference on computer science and network technology (ICCSNT), vol 3. pp 1841–1845