Characterizing information propagation patterns in emergencies
A case study with Yiliang Earthquake

Li, Lifang; Zhang, Qingpeng; Tian, Jun; Wang, Haolin

Published in:
International Journal of Information Management

Published: 01/02/2018

Published version (DOI):
10.1016/j.ijinfomgt.2017.08.008

Publication details:
Li, L., Zhang, Q., Tian, J., & Wang, H. (2018). Characterizing information propagation patterns in emergencies: A case study with Yiliang Earthquake. *International Journal of Information Management*, *38*(1), 34-41. https://doi.org/10.1016/j.ijinfomgt.2017.08.008

International Journal of Information Management 38 (2018) 34–41

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Research Note

Characterizing information propagation patterns in emergencies: A case study with Yiliang Earthquake

Lifang Li[a,b,c], Qingpeng Zhang[b,c,*], Jun Tian[a], Haolin Wang[b,d]

[a] School of Management, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[b] Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong SAR, China
[c] Shenzhen Research Institute of City University of Hong Kong, Shenzhen, Guangdong 518057, China
[d] Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China

ARTICLE INFO

Keywords:
Social networks
Emergency management
Information propagation
Social media analytics

ABSTRACT

Social media has been playing an increasingly important role in information publishing and event monitoring in emergencies like natural disasters. The propagation of different types of information on social media is critical in understanding the reaction and mobility of social media users during natural disasters. In this research, we analyzed the dynamic social networks formed by the reposting (retweeting) behaviors in Weibo.com (the major microblog service in China) during *Yiliang* Earthquake. We developed a Multinomial Naïve Bayes Classifier to categorize the microblog posts into five types based on the content, and then characterized the information propagation patterns of the five types of information at different stages after the earthquake occurred. We found that the type of information has significant influence on the propagation patterns in terms of scale and topological features. This research revealed the important role of information type in the publicity and propagation of disaster-related information, thus generated data-driven insights for timely and efficient emergency management using the publicly available social media data.

1. Introduction

The emergence of social media platforms such as Twitter and Facebook changed the way we create, distribute, and share emergency information during emergencies, like natural disasters and social movements (Palen, Vieweg, Liu, & Hughes, 2009; Shklovski, Palen, & Sutton, 2008; Yates & Paquette, 2010). Social media has been successfully used for a variety of applications for emergency management, including disaster surveillance, disaster-related information retrieval and propagation, behavioral modeling, social support, and mental health management (Chen and Sakamoto, 2014; Fung, Tse, Cheung, Miu, & Fu, 2014; Gruebner et al., 2016; Imran, Castillo, Diaz, & Vieweg, 2015; Lu & Yang, 2011; Plotnick & Hiltz,2016). Among them, users' information exchanging behavior after disaster is particularly important because it could inform the emotions, opinions, and actions of the public towards disasters, and presents the propagation of critical information. It has long been recognized that people close to the event are more interested in emergency services and evacuation procedures, while people far from the event are usually more interested in general details of the earthquake (e.g. casualties) and how they could help those affected by it (e.g. donations) (Chen, Sharman, Rao, & Upadhyaya, 2013; Yom-Tov & Diaz, 2013). A good understanding of such information exchanging behavior is the basis for the research on other applications (e.g. disaster surveillance, social support, etc.), and could further improve the effectiveness and efficiency of information publication by emergency services (Ludwig, Reuter, & Pipek, 2015; Reuter, Heger, & Pipek, 2013).

There is rich literature in examining social media users' information exchanging behavior during emergencies using social media data. A study using Twitter data during Woolwich terrorist attack revealed a number of important factors that positively influence the reposting behaviors, including emotions, time lag of reposting, and the co-occurrence of URLs and hashtags (Burnap et al., 2014). Another study analyzed the copying and pasting behaviors after *Tohuku* earthquake through examining users' reposting behaviors on Twitter, and observed the diminishing effects of retweeting behaviors after disasters (Kim, 2014). A survey based study analyzed the influence of social capital in social media users' information exchange behavior under extreme disaster conditions, and found that the structural capital increases the quantity of information, while cognitive capital and relational capital increase the quality of information (Lu & Yang, 2011). Another content analysis of Chinese Weibo data depicted users' response to different

* Corresponding author at: Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong SAR, China.
E-mail addresses: lifangli2-c@my.cityu.edu.hk (L. Li), qingpeng.zhang@cityu.edu.hk (Q. Zhang), tianjun@mail.xjtu.edu.cn (J. Tian), haolin.wang@my.cityu.edu.hk (H. Wang).

http://dx.doi.org/10.1016/j.ijinfomgt.2017.08.008
Received 3 January 2017; Received in revised form 29 August 2017; Accepted 30 August 2017
Available online 22 September 2017

topics during haze disasters (Wang & Bai, 2014).

In addition to aforementioned empirical analyses of factors affecting information propagation patterns, sub-graphs and network motifs are popular tools to examine such patterns in more detail. Subgraphs of social networks contain rich data to track human interactive activities, thus offered a way to characterize the information propagation patterns with higher resolution (Zhang, Li, Xu, & Vasilakos, 2015). Three-node subgraphs are frequently examined because of their strong interpretability and generalizability of the complex structure of networks (Itzkovitz & Alon, 2005; Milo, Shen-Orr, Itzkovitz, & Kashtan, 2002). The motif of subgraphs (a subgraph that occurred significantly more often in real-world networks than in random networks) is a powerful tool to define elementary structures and identify the underlying interactive activities that generate real-world networks (Milo et al., 2002). For instance, food webs possess motifs that allow the energy to flow from the bottom to the top of food chains (Milo et al., 2002); most social networks usually have motifs that allow the information to be broadcasted from the top to the bottom in the hierarchy of the community (Kempe, Kleinberg, & Tardos, 2015; Leskovec, McGlohon, Faloutsos, Glance, & Hurst, 2007; Zhou, Bandari, Kong, Qian, & Roychowdhury, 2010); human communication networks (like instant messaging) have been found to have reciprocal motifs that enable mutual conversations (Zhang et al., 2015).

The context and content has been proved to have significant influence on whom and how fast people forward information (Sutton, Palen, & Shklovski, 2005; Wang, Tong, & Lin, 2011). To understand how the content of information influences social media users' content sharing behavior during emergencies, a number of studies classified useful disaster-related information to typical types based on its content. In addition to the manual labeling of the types by domain experts, popular supervised learning methods (e.g. Naïve Bayes Classifier, Support Vector Machine, etc.) were widely adopted to label the type of information in large-scale. The popularities of different types of information were found to be heterogeneous in different events (Vieweg, 2012). A follow-up study found that five main types of informative information could contribute to situational awareness and enhance the social resilience during emergencies (Imran et al., 2015; Oh, Kwon, & Rao, 2010). A recent study further examined the interests of users in different types of information, and demonstrated that incorporating specific users' interests could enhance the performance of predicting the scale of information propagation (Hoang & Lim, 2016).

Because of the content sharing nature of information propagation on social media, *social exchange theory* has been widely used to explain the observations. Social exchange theory suggests that relationship decisions is based on the outcomes from social behaviors (Blau, 1964). There are both intrinsic and extrinsic motivations to engage in exchange behaviors. Social media users' online social behaviors (e.g. information propagation and social networking) are typical social exchange behaviors associated with a variety of outcomes, including sense of belonging, reputation, self-esteem, feeling of obligation, altruism, reciprocity, etc. (Ngai, Tao, & Moon, 2015; Shi, Rui, & Whinston, 2014; Wang, Wang, Li, Abrahams, & Fan, 2014). In particular, Twitter is a social system with both social networking and information sharing functions. There are both broadcasting behaviors and mutual conversations among Twitter users, making it an ideal platform to study the social exchanges in social systems (Kwak, Lee, Park, & Moon, 2010; Shi et al., 2014; Wu, Hofman, Mason, & Watts, 2011).

As suggested by social exchange theory, the content of information determines social media users' motivation to exchange the information during emergencies. Exchanging certain types of information could generate rewards for both the community and individual users at certain stages. For example, sharing information about casualties and damage right after an earthquake could help the community get up-to-date information; sharing information about donations in the recovery stage could help the community encourage more donations and improve the transparency of donations, and help individuals reinforce their feeling of obligation and gain reputation. However, research on the dynamic propagation patterns of different types of information after disasters is rare. To address this challenge, this research aims to study the following research questions:

- What are the differences in the propagation scale, speed, and efficiency of different types of disaster-related information?
- What are the differences in the interaction patterns of the social media users when sharing different types of disaster-related information at different stages?

The contribution of this research is twofold. First, by applying text mining and social network analysis techniques, we uncovered the difference in the interaction patterns among five types of disaster-related information propagation networks. Second, subgraph and motif analysis revealed the evolution of communication patterns from uni-directional broadcasting shapes to bi-directional conversational shapes. This data-driven research using social media data could shed light on the in-depth understanding of the information propagation patterns, thus providing critical decision support for retrieving and publishing disaster-related information on social media during emergencies.

2. Data and methodologies

2.1. Data collection

To examine the differences in the information propagation patterns during emergencies, we collected a dataset of Yiliang earthquake[1] (a series of earthquakes occurred in Yiliang on September 7, 2012) from Sina Weibo, a Twitter equivalent in China. We used the public APIs (Application Program Interfaces) provided by Weibo for data collection. We first crawled all *original posts* (9636) containing terms "彝良" (Yiliang) and one of the following two terms "震" (earthquake) and "灾" (disaster) from September 7, 2012 to April 30, 2013. Then, we collected all *reposts* (407,584) of the original posts. Eventually, the dataset contains 417,220 microblog posts, and 315,192 unique user IDs. Among all users, around 8% were residents of directly affected towns/counties, the rest were not directed affected by the earthquake. Therefore, this research investigates the information propagation of the general public, instead of focusing on the specific group of victims. It is worth noting that, we focused on the repost relationship to present information propagations. We did not use the mentioning relationship (the use of @ symbol in text) because: (a) only less than 10% of posts contained mentioning information; (b) the mentioning relationship was almost fully covered by the repost relationship (only one mentioned user was not covered).

2.2. Content annotations and classification

Exchanging different types of information could generate different rewards for people (Constant, Kiesler, & Sproull, 1994). In the context of natural disasters, social media users could not only retrieve useful information to ease their anxiety, but also obtain social approval from posting disaster-related information (Emerson, 1976). For example, exchanging information of casualties could help users and the community to get up-to-date information right after the earthquake occurred, while this type of information would become less important during post-disaster reconstruction phase (a few weeks or months later) (Shklovski et al., 2008). On the other hand, information related to

[1] Yiliang earthquake: On September 7, 2012, a series of earthquakes occurred in Yiliang (彝良) of Yunnan province. Till September 8, 2012, the earthquake had caused a total of 183,000 people affected, including 80 deaths. 7138 houses were collapsed, and 30,600 rooms were damaged.

donations often plays a more significant role in latter phases, because of the need for reconstructions (Chae et al., 2014). Therefore, we need to identify the types of information for an in-depth understanding of the information propagations.

To identify the type of posts, we randomly sampled 1000 original posts as the training set, and then manually labeled them to the following five types, which were derived by adapting domain knowledge and existing classifications from emergency management field (Imran, Castillo, Meier, & Diaz, 2013; Vieweg, 2012). We recruited three graduate students to manually perform the labeling work in two rounds. In the first round, we did a pilot labeling work of 100 posts, and obtained an average Cohen kappa of 0.619, which was within the substantial interval [0.61, 0.80] characterized by (Krippendorff, 2004; Landis & Koch, 2008). In the second round, we labeled the remaining 900 posts based on the agreement obtained from the first round. Eventually, we obtained an improved kappa value of 0.716 for all posts, indicating a high agreement in manual classification results. The five types are as follows:

- Type 1, *Personal* related posts that do not convey useful information to others;
- Type 2, *Caution and Advice* posts containing warning or advice about a possible hazard of an incident;
- Type 3, *Casualties and Damage* posts conveying information of casualties or damages;
- Type 4, *Donation of Money, Goods or Services,* posts discussing donations/goods/services offered to or requested by victims;
- Type 5, *Seeking for Help* posts reporting missing or found personals, and information sources to help people find missing people or provide helps.

Then, we developed a Multinomial Naïve Bayes Classifier using the training set (1000 sampled original posts), and used it to label the type of all the remaining 8636 original posts in the dataset. Each post was only classified to one type. Our classifier obtained a good average accuracy of 80% using 5-fold cross validation. Last, we set the type of reposts the same as the original posts. Manual and automatic classification results are shown in Table 1.

Naïve Bayes Classifier is a well-established statistical model that has been widely used in labeling the type of free text. It has the important advantage in achieving good accuracy with computational efficiency on large scale data (Ting, Ip, & Tsang, 2011). As compared with unsupervised topic modeling methods like the latent Dirichlet allocation model (LDA), the supervised Naïve Bayes Classifier fits this research better because of its capability to classify the content to pre-defined types or topics (identified by domain experts) (Dai, Xue, Yang, & Yu, 2007). It is worth noting that we could enhance the classification performance via incorporating the semantics using ontology-based models. In this research, we adopted Naïve Bayes Classifier for its generality and popularity, instead of focusing on methodological innovation for text mining.

Table 1
Classification results of Yiliang earthquake related posts.

Type of the content	Manual labeling results	Automatic labeling results	# of original posts of each Type
1-Personal related posts and others	365	1510	1875
2-Caution and Advice	39	31	70
3-Casualties and damage	206	2216	2422
4-Donation of Money, Goods or Services	331	4465	4796
5-Seeking for Help	59	134	193
In total	1000	8636	9636

2.3. Network construction

In microblog context (like Weibo and Twitter), the repost/retweet relationship between two users explicitly represents the information propagation from one user (who posted the original information) to the other (who reposted the information). In social network research, it has been well recognized as a reliable way to model and evaluate the information propagation and diffusion in microblog (Guan et al., 2014; Wu & Shen, 2015; Yang & Counts, 2010). Given current text mining capabilities, it is difficult to identify implicit information propagation without such explicit repost/retweet structure. Therefore, we used repost/retweet relations to construct information propagation networks.

In this research, we constructed the networks for the five types of posts to model users' reposting behavior. For each type, we constructed four networks representing the information propagation patterns at four stages according to the commonly used critical rescue time periods (Ambe, Weber, Christ, & Issenberg, 2014; Feng & Wang, 2003; Zook, Graham, Shelton, & Gorman, 2010):

- Stage 1: First 24 h, the "golden" rescue time.
- Stage 2: 25th-72nd hours, the "silver" rescue time.
- Stage 3: 4th day to 3rd month, the critical time point for earthquake recovery.
- Stage 4: 4th-7th month, which is beyond the critical recovery time.

In total, we constructed 20 networks for the five types in four stages. In the constructed information propagation networks, each node represents a unique user ID and a directed edge represents that the sink node reposted one or more original posts from the source node. For example, if user A reposts the original posts of user B, an edge is drawn from node B to node A, indicating that the information propagates from B to A. Fig. 1 visualized the evolution of these networks (using Gephi toolkit https://gephi.org/).

2.4. Extraction of three-node connected subgraphs and motifs

Three-node connected subgraphs in the networks represent the detailed information propagation patterns among three users. Following standard network science practice, we extracted all 13 connected three-node subgraphs and classified them into two main categories: cascading patterns or reciprocal patterns (Lancichinetti, Kivela, Saramaki, & Fortunato, 2010; Milo et al., 2002; Onnela et al., 2012). Cascading patterns refer to the connected subgraphs without reciprocal edges nor loops. Such patterns represent the information propagation from sources to others. C1 to C4 represent four Cascading patterns, in which there is no loop or reciprocal interactions between nodes. R1 to R9 represent nine Reciprocal patterns, which are formed by either immediate reciprocal interaction between two nodes, or a loop. For example, let us assume we have three nodes in the subgraph, node A, node B, and node C. The reciprocal patterns are as follows:

- R1 to R5 contain the immediate reciprocal interactions between two nodes, like A- > B and B- > A;
- R6 to R8 contain both the immediate reciprocal interactions and directed loops. For example, in addition to reciprocal interactions, like A- > B and B- > A, there is also a loop in the subgraph, like A- > B, B- > C, and C- > A;
- R9 represents the pattern with only a loop, but without immediate reciprocal interaction. There is only one pattern in this sub-category: A- > B, B- > C, and C- > A.

Reciprocal patterns refer to the connected subgraphs with reciprocal edges, representing the mutual communications and conversations among users. Fig. 2 visualizes the cascading patterns (red) and the reciprocal patterns (blue). High occurrence of cascading patterns indicates that users mainly repost others' posts without involving mutual

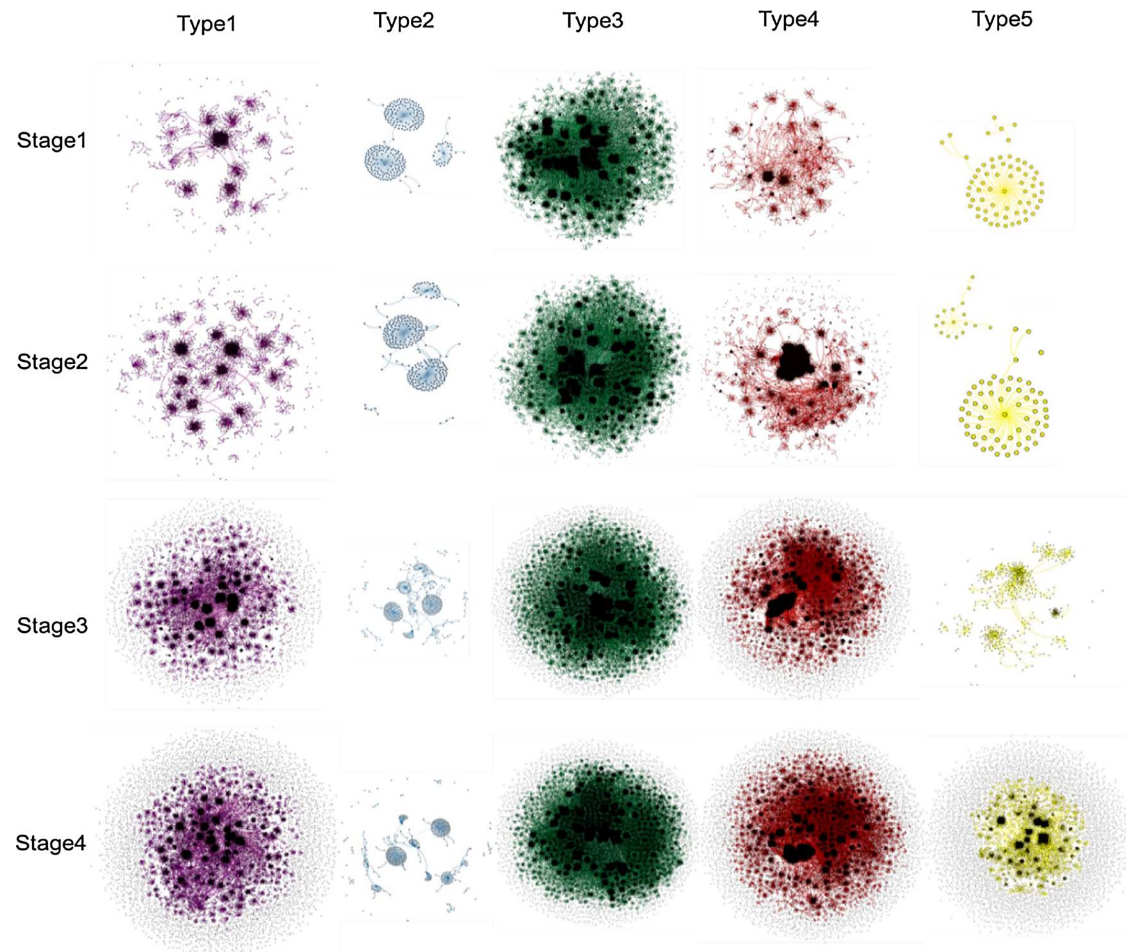

Fig. 1. The evolution of information propagation networks at four stages.

communications. High occurrence of reciprocal patterns indicates the more frequent mutual communications among users (Faraj & Johnson, 2011; Leskovec et al., 2007; Zhou et al., 2010; Wang et al., 2014). Motifs of a network refer to the connected subgraphs that occur significantly more frequent in the specific network than they do in the random networks of similar size (Milo et al., 2002). Therefore, motifs have been widely adopted as a tool to depict the characteristics of real-world networks.

3. Results and discussion

We performed a series of quantitative anayses to answer the proposed research questions through characterizing the scale, growth, efficiency of information propagation networks.

3.1. Scale, growth, and efficiency of information propagation networks

To measure the scale of information propagation, we summarized the number of nodes of each network (Table 2 and Fig. 3A). To measure the growth speed, we calculated the number of new nodes per day of each network (Table 2 and Fig. 3B). To measure the efficiency of information propagation, we calculated the average efficiency of each network, shown in Table 2 and Fig. 3C (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006). The efficiency between two nodes is the inverse of the shortest path length between them. The efficiency of a network is the average efficiencies between all possible nodes.

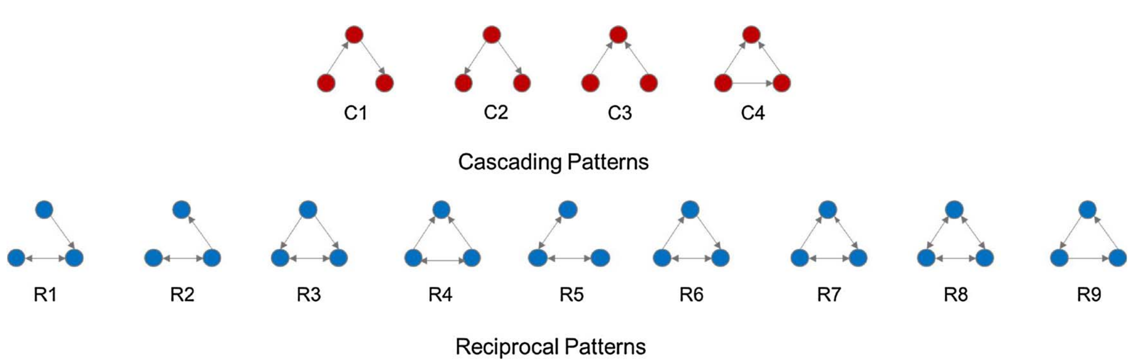

Fig. 2. Visualization of three-node connected subgraphs. Red subgraphs represent cascading patterns. Bluesubgraphs represent reciprocal patterns. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
The scale, growth, and efficiency of the information propagation networks at 4 stages.

		Type1	Type2	Type3	Type4	Type5
Stage1	Scale	1650	286	106781	3536	64
	Growth	1650	286	106781	3536	64
	Efficiency	0.186	0.208	0.123	0.112	0.253
Stage2	Scale	3717	299	132140	42618	77
	Growth	1034	7	12679	19541	7
	Efficiency	0.185	0.206	0.122	0.132	0.253
Stage3	Scale	24065	471	147755	93019	564
	Growth	234	2	179	579	6
	Efficiency	0.098	0.200	0.119	0.1192	0.124
Stage4	Scale	30221	483	169194	106089	19118
	Growth	43	0.08	151	92	130
	Efficiency	0.094	0.200	0.116	0.1114	0.106

Therefore, the larger the efficiency is, the easier it is for information to propagate within the network.

We found that, within the first 24 h of earthquake, the propagation scale of Type 3 information (*Casualties and Damage*) is the largest. It indicates that immediately after the disaster, users tend to retrieve and share up-to-date information of the event through social media. This finding echoed the previous research and indicated that people use social media as information source to ease their anxiety caused by missing information during natural disasters (Oh, Agrawal, & Rao, 2011; Shklovski et al., 2008). In the following stage (25th-72nd h of the earthquake), Type 4 information (*Donation of Money, Goods or Services*) propagated much faster than other types of information, indicating that people focused on granting help after the first 24 h. After 72 h (when traffic, electricity, water supply and communication level are restored), information propagated much slower. However, in the last stage (4–7 months after the earthquake), Type 5 information (*Seeking for Help*) propagated widely, with 71% of the posts related to a teacher, Zhu Yinquan, who was a hero in rescuing students after the earthquake and got sick in January 2013. Social media users widely spread the information of his sickness to seek help and donation for him. In terms of efficiency, we found that all networks except Type 4 were becoming less efficient over time, because of the growth in scale. However, the efficiency of Type 4 network grew at Stage 2, verifying our findings of the popularity of information exchange about donations.

3.2. The interaction patterns in information propagation networks

To figure out how information propagated in detail, we further analyzed the interaction patterns of users to reveal the basic structural elements of the constructed networks. In particular, users' specific behaviors discussing certain types of information were revealed through extracting three-node connected subgraphs defined in Section 2.4 (Fig. 2). The occurrences of different subgraphs at four stages are presented in Fig. 4. Dots on the left of the gray dashed line represent the occurrences of cascading patterns. Dots on the right represent the occurrence of reciprocal patterns. In addition, we further identified network motifs to characterize the unique hidden structural patterns that differentiated these networks from others (shown in Table 3).

We found that, within the "golden" and "silver" rescue time (Stages 1 and 2), users mainly reposted others' posts, as indicated by the high occurrence of cascading patterns, particularly for Type 3 information (*Casualties and Damage*). This observation is consistent with existing research indicating that people mainly used social media as a medium to retrieve information related to the situation awareness of disasters (Oh et al., 2011; Simon, Goldberg, & Adini, 2015). In Yiliang earthquake, information of casualties is clearly the most important information that caused overwhelming information cascades within the "golden" and "silver" rescue time. In particular, the highest occurrence of interaction pattern C2 (A- > B and A- > C) showed that after people tended to "broadcast" critical information to others in order to help the community to fill the information gap of the traditional mass media (Tapia, Bajpai, Jansen, Yen, & Giles, 2011).

After 72 h of the earthquake, more mutual communications and conversations among users occurred for all types of disaster-related information, particularly for Type 4 information (*Donation of Money, Goods or Services*). The high occurrence of the reciprocal interactions (such as, R1, R2, R3, R5, R7, R9) and motifs of Type 4 information at Stages 3 and 4 indicated that when the "golden rescue time" passed, social media users tended to communicate with each other about post-disaster reconstruction, like granting help to victims through donations and other services (e.g. mental health counselling).

Subgraph and motif analysis in this section revealed the evolution of communication patterns from uni-directional broadcasting (cascading) shapes to bi-directional conversational (reciprocal) shapes, indicating the shifting needs of social media users for information exchange at different stages after natural disasters, which will be further discussed in the following section.

3.3. Interpretations with social exchange theory

Social exchange theory has been widely adopted to understand the content sharing and information propagation behaviors. It suggests that social exchange behaviors are mainly driven by a set of key motivations including the expectation of reciprocity, altruism, rationality, sense of belonging, reputation, self-esteem, feeling of obligation etc. (Blau, 1964; Cheung, Liu, & Lee, 2015; Meeker, 1971; Wang et al., 2014). Experiments verified that people have different motivations to share different types of information.

In the emergency management context (like Yiliang earthquake), the motivations of sharing different types of information are diverse. During disasters, people's information sharing behaviors were motivated by multiple rewards, including the improvement of reputation and social status, and the intrinsic satisfaction achieved from helping others (Shklovski et al., 2008; Constant et al., 1994). When the earthquake just occurred (Stage 1), there was a lack of information, especially casualties and damage caused by the earthquake. The altruism, sense of belonging or obligation of people generated a strong motivation for social media users to repost Type 3 information (Casualties and Damage) to help the communities fill the information gap and reduce anxiety (Chiu, Hsu, & Wang, 2006; Oh et al., 2010). As a result, we observed a rapid growth of Type 3 network, and prevalent cascading patterns in the network.

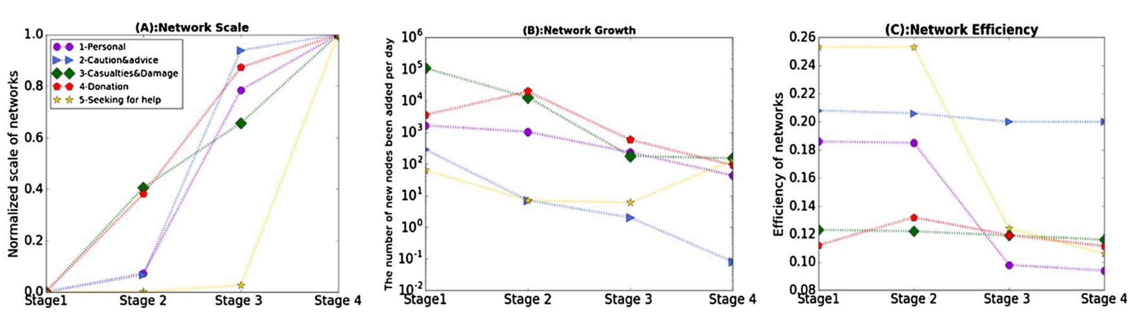

Fig. 3. The number of nodes (Fig. 3.A), network growth (Fig. 3.B), and efficiency (Fig. 3.C) of the information propagation networks at 4 stages.

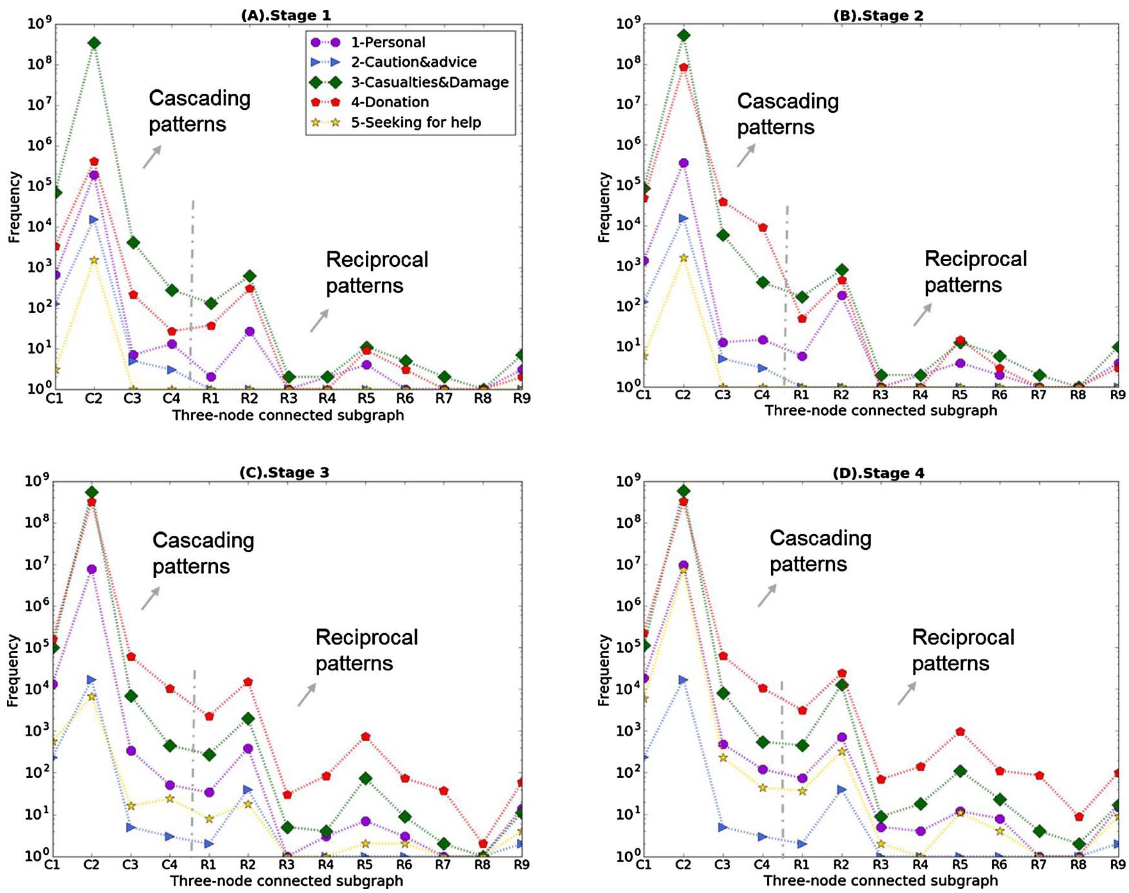

Fig. 4. The frequency of three-node connected subgraphs over time.

Table 3
Network Motifs of 5 information propagation networks over time.

	Type1	Type 2	Type 3	Type 4	Type 5
Stage 1	R2	None	R1, R2	R1, R2	None
Stage 2	R2	None	R1, R2	R2	None
Stage 3	None	R2	R1, R2, R5	R1, R3, R9, R5, R7	None
Stage 4	None	None	R1, R2, R5	R1, R3, R9, R5, R7	R1, R2

In the latter rescue and recovery stages (Stages 2 and 3) when the emergency condition was mitigated, victims (many were homeless) needed help to recover and rebuild from the earthquake (Lovejoy & Saxton, 2012). Such needs inspired the strong feeling of obligation in addition to the altruism and sense of belonging leading to the wide spread of Type 4 information (Donation of Money, Goods or Services). The high occurrence of reciprocal patterns demonstrated that donation related discussions relied on social solidarity and trust (Faraj & Johnson, 2011; Luo, Zhang, & Marquis, 2016; Meeker, 1971). At Stage 4, the sense of obligation made social media users spread the information of the severe sickness of a rescue hero, resulting in the unusual growth of Type 5 information (Seeking for Help).

4. Implications and limitations

4.1. Implications for research

Our study contributes to the literature on emergency management through a novel social media research. First, this study demonstrated the feasibility of (a) adapting text mining methodologies to automatically classify large amount of disaster-related social media content into different types; (b) using social networks to model and analyze the information propagation patterns over time. The combination of text mining and social network analysis enabled us to investigate the detailed propagation scale, speed, and efficiency of different types of disaster-related information over time. This research showed that the adaptation of these new methodologies could help information systems researchers to dive into the large-scale social media data to discover useful data-driven insights, which complement conventional emergency management research.

Second, this study bridged the knowledge gap between the current disaster-related social media research and information propagation patterns in social networks. Past disaster-related social media research focused on context analysis, without examining the detailed information propagation patterns (Imran et al., 2015; Vieweg, 2012, Castillo, Meier, & Diaz, 2013). On the other hand, social network research mainly focused on the propagation patterns of disaster-related information without differentiating the types of information (Boccaletti et al., 2006; Kashtan, Itzkovitz, Milo, & Alon, 2004; Milo et al., 2002; Wang et al., 2014), thus lost substantial details of the potentially diversified propagation patterns. In this study, the content classification enabled us to capture the differences in the information propagation patterns of different types of information. Furthermore, the dynamic network analysis revealed the changes of information propagation patterns over time. Consequently, this research produced new insights that further enriched the applications of social exchange theory and social network analysis in emergency management context. Notably, previous research found that people close to the event were more interested in emergency services, while people far from the event were more focused on details of the earthquake and how to help victims through donations (Ma & Chan, 2014; Yom-Tov & Diaz,2013). This research focused on social media users not directly affected by the earthquake. Through categorizing the information and analyzing its propagation patterns over time, we obtained results that echoed the fact that non-affected people were interested in details of the event (casualties) and ways to help victims (donations). Furthermore, the results extended our understanding of social media users' behavior by revealing the shift of their interests from information about casualties to donations.

Third, the present study examined social media users' interaction behaviors by identifying the subgraphs and motifs in the information propagation networks. The analysis of subgraphs and motifs revealed evolution of communication patterns from uni-directional broadcasting shapes (cascading) to bi-directional conversational shapes (reciprocal). Combining with the finding of interest shift, this research further enriched the understanding of the information propagation patterns on social media through revealing the shift of users' behavior from mainly disseminating information about casualties to forming conversations to discuss topics about donations and post-disaster reconstruction.

4.2. Implications for practice

Social media has been playing an increasingly important role in emergency management. Social media data could be used to enhance the situational awareness during disasters, monitor the public responses, and carry potential courses of action, etc. However, the majority of government authorities are still relying on human power to retrieve such information from social media, without a proper tool to automatically extract key information from such huge amount of data (Plotnick, Hiltz, Kushma, & Tapia, 2015). Decisions in publicizing information on social media is mainly based on the experience of experts.

Although existing research showed that the discussions on social media could help reduce the amount of misinformation and identify critical information during and after crisis (Chen and Sakamoto, 2014), there exist a few technological barriers (e.g. the lack of skills and training) that prevent emergency managers from using social media for information retrieval and dissemination (Plotnick & Hiltz, 2016). A recent survey-based study found the positive attitudes of emergency service staff towards the use of social media in emergencies (Reuter, Ludwig, Kaufhold, & Spielhofer, 2016). This study contributed to the practice through filling this technological gap with automatic content classification and network analysis. The outcome of this research could enable decision makers to identify critical and actionable information from social media for a more effective mitigation and relief programs.

4.3. Limitations and future research

This study has limitations in data and methods, which provide opportunities for future research. First, we classified the type of posts based on the content of original posts, because this study focused on the information propagation. Information could be added or distorted during the reposting process. Therefore, future research may consider performing indepth analysis of the content added by users who reposted the original posts. Second, we constructed the information propagation networks based on the reposting behaviors, while ignoring the mentioning behaviors due to its rare incidence in this dataset. The mentioning behaviors could present another type of information propagation in other dataset. Given sufficient data of both reposting and mentioning behaviors, we could draw conclusions from a more comprehensive perspective. We suggest social media researchers to examine multiple types of relationships in similar studies. Third, we did not examine the friendship network (defined by follower relationships), because of the limitations of Weibo API. In our future research, we will seek collaborations with Weibo to retrieve friendship information, so that we could model the information propagation within an established friendship-based social network. Fourth, the majority of users in our dataset were not directly affected by the earthquake. Thus, this research is focused on the information propagation and exchange patterns of the general public, instead of actual victims. In the follow-up research, we plan to examine the difference in information propagations between victims and non-affected populations.

In addition to the aforementioned limitations, our future research will focus on the classification of information with a higher resolution. We plan to extend this research to identify information for disaster mitigation, malicious information, and rumors, and identify key users with high influence on information propagation. We will also develop proper information publication policies to facilitate rescue and post-disaster reconstruction through minimizing the influence of false information and maximizing the influence of helpful information including donations, mental health advices, etc.

5. Conclusions

This paper characterized the information propagation patterns of different types of disaster-related information during emergencies. Through quantitative network analysis, we found diverse growth patterns of different types of information over time in both the growth of networks and the interaction patterns within networks. Through text mining and social network analysis of publicly available social media data, we uncovered the difference in the interaction patterns among five types of disaster-related information propagation networks. The analysis of interaction patterns with subgraphs and motifs revealed the unique evolution of information propagation patterns from uni-directional broadcasting shapes to bi-directional conversational shapes. From both perspectives, we found the evidence showing the change of people's behavior of using social media during disasters: they first used social media as a source of disaster-related information at beginning of the event ("golden rescue time"), and then as a platform for them to discuss about donations and post-disaster reconstruction.

To the best of our knowledge, this is a first attempt to quantitatively characterize the difference in the information propagations of multiple types of disaster-related information. Outcomes of this research shed light on deriving data-driven insights for emergency management using the publicly available social media data. Formal response agencies (like FEMA of the U.S. and China Earthquake Administration) and humanitarian organizations (like U.N. and NGOs) could adapt similar approaches as proposed in this paper to retrieve useful disaster-related information from social media. In addition, empirical studies of the information propagation networks and the text content could generate practical insights on how to use social media as a medium to publicize information more efficiently. These data-driven insights could help them to disseminate specific types of information to people in need of it, and to increase the effectiveness of donation gatherings, reconstructions, and other activities.

In particular, in the latter stages of post-disaster reconstruction, humanitarian organizations could improve the trust and solidarity among social media users through encouraging mutual communications about donation related information. This research also demonstrates the feasibility of leveraging automatic text mining and social network analysis techniques to fulfill the potential of social media data, thus calls for future research on developing advanced and user-friendly decision support tools for authorities.

Acknowledgements

The authors thank the constructive comments and suggestions from anonymous reviewers. This work was supported by The National Natural Science Foundation of China (NSFC) Grant Nos. 71402157, 71672163, and 71171157, and the Guangdong Provincial Natural Science Foundation No. 2014A030313753.

References

Ambe, P., Weber, S. A., Christ, H., & Wassenberg, D. (2014). Cholecystectomy for acute cholecystitis. How time-critical are the so called golden 72 hours? Or better golden 24 hours ansilver 25–72 hour? A case control study. *World Journal of Emergency Surgery: WJES, 9*(1), 60. http://dx.doi.org/10.1186/1749-7922-9-60.

Blau, P. M. (1964). *Exchange and power in social life.* Transaction Publishers.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(4–5), 175–308. http://dx.doi.org/10.1016/j.physrep.2005.10.009.

Burnap, P., Williams, M. L., Sloan, L., Rana, O., Housley, W., Edwards, A., et al. (2014). Tweeting the terror: modelling the social media reaction to the Woolwich terrorist

attack. *Social Network Analysis and Mining, 4*(1), 1–14. http://dx.doi.org/10.1007/s13278-014-0206-4.
Chae, J., Thom, D., Jang, Y., Kim, S., Ertl, T., & Ebert, D. S. (2014). Public behavior response analysis in disaster events utilizing visual analytics of microblog data. *Computers & Graphics, 38*, 51–60.
Chen, R., & Sakamoto, Y. (2014). Feelings and perspective matter: Sharing of crisis information in social media. *System sciences (HICSS), 2014 47th hawaii international conference on. IEEE,* 1958–1967 [January].
Chen, R., Sharman, R., Rao, H. R., & Upadhyaya, S. J. (2013). Data model development for fire related extreme events: An activity theory approach. *Mis Quarterly, 37*(1), 125–147.
Cheung, C. M. K., Liu, I. L. B., & Lee, M. K. O. (2015). How online social interactions influence customer information contribution behavior in online social shopping communities: A social learning theory perspective. *Journal of the Association for Information Science and Technology, 66*(12), 2511–2521. http://dx.doi.org/10.1002/asi.23340.
Chiu, C.-M., Hsu, M.-H., & Wang, E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems, 42*(3), 1872–1888. http://dx.doi.org/10.1016/j.dss.2006.04.001.
Constant, D., Kiesler, S., & Sproull, L. (1994). What's mine is ours, or is it? A study of attitudes about information sharing, information systems research What's Mine Is Ours, or Is It? A Study of Attitudes about Information Sharing. *Information Systems Research, 5*(4), 400–421. http://dx.doi.org/10.1287/isre.5.4.400.
Dai, W., Xue, G., Yang, Q., & Yu, Y. (2007). Transferring naive bayes classifiers for text classification. *AAAI conference on artificial intelligence,* 540–545. http://dx.doi.org/10.1109/TED.2007.900976.
Emerson, R. M. (1976). Social exchange theory. *Annual Review of Sociology, 2*(1), 335–362.
Faraj, S., & Johnson, S. L. (2011). Network exchange patterns in online communities. *Organization Science, 22*(6), 1464–1480. http://dx.doi.org/10.1287/orsc.1100.0600.
Feng, C.-M., & Wang, T.-C. (2003). Highway emergency rehabilitation scheduling in post-earthquake 72 hours. *Journal of the Eastern Asia Society for Transportation Studies, 5*, 3276–3285.
Fung, I. C. H., Tse, Z. T. H., Cheung, C. N., Miu, A. S., & Fu, K. W. (2014). Ebola and the social media. *The Lancet, 384*(9961), 2207. http://dx.doi.org/10.1016/S0140-6736(14)62418-1.
Gruebner, O., Sykora, M., Lowe, S. R., Shankardass, K., Trinquart, L., Jackson, T., et al. (2016). Mental health surveillance after the terrorist attacks in Paris. *The Lancet, 387*(10034), 2195–2196.
Guan, W., Gao, H., Yang, M., Li, Y., Ma, H., Qian, W., et al. (2014). Analyzing user behavior of the micro-blogging website Sina Weibo during hot social events. *Physica A: Statistical Mechanics and Its Applications, 395*(71101140), 340–351. http://dx.doi.org/10.1016/j.physa.2013.09.059.
Hoang, T. A., & Lim, E. P. (2016). Microblogging content propagation modeling using topic-specific behavioral factors. *IEEE Transactions on Knowledge and Data Engineering, 28*(9), 2407–2422. http://dx.doi.org/10.1109/tkde.2016.2562628.
Imran, M., Castillo, C., Meier, P., & Diaz, F. (2013). Extracting information nuggets from disaster related messages in social media. *Iscram,* 791–800.
Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2015). A processing social media messages in mass emergency: A survey. *ACM Computing Surveys, 47*. http://dx.doi.org/10.1145/2771588.
Itzkovitz, S., & Alon, U. (2005). Subgraphs and network motifs in geometric networks. *Physical Review E – Statistical, Nonlinear, and Soft Matter Physics, 71*(2), 1–9. http://dx.doi.org/10.1103/PhysRevE.71.026117.
Kashtan, N., Itzkovitz, S., Milo, R., & Alon, U. (2004). Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics, 20*(11), 1746–1758. http://dx.doi.org/10.1093/bioinformatics/bth163.
Kempe, D., Kleinberg, J., & Tardos, E. (2015). Maximizing the spread of influence through a social network. *Theory of Computing, 11*(1), 105–147. http://dx.doi.org/10.4086/toc.2015.v011a004.
Kim, T. (2014). Observation on copying and pasting behavior during the Tohoku earthquake: Retweet pattern changes. *International Journal of Information Management, 34*(4), 546–555. http://dx.doi.org/10.1016/j.ijinfomgt.2014.03.001.
Krippendorff, K. (2004). Content analysis: An introduction to its methodology. *Education, 79*. http://dx.doi.org/10.2307/2288384.
Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is twitter, a social network or a news media? *The international world wide web conference committee (IW3C2),* 1–10. http://dx.doi.org/10.1145/1772690.1772751.
Lancichinetti, A., Kivela, M., Saramaki, J., & Fortunato, S. (2010). Characterizing the community structure of complex networks. *PloS One, 5*(8), e11976. http://dx.doi.org/10.1371/journal.pone.0011976.
Landis, J. R., & Koch, G. G. (2008). The measurement of observer agreement for categorical data. *Society, 33*(1), 159–174. http://dx.doi.org/10.2307/2529310. Published by: International Biometric Society Stable URL: http://www. jstor. org/stable/2529310.24.
Leskovec, J., McGlohon, M., Faloutsos, C., Glance, N., & Hurst, M. (2007). *Patterns of Cascading Behavior in Large Blog Graphs.* 1–21. http://dx.doi.org/10.1137/1.9781611972771.60.
Lovejoy, K., & Saxton, G. D. (2012). Information, community, and action: How nonprofit organizations use social media. *Journal of Computer-Mediated Communication, 17*(3), 337–353. http://dx.doi.org/10.1111/j.1083-6101.2012.01576.x.
Lu, Y., & Yang, D. (2011). Information exchange in virtual communities under extreme disaster conditions. *Decision Support Systems, 50*(2), 529–538. http://dx.doi.org/10.1016/j.dss.2010.11.011.
Ludwig, T. C., Reuter, C., & Pipek, V. (2015). Social haystack: dynamic quality assessment of citizen generated content during emergencies. *ACM Transactions on Computer-Human Interaction, 22*(17), 1–27. http://dx.doi.org/10.1145/2749461.
Luo, X. R., Zhang, J., & Marquis, C. (2016). Mobilization in the internet age: internet activism and corporate response. *Academy of Management Journal, 59*(6), 0693. http://dx.doi.org/10.5465/amj.2015.0693.
Ma, W. W., & Chan, A. (2014). Knowledge sharing and social media: altruism, perceived online attachment motivation: and perceived online relationship commitment. *Computers in Human Behavior, 39*, 51–58.
Meeker, B. F. (1971). Decisions and exchange. *American Sociological Review, 36*(3), 485–495.
Milo, R., Shen-Orr, S., Itzkovitz, S., & Kashtan, N. (2002). Network motif: simple building blocks of complex networks. *Science, 298*(5594), 298. http://dx.doi.org/10.1126/science.298.5594.824.
Ngai, E. W. T., Tao, S. S. C., & Moon, K. K. L. (2015). Social media research: Theories, constructs, and conceptual frameworks. *International Journal of Information Management, 35*(1), 33–44. http://dx.doi.org/10.1016/j.ijinfomgt.2014.09.004.
Oh, O., Kwon, K. H., & Rao, H. R. (2010). An exploration of social media in extreme events: Rumor theory and twitter during the Haiti earthquake 2010. *Icis,* 231.
Oh, O., Agrawal, M., & Rao, H. R. (2011). Information control and terrorism: Tracking the Mumbai terrorist attack through twitter. *Information Systems Frontiers, 13*(1), 33–43. http://dx.doi.org/10.1007/s10796-010-9275-8.
Onnela, J. P., Fenn, D. J., Reid, S., Porter, M. A., Mucha, P. J., Fricker, M. D., & Jones, N. S. (2012). Taxonomies of networks from community structure. *Physical Review E – Statistical, Nonlinear, and Soft Matter Physics, 86*(3), http://dx.doi.org/10.1103/PhysRevE.86.036104.
Palen, L., Vieweg, S., Liu, S. B., & Hughes, A. L. (2009). Crisis in a networked world: features of computer-mediated communication in the April 16, 2007, Virginia Tech Event. *Social Science Computer Review, 27*(4), 467–480. http://dx.doi.org/10.1177/0894439309332302.
Plotnick, L., & Hiltz, S. R. (2016). Barriers to use of social media by emergency managers. *Journal of Homeland Security and Emergency Management, 13*(2), 247–277.
Plotnick, L., Hiltz, S. R., Kushma, J. A., & Tapia, A. (2015). Red tape: attitudes and issues related to use of social media by US county-level emergency managers. *Proceedings of the information systems for crisis response and management (ISCRAM)*.
Reuter, C., Heger, O., & Pipek, V. (2013). Combining real and virtual volunteers through social media. *Iscram,* 780–790. http://dx.doi.org/10.1126/science.1060143.
Reuter, C., Ludwig, T., Kaufhold, M.-A., & Spielhofer, T. (2016). Emergency services 'attitudes towards social media: A quantitative and qualitative survey across Europe. *International Journal of Human Computer Studies, 95*, 96–111. http://dx.doi.org/10.1016/j.ijhcs.2016.03.005.
Shi, Z., Rui, H., & Whinston, A. (2014). Content sharing in a social broadcasting environment: evidence from twitter. *Mis Quarterly, 38*(1), 123–142. http://dx.doi.org/10.2139/ssrn.2341243.
Shklovski, I., Palen, L., & Sutton, J. (2008). Finding community through information and communication technology during disaster events. *2008 ACM conference on computer supported cooperative work,* 127–136. http://dx.doi.org/10.1145/1460563.1460584.
Simon, T., Goldberg, A., & Adini, B. (2015). Socializing in emergencies − A review of the use of social media in emergency situations. *International Journal of Information Management, 35*(5), 609–619. http://dx.doi.org/10.1016/j.ijinfomgt.2015.07.001.
Sutton, J., Palen, L., & Shklovski, I. (2005). Backchannels on the front lines: emergent uses of social media in the 2007 southern california wildfires. *Proceedings of the 5th international ISCRAM conference* (pp. 624–632).
Tapia, A. H., Bajpai, K., Jansen, J., Yen, J., & Giles, L. (2011). Seeking the trustworthy tweet: can microblogged data fit the information needs of disaster response and humanitarian relief organizations. *Proceedings of the 8th international ISCRAM conference* (pp. 1–10). . Retrieved from http://irevolution.files.wordpress.com/2011/06/tapia-seeking_the_trustworthy_tweet-161_a.pdf.
Ting, S. L., Ip, W. H., & Tsang, A. H. C. (2011). Is Naïve bayes a good classifier for document classification? *International Journal of Software Engineering and Its Applications, 5*(3), 37–46.
Vieweg, S. E. (2012). *Situational awareness in mass emergency: a behavioral and linguistic analysis of microblogged communications.* 1–300.
Wang, S. Y., & Bai, H. (2014). A case study on micro-blog communication characteristics related to haze during the Harbin haze disaster. *International conference on management science and engineering – annual conference proceedings,* 120–125. http://dx.doi.org/10.1109/ICMSE.2014.6930218.
Wang, D., Tong, H., & Lin, C. (2011). Information spreading in context. *Human Factors, 82*, 735–744. http://dx.doi.org/10.1145/1963405.1963508.
Wang, G. A., Wang, J. H., Li, J., Abrahams, A., & Fan, W. (2014). An analytical framework for understanding knowledge-sharing processes in online Q & A communities. *ACM Transactions on Management Information Systems, 5*(4), 31. http://dx.doi.org/10.1145/2629445.
Wu, B., & Shen, H. (2015). Analyzing and predicting news popularity on Twitter. *International Journal of Information Management, 35*(6), 702–711.
Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who says what to whom on twitter. *Proceedings of the 20th international conference on world wide web,* 705–714. http://dx.doi.org/10.1145/1963405.1963504.
Yang, J., & Counts, S. (2010). Predicting the speed, scale: and range of information diffusion in twitter. *Icwsm, 10*, 355–358.
Yates, D., & Paquette, S. (2010). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *Proceedings of ASIST 2010, 31*(1), 6–13. http://dx.doi.org/10.1016/j.ijinfomgt.2010.10.001.
Yom-Tov, E., & Diaz, F. (2013). The effect of social and physical detachment on information need. *ACM Transactions on Information Systems (TOIS), 31*(1), 4.
Zhang, Y. Q., Li, X., Xu, J., & Vasilakos, A. (2015). Human interactive patterns in temporal networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45*(2), 214–222. http://dx.doi.org/10.1109/TSMC.2014.2360505.
Zhou, Z., Bandari, R., Kong, J., Qian, H., & Roychowdhury, V. (2010). Information resonance on twitter. *Proceedings of the first workshop on social media analytics – SOMA'10,* 123–131. http://dx.doi.org/10.1145/1964858.1964875.
Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: a case study of the haitian earthquake. *World Medical & Health Policy, 2*(2), 6–32. http://dx.doi.org/10.2202/1948-4682.1069.