

A new crowdsourcing model to assess disaster using microblog data in typhoon Haiyan

Qing Deng¹ · Yi Liu² · Hui Zhang¹ · Xiaolong Deng³ · Yefeng Ma⁴

Received: 28 October 2015 / Accepted: 24 July 2016 / Published online: 29 July 2016
© Springer Science+Business Media Dordrecht 2016

Abstract Risk prediction and damage assessment play critical roles in disaster response to reduce losses. Social media can serve as crowdsourcing platforms for disaster information dissemination and data mining. Using typhoon Haiyan as an example, a close relationship between social media and disaster damage estimation is demonstrated, which provides a new perspective for disaster preparedness and response. Based on disaster-related social media data, a new index model is developed for situation awareness and damage assessment before, during, and after disasters. The difference between the new index model and traditional ones is that the new index is extracted from microblogs using semantic analysis method. The score of each index is determined by the emergency management experts. The weight is calculated based on TF-IDF method, a classical term frequency weight method. Based on the new index model, quantitative assessment is added to qualitative analysis. The assessment result is consistent with actual situation, which underlines the feasibility of implementation of the new model.

Keywords Damage assessment · Microblog data · Index model · Crowdsourcing · Typhoon response

✉ Hui Zhang
zhhui@tsinghua.edu.cn

¹ Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, China

² Public Order School, People's Public Security University of China, Beijing, China

³ Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing, China

⁴ School of Economics and Management, Tsinghua University, Beijing, China

1 Introduction

Social media, more specifically twitter and microblog, play more and more important roles in emergency management (Starbird and Palen 2012; Merchant et al. 2011; Pennington-Gray et al. 2013). Applying social media data to crisis management has become a growing trend since social media are regarded as knowledge management systems in 2010 Haitian earthquake (Couling 2014; Yates and Paquette 2011). Current studies mainly focus on four aspects. First, social media data were applied to early disaster detection and early warning (Sakaki et al. 2010; Abel et al. 2012; Abdelhaq et al. 2013). Each user shared what he/she observed and experienced through social media anytime and anywhere. Social media offered early warning information and sometimes reported breaking news more rapidly and timely than mainstream media (Guan and Chen 2014). Second, researches were looking into the mechanism of online disaster-related individual and group behaviors, such as user motivation, tweet- or retweet-behavior, and various factors on social media behaviors from the perspective of twitter and microblog contents (Correa et al. 2010; Stieglitz and Dang-Xuan 2013; Rudat et al. 2014; Morchid et al. 2014; Pennington-Gray et al. 2013). Third, researches tried to discover the mechanism of online social network and information propagation during disaster (Starbird et al. 2010; Bakshy et al. 2012). As an important information spreading platform, retweet attracted much attention, such as waiting time of a retweet from the perspective of the users, external context, and prediction whether a tweet will be retweeted based on its owner and content (Spiro et al. 2012; Starbird and Palen 2012; Zaman et al. 2010). Fourth, social media also supported emergency decision-making, and served as crowdsourcing platform for disaster relief mission (Gao et al. 2011a, b), emergency preparedness (Merchant et al. 2011), situation awareness (Yin et al. 2012), and disaster response planning through knowledge sharing (Yates and Paquette 2011). These studies show a close relationship between social media and disaster detection, users' behavior, information propagation, decision-making. However, with twitter data being widely used, Chinese microblog (e.g., Sina Weibo) has been paid relatively less attention in emergency management.

Sina Weibo, one of the most popular Chinese microblogs, is a Twitter-like Chinese microblog (Gao et al. 2012; Chen et al. 2011). Sina Weibo was launched by Sina in August 2009 while Twitter was established in March 2006. Each user can release and receive short messages less than 140 characters by PCs, mobile phones or other media anytime and anywhere. When one user follows other users, he/she can forward and comment on the microblog messages. When many users have heated discussion on a same topic, a hot topic can be established by using a hastag, which is marked by "#". The open API (an open Application Programming Interface) of Sina Weibo provides an opportunity to capture microblog data.

Sina Weibo has a strong influence on emergency management. Some researchers have started to enhance emergency response using Sina microblog data. Statistical analysis was conducted to explore some characteristics of hot social events based on Sina Weibo data (Guan et al. 2014). The evolving process of collective emotions was recognized using Sina Weibo data analysis (Xiong et al. 2013). A keyword-based semantic analysis was conducted on Sina microblog data to obtain public opinions of online collective behaviors and to support decision-making (Ma et al. 2014). Content, trend and information spreading were analyzed based on Sina microblog data for 2010 Yushu Earthquake (Qu et al. 2011). Qualitative methods are generally employed.

Crowdsourcing has emerged for many years. Crowdsourcing means distributing a query to several Twitter users to aggregate the results and exploit the wisdom-of-crowds effect (Demirbas et al. 2010; Doan et al. 2011). It includes three elements: crowd, initiator and process (Estellés-Arolas and González-Ladrón-de-Guevara 2012). Crowdsourcing is widely used because it is generally believed that the aggregated opinion of many non-experts from the local community with local knowledge will approach the quality of the opinion of an expert (Eskenazi et al. 2013). The wisdom of crowds also supports emergency preparedness and emergency response. Rapid development of social network offers an important platform for the expression of public wisdom and for data mining and information extraction during crisis. Twitter, Ushahidi, and Sahana are some widely used crowdsourcing applications in western countries (Chan 2014). These applications can provide accurate and timely information during crisis, which is the foundation of a disaster relief information system (Gao et al. 2011a, b).

The consequences of the disasters, such as, casualty, destruction of infrastructure systems, loss of properties, and chaos of daily life, challenges to disaster response of various levels of government agencies. Social media are very helpful for information dissemination, situation awareness, early warning, risk assessment, damage evaluation, command and control. During crisis, the affected people use Twitter, Facebook, Microblog to describe situation and keep in touch with others. Each user acts as a moving sensor. A large number of the affected people offer massive information; therefore, social media imply tremendous valuable information to support emergency response. In this study, a two-stage model is proposed to assess risk and damage of a disaster. Degree of risk and degree of damage are constructed as two composite scores to measure risk level before and after disaster. Composite measures can be simplified to two single measures based on microblog data since there is a close relationship between social media data and severity of impact. Degree of risk is used to describe all potential risks when a disaster is coming, while degree of damage describes its impact on local society, infrastructure systems, and mentality of the affected people. They are estimated by an index system. All indices in the index system are extracted from Chinese microblog data collected from Sina Weibo. From the spatial perspective, the degrees of risk and damage are distinguished at different regions. According to the degrees of risk and damage at different regions, risk-map and damage-map can be drawn to support disaster preparedness and response. Typhoon Haiyan (one of the strongest tropical cyclones ever recorded in the Philippines history) is selected as an example to provide an empirical illustration of the new model.

2 Methodology

2.1 Fuzzy comprehensive evaluation

Fuzzy comprehensive evaluation is a quantitative assessment model to solve a complex problem described by a series of indices. Each index is assessed a score and assigned a weight based on its contribution to the problem. It is widely used in the quantitative assessment of many complex systems, such as urban response ability (Tian and Yang 2008) and power system (Zhao et al. 2008). With rapid growth of online collectives, the model is also introduced to unconventional emergency. It focuses on the statistical analysis rather than content analysis on social media data (Lan 2011). This paper concentrates on the semantic analysis on microblog data to extract wisdom of crowds and support disaster

response. The framework of research method is shown in Fig. 1. A large number of users express their opinions on Sina Weibo. Based on semantic analysis on the collected data, risk and damage can be evaluated using the index model to support emergency decision-making.

2.2 Proposed evaluation model

Public opinion online can facilitate situation awareness and damage assessment. People, both disaster managers and general public express their opinions on social media before, during and after disaster. Government officials stay in the decision-making roles in disaster management. The messages from government officials are removed from the datasets. Messages posted by general public are treated as the crowdsourcing data to collect public opinion. Before disaster, public opinion expresses individuals’ risk perception through their descriptions on disaster, their mentality and activities. For example, if many people in a region express that they bought some emergency goods before typhoon, such as food, water, and radio, it indicates that people in the area are relatively well prepared. During or after disaster, people tend to express the current situation, future perspective, response measures and their mentality to reflect the damage level. For instance, the extent of damage may be high in a place where many people are calling for help or saying lack of food. Therefore, risk and damage can be assessed during crisis using microblog data.

Based on microblog data mining and semantic analysis, a new model is developed to predict risk before disaster and assess damage after disaster. The model is an index system based on microblog data. It is quite different from traditional ones because the indices here are the terms or statements extracted directly from microblog data through semantic analysis. Semantic analysis is a content analysis method to infer the meaning of a message

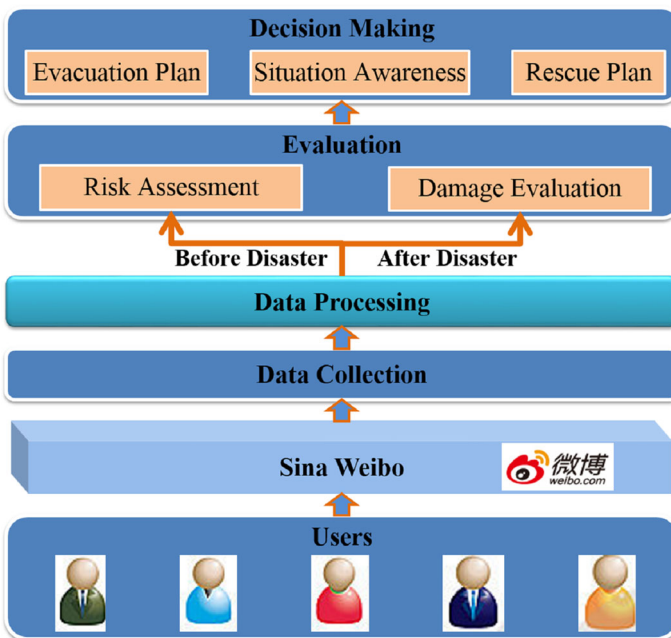


Fig. 1 Framework of crowdsourcing for disaster response

and mine the relation between messages (Ma et al. 2014). For an index system, four aspects are determined: (1) index selection, (2) score of index, (3) weight of index, and (4) evaluation function. These four aspects will be discussed one-by-one as follows.

2.2.1 Index selection

The messages on social media contain a large amount of information. The valuable information during crisis is divided into three categories: disaster description, public demands and mitigation measures. The mentioned information is described and reflected in the disaster-related messages to contribute to risk assessment and damage evaluation.

2.2.1.1 Disaster description Physical parameters are direct variables to measure the intensity of a disaster. They mainly include disaster intensity (e.g., magnitude and intensity of a disaster, affected people and scope), tendency of intensity, disaster-related location or time, secondary or derivative disasters and destruction of infrastructures. The severity of a disaster is described by general public to represent their risk perception for the disaster or actual damage degree.

2.2.1.2 Public demands The demands of the general public reflect their risk perception and damage level. The general public talk about their daily routines and what they are experienced on the Internet. Messages convey what they need through the Internet if they cannot obtain direct aids during disaster. The demands often include emergency supplies, mentality, activities and help from government. The demands indicate that their risk awareness and severity of situation surrounding them.

2.2.1.3 Mitigation measures Mitigation measures are actions to reduce risks and avoid severe damage during disaster. If the general public are well prepared, such as prepared adequate food, water and other supplies, the risk level may be reduced. Some official measures, such as reporting instant news, offering alert information, communicating with the general public, and seeking psychological counseling, can also mitigate risk. The measures and the participation of governments in disaster response are reflected by public opinion. For example, if many people complain, called for help and lived without food, government response is considered as ineffective, and risk and damage levels are high. People also express the aids from governments which mitigate the risk and potential damage. Therefore, mitigation measures include government effective response and public well preparedness.

The new model is developed based on three-category information, as shown in Fig. 2. Three categories are disaster description, public demands and mitigation measures. Each category is composed of several subcategories which are the first-grade indices. The first-grade indices of disaster description include intensity (I_1), tendency of intensity (I_2), key locations or time (I_3), secondary or derivative disasters (I_4) and infrastructure destruction (I_5). Four first-grade indices of public demands are supply demands (I_6), mentality (I_7), activities (I_8) and organizations (I_9). Two first-grade indices of mitigation measures are government measures (I_{10}) and public preparedness (I_{11}). There are total eleven first-grade indices in this model. The following analysis is based on these first-grade indices.

Each index contains many terms which are the second-grade indices, and each term represents a similar meaning. For instance, the infrastructure systems (I_5) include the

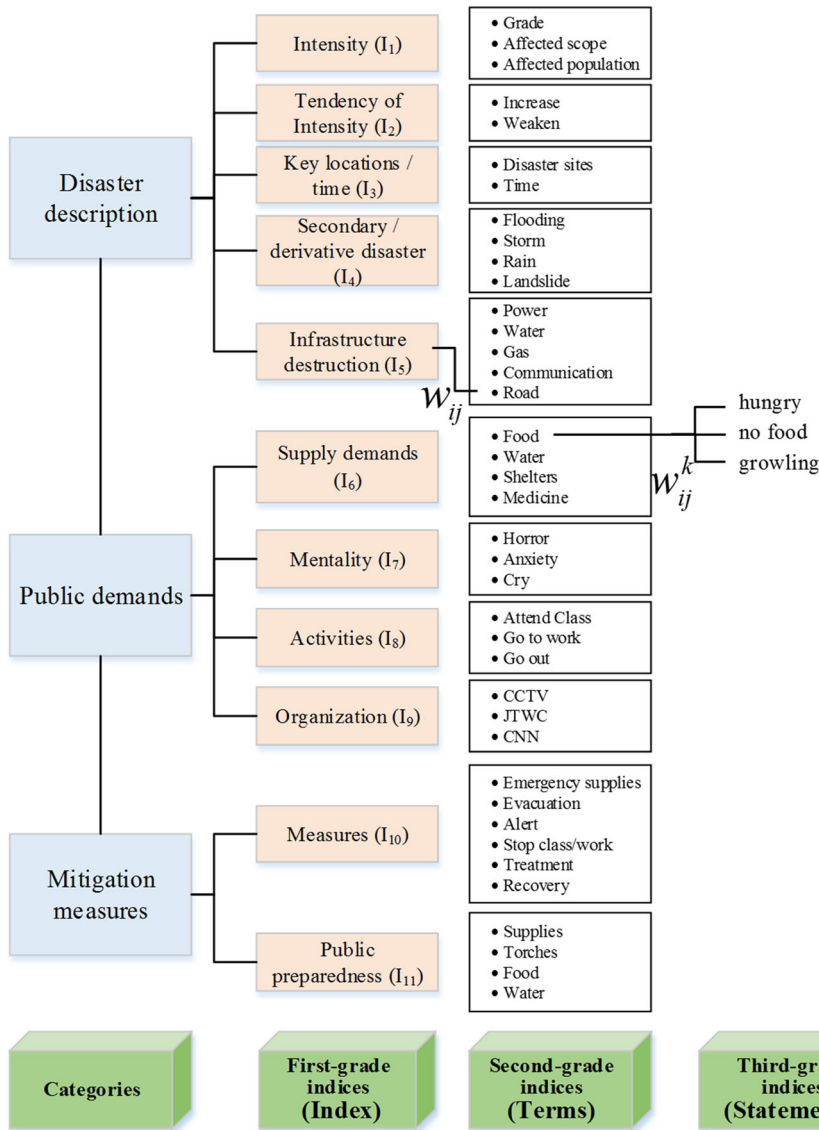


Fig. 2 Index model of risk prediction/damage assessment

networks of power, water, gas, communication and road, et al. The i th index can be obtained and shown in Eq. (1).

$$\text{index}_i = \{ \text{term}_{i1}, \text{term}_{i2}, \dots, \text{term}_{ij}, \dots \} \quad (i = 1, 2, \dots, 11; j = 1, 2, \dots, n) \quad (1)$$

where index_i is the i th index and term_{ij} represents all potential terms contained in index_i .

There is an important feature of Chinese language, which is synonym. It means that several words express one meaning. Then, a certain term is just a collection of different

statements with similar meaning. For example, the statements “饿 (hungry)”, “没有食物 (no food)” and “肚子叫 (growling)” describe the meaning “缺乏食物 (lack of food)”. In this study, several statements with a high similarity are counted as a same term. The j th term in index, can be obtained and shown in Eq. (2).

$$\text{term}_{ij} = \left\{ \text{statement}_{ij}^1, \text{statement}_{ij}^2, \dots, \text{statement}_{ij}^k, \dots \right\} \quad (2)$$

$(i = 1, 2, \dots, 11; j = 1, 2, \dots, n; k = 1, 2, \dots, m)$

where statement_{ij}^k ($k = 1, 2, \dots, m$) represents all statements that occur and have a similar meaning with term_{ij} .

2.2.2 Score

Each statement (third-grade index) is estimated a score based on its contributions to the degree of risk or damage. Different statements reflect different degrees of risk or damage, which is the standard for measuring scores. When a certain emotional statement is referred in many messages, it indicates that many people have resonated. Once the emotion is pessimistic, some irrational behaviors are easy to be triggered and risk may be expanded. The score of this statement is therefore relatively high. Different statements have different contributions to risk and damage levels, and they will be assigned for different scores.

The expert evaluating method is used to estimate the score of each statement. The score of a statement is determined by the disaster experts based on its contribution to disaster risk or damage level. The score range is from -5 to 5. When a statement has opposite contribution to the risk or damage level, its score is assigned to a negative value. For example, the statement “power recovery” decreases the risk or damage level, so its score is a negative value. The larger the absolute value of the score is, the bigger its contribution to the degree of risk or damage is.

2.2.3 Weight

There are two kinds of weight. One is the weight between first-grade index and second-grade (term), which is denoted by w_{ij} . The other is the weight between first-grade index and second-grade (statement), which is denoted by w_{ij}^k .

Based on term frequency weight method (Salton and Buckley 1988), the weight of an index can be calculated based on its frequency to a certain degree. For instance, a statement representing a necessity is repeated for many times, which shows the necessity is required by a large scale of people. It causes a relatively higher risk and more severe damage for lack of emergency supplies and poor preparation. Absolute frequency and relative frequency are widely used in Chinese natural language processing. However, absolute frequency is not appropriate since less frequent statements from crowd are not necessarily uncritical and more frequent statements are not always significant. Relative frequency is an alternative method. Term frequency-inverse document frequency (TF-IDF), as a widely used method of relative frequency calculation, characterizes the text better than term frequency (TF) method and is more appropriate to actual situation (Forman 2008; Xie et al. 2015).

TF-IDF method is a typical method of calculating the values of important terms in a document through an inverse proportion of the frequency of the term in a particular document to the proportion of that term over the entire document corpus (Chen and Xiao

2016; Ramos 2003). Equation (3) is the classical formula of TF-IDF used for term weighting.

$$w_{ij} = tf_{ij} \times \log\left(\frac{N}{df_i}\right) \tag{3}$$

where w_{ij} is the weight for term i in document j , tf_{ij} is the term frequency of term i in document j , N is the total amount of documents in the document corpus and df_i is the number of document containing term i in the document corpus.

In this study, a microblog message is a document. Term i is not always contained in only one document (e.g., document j). For a document corpus, document j and other documents may contain term i . A weight shall be calculated for the term in each document containing it. The weight for the term in the document corpus is the average of the weights for the term in all documents containing it. The formula of the weight for a term in document corpus is calculated as Eq. (4).

$$w_i = \frac{1}{df_i} \times \left(\sum_{j \in D_i} tf_{ij}\right) \times \log\left(\frac{N}{df_i}\right) \tag{4}$$

where w_i is the weight for term i in document corpus and D_i is the collection of documents containing term i .

Terms common in a single or a small group of documents tend to have higher TF-IDF values than common terms. For instance, the term “typhoon”, common in many messages in typhoon situation, will not be given high weight. In addition, some less frequent but important terms are not ignored because of their low df_i values. TF-IDF can calculate the weight of a term which is appropriate to actual situation. Therefore, the method of TF-IDF is employed to assign the weight of a statement.

The value of w_{ij} is simply set as unity in the case study for the limitation of data. It means that the importance of all terms (second-grade indices) is considered similar. Several cases have been used to test the method. The results have a relatively high quality and are more appropriate to actual situation. This method is applied to the case study.

2.2.4 Evaluation function

The degree of risk or damage is the total value of all indices. The degree of statement ^{k} _{ij} is the result of multiplying its score by its weight. The degrees of term _{ij} , index _{i} are defined in Eq. (5) and Eq. (6), respectively.

$$DTerm_{ij} = \sum_{k=1}^m w_{ij}^k S_{ij}^k \quad (i = 1, 2, \dots, 11; j = 1, 2, \dots, n; k = 1, 2, \dots, m) \tag{5}$$

$$DIndex_i = \sum_{j=1}^n w_{ij} DTerm_{ij} \quad (i = 1, 2, \dots, 11; j = 1, 2, \dots, n) \tag{6}$$

where $DTerm_{ij}$ is the degree of term _{ij} , S_{ij}^k is the score of statement ^{k} _{ij} , w_{ij}^k is the weight of statement ^{k} _{ij} , $DIndex_i$ is the degree of index _{i} and w_{ij} is the weight of term _{ij} .

The degree of risk or damage is the sum of the degrees of all indices, and it is defined in Eq. (7).

$$DRisk/DDamage = \sum_{i=1}^{11} DIndex_i \quad (i = 1, 2, \dots, 11) \tag{7}$$

where DRisk is the degree of risk and DDamage is the degree of damage.

In conclusion, the degree of risk or damage is defined in Eq. (8).

$$DRisk/DDamage = \sum_{i=1}^{11} \sum_{j=1}^n \left(w_{ij} \sum_{k=1}^m w_{ij}^k S_{ij}^k \right) \tag{8}$$

(i = 1, 2, ..., 11; j = 1, 2, ..., n; k = 1, 2, ..., m)

3 Case study

Recently, natural disasters happen frequently and cause huge losses for local society. On November 7, 2013, Typhoon Haiyan, one of the strongest tropical cyclones ever recorded, landed in Philippines at Guiuan (Eastern Samar). The storm’s maximum sustained wind was at 275 km/h, and the affected population was up to 9.6 million; subsequently, it grazed in Hainan at 16:00 on November 10 with the wind speed of 121 km/h; it landed in Guangxi at 5:00 on November 11 with the speed of 84 km/h. After it made landfall in China, its intensity was weakened. Its track was shown in Fig. 3. Without early situation awareness to support evacuation before it landed in China and effective damage assessment to help rescue after it made landfall, local society suffered heavy losses. Its early risk prediction and effective damage assessment caught wide attention. For this goal, a new model with statement frequency analysis of the most common statements and comparative analysis is proposed to evaluate the degrees of risk and damage based on the disaster-related microblog data.

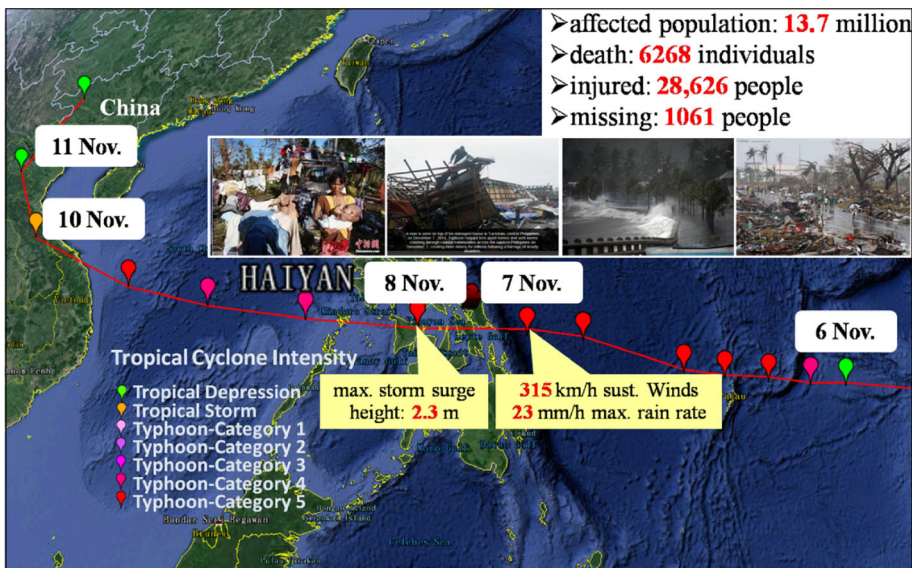


Fig. 3 Path of typhoon Haiyan

3.1 Data collection

Large quantities of microblog messages are collected for the life-cycle response. Sina Weibo is selected as the source to collect disaster-related microblog data. The keyword-based method is applied to capture microblog data by using Weibo API and some web page crawling algorithms. The terms “Haiyan”, “typhoon” and “typhoon Haiyan” are the keywords to collect microblog posts and the period of data collection is from 3 November, 2013 to 15 November, 2013. 3098 pieces of keyword-related microblog messages are collected. The detailed information of the collected data includes author’s name, ID, user’s region, content, post time and URL. However, some may be noise data, such as data keyword-related but not disaster-related or data posted by governments. In order to ensure the validity of the data as far as possible, manual sorting method is also used to filter noise data. All collected data are classified into related and unrelated categories manually. Then each piece of unrelated microblog posts is double-checked in case of wrong selection. Finally, 2539 pieces of related messages are remained with geographic location of the author and post time attached for the following analysis.

Hainan and Guangxi provinces are considered as two of the worst-hit provinces in China. When typhoon Haiyan landed in Philippines and is on its way to China, the number of related microblogs increases greatly in China, as shown in Fig. 4. It grazes in Hainan at 16:00 on November 10 and lands in Guangxi at 5:00 on November 11. Temporally, within the 24 h before and after its landfall, the microblog data are collected to help risk prediction and damage assessment, respectively. For Hainan, the period of data collection for risk prediction is from 16:00 on November 9 to 16:00 on November 10, and it for damage assessment is from 16:00 on November 10 to 16:00 on November 11. Corresponding to Hainan, the period of data collection for risk prediction is from 5:00 on November 10 to 5:00 on November 11, and for damage assessment, it is from 5:00 on November 11 to 5:00 on November 12 in Guangxi.

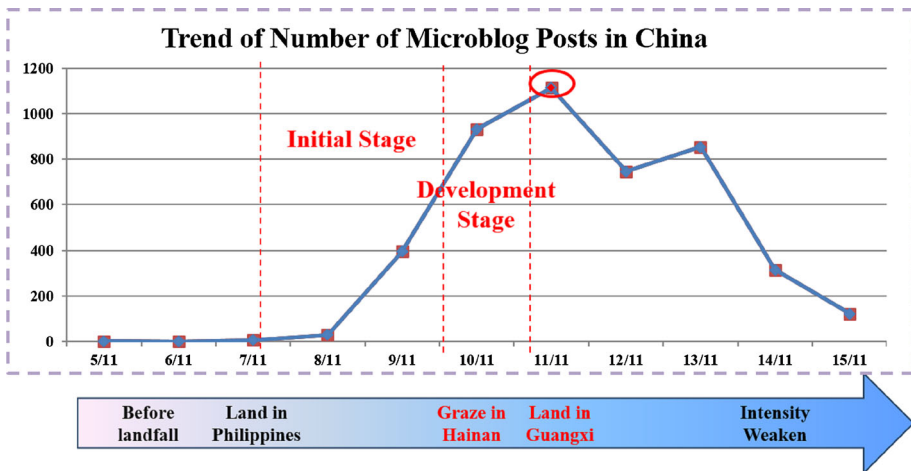


Fig. 4 Trend of number of microblog posts in China in typhoon Haiyan

3.2 Data pre-processing

After collecting microblog messages from Sina Weibo, the collected messages are processed and analyzed, including classification, segmentation, stopwords removing and term frequency analysis.

The collected microblog messages are classified based on the geographic location of the users. A new user registers geographic information in Sina Microblog and it appears in the user’s homepage. Most users register their provinces but not offer detailed address information. Therefore, province is treated as the smallest study area in this study. Microblog messages posted by Hainan and Guangxi citizens are analyzed. The messages posted by Hainan and Guangxi citizens for four specific periods are picked out. Based on the post time, the selected messages are divided into four datasets: (1) before grazing in Hainan Province (16:00 on November 9 to 16:00 on November 10); (2) before landing in Guangxi Province (5:00 on November 10 to 5:00 on November 11); (3) after grazing in Hainan Province (5:00 on November 10 to 5:00 on November 11); (4) after landing in Guangxi Province (5:00 on November 11 to 5:00 on November 12).

Segmentation means dividing a complete sentence into a series of terms or phrases based on some Chinese language rules. In these terms and phrases, some words (called stopwords, e.g., a, the, an) have small significance for the expression of a sentence while their frequency is very high, so this kind of words are removed. Relative frequency analysis is conducted to count the frequency of a term and obtain its relative frequency. The terms are ranked based on their frequency, and the most common terms (e.g., in the top 100 in each dataset) are selected to evaluate the degree of risk or damage.

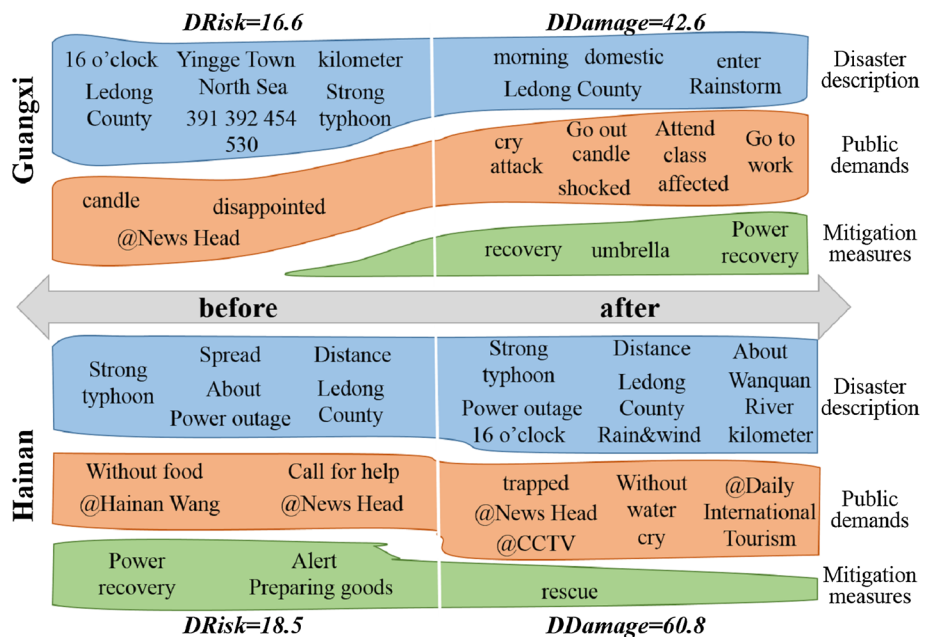


Fig. 5 The most common terms and results of risk and damage degrees. The terms “cry”, “candle”, “shocked” and “disappointed” are four emotions translated into their meanings

4 Results

4.1 Semantic analysis

Semantic analysis helps governments know more about what citizens concerned in different areas and at different stages. The concerns of citizens are inconsistent at different areas. Knowing what they need, some targeted measures can be executed to them. Figure 5 provides some of the most common terms in four datasets, and they are before grazing in Hainan (Dataset 1), before landing in Guangxi (Dataset 2), after grazing in Hainan (Dataset 3) and after landing in Guangxi (Dataset 4). Some terms, such as the name of typhoon or places (e.g., Haiyan, Hainan or Gaungxi), are frequently occurred in almost all messages. Their contributions to the degree of risk or damage are negligible, so they are removed when Fig. 5 is drawn.

4.1.1 Semantic analysis before typhoon landing

Before typhoon Haiyan lands in Hainan, the terms “power outage”, “preparing goods”, “without water”, “hungry”, “call for help” and “alert”, “distance” are frequent in the microblogs posted by Hainan citizens. It indicates that Hainan citizens focus on preparation, necessities, infrastructure systems and alert information. However, Guangxi citizens pays more attention on the key places or time (e.g., “Yingge Town”, “16 o’clock”), wind speed (e.g., “391”, “530”) and their mentality (e.g., “candle”, “disappointed”) before it lands in Guangxi. In addition, with the high frequency of “@News Head”, citizens in both provinces are calling for help from the official organizations. It indicates the important role of official social media in disaster response.

4.1.2 Semantic analysis after typhoon landing

After Typhoon Haiyan grazes in Hainan, the terms “power outage”, “Ledong County”, “Wanquan River”, “rescue” and “without water” are frequent in the messages posted by Hainan citizens, which indicates that the attention of Hainan citizens is paid to the damage of infrastructure systems (e.g., power, water, road, etc.), necessities (e.g., food, water, etc.), affected area (e.g., Wanquan River) and rescue mission, as shown in Fig. 5. However, citizens in Guangxi concern the storm caused by the typhoon, recovery after the typhoon landed and desire to go work or attend class, because the words “rain storm”, “umbrella”, “attend class”, “go to work” and “power recovery” have high frequency. At the same time, the mentality of Guangxi citizens is relatively unstable, because the emotions “cry”, “shocked” and “candle” for praying are frequent in their messages. Based on further comparative analysis, it is clear that Hainan citizens pay more attention to the basic needs while citizens in Guangxi on the storm, recovery and desire for daily activities. The official social media, such as News Head, CCTV, are popular for citizens to call for help after disaster.

4.1.3 Comparison of semantic analysis between before- and after-landing

Before- and after-landing in Hainan, the concerns of citizens are on different aspects. For disaster description, citizens focus on real-time situation and the tendency before it lands in Hainan, while they pay more attention on the affected areas after it makes landfall. For

mitigation measures, preparing measures and alert are important before it lands in Hainan while rescue mission is the most important after it makes landfall.

There is difference of citizens’ concerns between before- and after-landing in Guangxi. Before it makes landfall, citizens pay more attention on disaster description, such as typhoon speed, key time or places, while public demands (e.g., go out, go to work) and mitigation measures (e.g., power recovery) are more important after it lands in Guangxi.

4.2 Quantitative analysis

The paper provides a simple but complete demonstration of the system by using only one index in typhoon Haiyan to control the length of the paper and avoid repetition. The detailed illustration of the index model is shown in Table 1. The index-supply demands (I_6) —is chosen as an example to give the detailed information. The information includes the most common terms belonging to the index in the study period, the statements in each term, the score and weight of each statement, and the DTerm corresponding to each term.

The degree of the index-supply demands (I_6)—is estimated in Eq. (9), and the result is 3.50.

$$DIndex_6 = \sum_{j=1}^n w_{ij} DTerm_{ij} = 3.50 \tag{9}$$

In the case study, the value of w_{ij} is simply set as unity for the limitation of collected data.

4.2.1 Risk prediction before typhoon landing

When typhoon Haiyan is coming, risk analysis is conducted. Based on the index model, the risk degrees at different areas are estimated. Before it grazes in Hainan, the risk degree of Hainan is 18.5, while it is 16.6 before it lands in Guangxi, as the results shown in Fig. 5. It shows the risk of Hainan is higher than that of Guangxi. The reason may be when it enters Guangxi, its intensity is weaker (The wind speed decreases from 121 km/h when Typhoon Haiyan grazes in Hainan to 84 km/h when it lands in Guangxi), so public risk perception is decreased. When governments take measures, more attention is paid to the areas with high risk.

Table 1 An example of the index model. The data are from semantic analysis before Typhoon Haiyan grazing in Hainan (Dataset 1)

Category	Index	Term	Statement			
			Statements	Score	Weight	DTerm
Public demands	Supply demands (I_6)	Without water	□渴(thirsty)	5	0.155	0.775
			水(water)	4	0.187	0.748
		Without food	(肚子)饿(hungry)	5	0.196	0.980
			食物(food)	4	0.144	0.576
			吃的(something to eat)	3	0.139	0.417

4.2.2 Damage assessment after typhoon landing

After typhoon Haiyan lands in an area, it results in a serious damage to local society and infrastructure systems. All information can be reflected through social media data to help estimate the degree of damage. After it grazes in Hainan, the degree of damage is 60.8, while the degree of damage is 42.6 after it lands in Guangxi, as shown in Fig. 5. Though typhoon Haiyan has landed in Guangxi while just grazes in Hainan, damage in Hainan is much more severe than that in Guangxi. Many measures, such as power recovery, alert and rescue, are taken into action by Hainan governments before and after Typhoon Haiyan grazing in Hainan, which indicates that the risk and damage levels in Hainan are relatively high. According to the official statistics from GDACS, the affected population is 1.1 million when Typhoon Haiyan grazes in Hainan while there are no people affected when it lands in Guangxi. So the assessment results based on the index model are consistent with the losses in actual situation. The index model using disaster-related social media data supports emergency decision-making.

4.2.3 Comparison of quantitative analysis between before- and after-landing

Before typhoon landing, the ratio of risk degrees in these two provinces is 1.12, while the ratio of damage degrees is 1.43 after typhoon landing. It means that the difference between Hainan and Guangxi becomes bigger during the evolution of typhoon Haiyan. Before typhoon landing, citizens cannot accurately know where the typhoon may land, so the difference of two provinces is relatively small. When Haiyan makes landfall, citizens have a knowledge of the situation to a certain degree. Because the damage of Hainan is more severe than that of Guangxi, the difference of damage degrees becomes bigger.

5 Conclusions

Risk prediction is crucial for the evacuation plan, and damage assessment is required for rescue plan in disaster response. This paper proposes a new method for risk prediction and damage assessment. Based on social media data, what people concerned can be obtained from their messages. An index system is developed from three perspectives: disaster description, public demands and mitigation measures. The difference between the new model and traditional ones is that the new index is extracted directly from the disaster-related microblogs. The score of each index is determined by emergency management experts. The weight is calculated based on the TF-IDF method, which is a classical term frequency weight method. With the evaluation function, the degree of risk or damage can be estimated. By this way, once the microblog data are collected, the degree of risk and damage can be evaluated quickly to provide rapid risk prediction and effective damage assessment.

Using typhoon Haiyan as an example, this paper demonstrates the use of the new model. As two worst-hit provinces in China, risk prediction and damage assessment are conducted in Hainan and Guangxi provinces based on social media data. Before typhoon makes landfall, the degree of risk in Hainan is higher than that in Guangxi. After it makes landfall, the degree of damage in Hainan is also higher than that in Guangxi. Through comparison with actual situation, the assessment result based on the index model is consistent with actual situation, which underlines the feasibility of implementation of the new model.

Acknowledgments This work was partially supported by the National Natural Science Foundation of China (Grant No. 91224008). We thank the Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia at Beijing University of Posts and Telecommunications for providing the microblog data support. We also thank the reviewers and editors for their careful review and helpful suggestions.

References

- Abdelhaq H, Sengstock C, Gertz M (2013) Eventweet: online localized event detection from twitter. *Proc VLDB Endow* 6(12):1326–1329. doi:[10.14778/2536274.2536307](https://doi.org/10.14778/2536274.2536307)
- Abel F, Hauff C, Houben GJ et al (2012) Twitcident: fighting fire with information from social web streams. *Proceedings of the 21st international conference companion on World Wide Web. ACM*, pp 305–308. doi:[10.1145/2187980.2188035](https://doi.org/10.1145/2187980.2188035)
- Bakshy E, Rosenn I, Marlow C et al (2012) The role of social networks in information diffusion. *Proceedings of the 21st international conference on World Wide Web. ACM*, pp 519–528. doi:[10.1145/2187836.2187907](https://doi.org/10.1145/2187836.2187907)
- Chan CJ (2014) Crowdsourcing Disaster Response. *ResearchGate*
- Chen G, Xiao L (2016) Selecting publication keywords for domain analysis in bibliometrics: a comparison of three methods. *J Informetr* 10(1):212–223. doi:[10.1016/j.joi.2016.01.006](https://doi.org/10.1016/j.joi.2016.01.006)
- Chen S, Zhang H, Lin M et al (2011) Comparison of microblogging service between Sina Weibo and Twitter. *Computer science and network technology (ICCSNT), 2011 international conference on. IEEE*, vol 4, pp 2259–2263. doi:[10.1109/ICCSNT.2011.6182424](https://doi.org/10.1109/ICCSNT.2011.6182424)
- Correa T, Hinsley AW, De Zuniga HG (2010) Who interacts on the Web? The intersection of users' personality and social media use. *Comput Hum Behav* 26(2):247–253. doi:[10.1016/j.chb.2009.09.003](https://doi.org/10.1016/j.chb.2009.09.003)
- Couling M (2014) Tsunami risk perception and preparedness on the east coast of New Zealand during the 2009 Samoan Tsunami warning. *Nat Hazards* 71(1):973–986. doi:[10.1007/s11069-013-0945-y](https://doi.org/10.1007/s11069-013-0945-y)
- Demirbas M, Bayir MA, Akcora CG et al (2010) Crowd-sourced sensing and collaboration using twitter. *World of wireless mobile and multimedia networks (WoWMoM), 2010 IEEE international symposium on a. IEEE*, pp 1–9. doi:[10.1109/WOWMOM.2010.5534910](https://doi.org/10.1109/WOWMOM.2010.5534910)
- Doan A, Ramakrishnan R, Halevy AY (2011) Crowdsourcing systems on the world-wide web. *Commun ACM* 54(4):86–96. doi:[10.1145/1924421.1924442](https://doi.org/10.1145/1924421.1924442)
- Eskenazi M, Levow GA, Meng H et al (2013) Crowdsourcing for speech processing: applications to data collection, transcription and assessment. *Wiley*
- Estellés-Arolas E, González-Ladrón-de-Guevara F (2012) Towards an integrated crowdsourcing definition. *J Inf Sci* 38(2):189–200. doi:[10.1177/0165551512437638](https://doi.org/10.1177/0165551512437638)
- Forman G (2008) BNS feature scaling: an improved representation over tf-idf for svm text classification. *Proceedings of the 17th ACM conference on Information and knowledge management. ACM*, pp 263–270. doi:[10.1145/1458082.1458119](https://doi.org/10.1145/1458082.1458119)
- Gao H, Barbier G, Goolsby R et al (2011a) Harnessing the crowdsourcing power of social media for disaster relief. *Arizona State University, Tempe*
- Gao H, Wang X, Barbier G et al (2011) Promoting coordination for disaster relief—from crowdsourcing to coordination. *Social computing, behavioral-cultural modeling and prediction. Springer Berlin Heidelberg*, pp 197–204. doi:[10.1007/978-3-642-19656-0_29](https://doi.org/10.1007/978-3-642-19656-0_29)
- Gao Q, Abel F, Houben G J et al (2012) A comparative study of users' microblogging behavior on Sina Weibo and Twitter. *User modeling, adaptation, and personalization. Springer Berlin Heidelberg*, pp 88–101. doi:[10.1007/978-3-642-31454-4_8](https://doi.org/10.1007/978-3-642-31454-4_8)
- Guan XY, Chen C (2014) Using social media data to understand and assess disasters. *Nat Hazards* 74:837–850. doi:[10.1007/s11069-014-1217-1](https://doi.org/10.1007/s11069-014-1217-1)
- Guan W, Gao H, Yang M et al (2014) Analyzing user behavior of the micro-blogging website Sina Weibo during hot social events. *Physica A* 395:340–351. doi:[10.1016/j.physa.2013.09.059](https://doi.org/10.1016/j.physa.2013.09.059)
- Lan YX (2011) On Construction of Emergency Network Safety Evaluation Index System. *J Intell* 7:019 (in Chinese)
- Ma Y, Deng Q, Wang X et al (2014) Keyword-based semantic analysis of microblog for public opinion study in online collective behaviors. *Web-age information management. Springer International Publishing*, pp 44–55. doi:[10.1007/978-3-319-11538-2_5](https://doi.org/10.1007/978-3-319-11538-2_5)
- Merchant RM, Elmer S, Lurie N (2011) Integrating social media into emergency-preparedness efforts. *N Engl J Med* 365(4):289–291. doi:[10.1056/NEJMp1103591](https://doi.org/10.1056/NEJMp1103591)
- Morchild M, Dufour R, Bousquet PM et al (2014) Feature selection using principal component analysis for massive retweet detection. *Pattern Recogn Lett* 49:33–39. doi:[10.1016/j.patrec.2014.05.020](https://doi.org/10.1016/j.patrec.2014.05.020)

- Pennington-Gray L, Kaplanidou K, Schroeder A (2013) Drivers of social media use among African Americans in the event of a crisis. *Nat Hazards* 66(1):77–95. doi:[10.1007/s11069-012-0101-0](https://doi.org/10.1007/s11069-012-0101-0)
- Qu Y, Huang C, Zhang P et al (2011) Microblogging after a major disaster in China: a case study of the 2010 Yushu earthquake. *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. ACM, pp 25–34. doi:[10.1145/1958824.1958830](https://doi.org/10.1145/1958824.1958830)
- Ramos J (2003) Using tf-idf to determine word relevance in document queries. *Proceedings of the first instructional conference on machine learning*
- Rudat A, Buder J, Hesse FW (2014) Audience design in Twitter: retweeting behavior between informational value and followers' interests. *Comput Hum Behav* 35:132–139. doi:[10.1016/j.chb.2014.03.006](https://doi.org/10.1016/j.chb.2014.03.006)
- Sakaki T, Okazaki M, Matsuo Y (2010) Earthquake shakes Twitter users: real-time event detection by social sensors. *Proceedings of the 19th international conference on World Wide Web*. ACM, pp 851–860. doi:[10.1145/1772690.1772777](https://doi.org/10.1145/1772690.1772777)
- Salton G, Buckley C (1988) Term-weighting approaches in automatic text retrieval. *Inf Process Manage* 24(5):513–523. doi:[10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
- Spiro E, Irvine C, DuBois C, Butts C (2012) Waiting for a retweet: modeling waiting times in information propagation. In: 2012 NIPS workshop of social networks and social media conference. Accessed 12 Dec 2012
- Starbird K, Palen L (2012) Pass it on? Retweeting in mass emergencies. In: *Proceedings of conference on information systems on crisis response and management (ISCRAM 2010)*. Accessed Dec 12 2012
- Starbird K, Palen L, Hughes AL et al (2010) Chatter on the red: what hazards threat reveals about the social life of microblogged information. *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. ACM, pp 241–250. doi:[10.1145/1718918.1718965](https://doi.org/10.1145/1718918.1718965)
- Stieglitz S, Dang-Xuan L (2013) Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *J Manage Inf Syst* 29(4):217–248. doi:[10.2753/MIS0742-1222290408](https://doi.org/10.2753/MIS0742-1222290408)
- Tian YL, Yang Q (2008) Urban Disaster Response Capacity Evaluation Index system model design based on AHP-DELPHI method. *WUHAN LIGONG DAXUE XUEBAO JIAOTONG KEXUE YU GIONG-CHENG BAN* 32(1):168 (in Chinese)
- Xie X, Du X, Yang Q et al (2015) A study on process model of computing similarity between product features and online reviews. *Big data technology and applications*. Springer, Singapore, pp 298–308. doi:[10.1007/978-981-10-0457-5_27](https://doi.org/10.1007/978-981-10-0457-5_27)
- Xiong XB, Zhou G, Huang YZ et al (2013) Dynamic evolution of collective emotions in social networks: a case study of Sina weibo. *Sci China Inf Sci* 56(7):1–18. doi:[10.1007/s11432-013-4892-8](https://doi.org/10.1007/s11432-013-4892-8)
- Yates D, Paquette S (2011) Emergency knowledge management and social media technologies: a case study of the 2010 Haitian earthquake. *Int J Inf Manage* 31:6–13. doi:[10.1016/j.ijinfomgt.2010.10.001](https://doi.org/10.1016/j.ijinfomgt.2010.10.001)
- Yin J, Lampert A, Cameron M et al (2012) Using social media to enhance emergency situation awareness. *IEEE Intell Syst* 27(6):52–59
- Zaman TR, Herbrich R, Van Gael J et al (2010) Predicting information spreading in twitter. *Workshop on computational social science and the wisdom of crowds, nips*. vol 104, no 45, pp 17599–601
- Zhao WW, Zhang JH, Shang JF et al (2008) Power system large scale blackout emergency evaluation index system and its application. *Autom Electr Power Syst* 32(20):27–30 (in Chinese)