Chapter 7
The Fear of Ebola: A Tale of Two Cities in China

Xinyue Ye, Shengwen Li, Xining Yang, Jay Lee, and Ling Wu

Abstract Emerging social issues have often led to rumors breeding and propagation in social media in China. Public health-related rumors will harm social stability, and such noise negatively affects the quality of disease outbreak detection and prediction. In this chapter, we use the diffusion of Ebola rumors in social media networks as a case study. The topic of rumors is identified based on latent Dirichlet allocation method, and the diffusion process is explored using the space-time methods. By comparing Ebola rumors in the two cities, the chapter explores the relationship between the spread of rumors, user factors, and contents. The results show that: (1) rumors have a self-verification process; (2) rumors have strong aggregation characteristics, and similar rumors in different regions at the same period of time will lead to a synergistic effect; (3) non-authenticated users are more inclined to believe the rumors, while the official users play a major role in stopping rumors as they pay more attention to the fact; (4) the spread and elimination of rumors largely depend on the users who have more followers and friends; and (5) the topics of rumors are closely related to the local event.

Keywords Ebola • Rumor • LDA • Social media • China

X. Ye
Department of Geography, Kent Sate University, Kent, OH, USA
e-mail: xye5@kent.edu

S. Li
Department of Information Engineering, China University of Geosciences, Wuhan, China
e-mail: lsw4000@gmail.com

X. Yang (✉)
Department of Geography and Geology, Eastern Michigan University, Ypsilanti, MI, USA
e-mail: westoxgeorge@gmail.com

J. Lee
College of Environment and Planning, Henan University, Kaifeng, Henan, China

Department of Geography, Kent Sate University, Kent, OH, USA
e-mail: jlee@kent.edu

L. Wu (✉)
Department of Sociology, Kent State University, Kent, OH, USA
e-mail: lwu18@kent.edu

Abbreviations

ESDA   Exploratory spatial data analysis
LDA   Latent Dirichlet allocation

7.1 Introduction

In August 2014, the worst Ebola outbreak on record conceded by the World Health Organization had become an international public health emergent event. Communication channels during times of emerging social crises and natural disasters play an important role before, during, and after such events, and social media have become important channels for communications with the advancement of Internet communication technology. However, emerging social issues have often led to rumor breeding and propagation in social media in China. Public health-related rumors will harm social stability, and such noise negatively affects the quality of disease outbreak detection and prediction based on social media. To study the diffusion of social media rumor towards such social events, in this chapter we present a case study of the diffusion of Ebola rumors in social media networks in China along with the outbreak period.

This chapter contributes to the literatures as an empirical study to apply social media and volunteered geographic data to understand human behavior and social phenomena. Recently a new source of geographic information has become available through user-generated efforts, enabled by technologies loosely regarded as Web 2.0 (Goodchild 2007). Geographic information can now be found in the contents of wikis, blogs, tweets, photos, and many other forms of user-generated content. Goodchild defines this new source of data as volunteered geographic information (VGI), which emphasizes the role of voluntary efforts in producing geographic information, to differentiate with the conventional geographic information production process. Sui (2008) further pointed out that VGI has transformed the conventional way in how geographic information is created and used so it represents a "wikification" of GIS, which is not simply confined to new ways of data productions but also includes the development of GIS software, hardware, organizations, and people. Already we have witnessed the popularity of VGI has provided many research opportunities for geographers and GIScientists in the digital age. Many studies found out VGI to be a valuable source because of its huge potential to engage citizens and to be a significant, timely, and cost-effective source for geographers' understanding of the earth (Yang et al. 2016). Our study resonates some other recent successful cases using VGI to help disaster relief, such as the provision of useful location-based information to the victims after the 7.0 earthquake in Haiti in 2010 and the Great East Japan Earthquake and Tsunami in 2011 (Zook et al. 2010; Yap et al. 2012).

7.2 Literature Review

7.2.1 Social Media and Public Health

Social media produce large volumes of real-time information every day. Messages via social media have supported research and analysis in a wide variety of fields, such as crisis analysis (MacEachren et al. 2011), detecting opinion events (Maynard et al. 2014), assessment of disaster damages (Sakaki et al. 2013), strategic environmental assessment (Floris and Zoppi 2015), and air pollution (Kay et al. 2014), among others. Scholars have also explored how social media information conveyed issues in public health (Hay et al. 2013). For example, Windener (Widener and Li 2014) used sentiment analysis to explore the prevalence of healthy and unhealthy food in the USA. King et al. (2013) analyzed the role of social media in informing, debating and influencing opinions in a specific area of health policy.

People began to realize that we have an extremely poor knowledge of how the vast majority of infectious diseases spread across spatial scales (Hay et al. 2013). The process of understanding starts with records of disease occurrences estimated from the literature (Rogers et al. 2006), reports of disease distribution on the web (Brownstein et al. 2008), and the trends suggested by GenBank (Benson et al. 2012). Together, these could be used to define a definitive extent of the disease (Brady et al. 2012) and to populate a database of occurrence locations (i.e., points) where the disease had been reported. For example, Messina et al. (2014) manually built a comprehensive global database of confirmed human dengue infection in March 2014, consisting of 8309 geo-positioned occurrence. However, it is labor-intensive to keep track of the current trend of infectious diseases and distribution. It should be noted that social media messages are self-reported volunteered information, unlike those from official or government outlets with delayed announcements. Tweets can be used to track and predict the emergence and spread of an epidemic (Achrekar et al. 2011; Velardi et al. 2014). The spatial pattern of flu risk was dynamically mapped across various spatial and temporal scales by FluMapper (Padmanabhan et al. 2014).

7.2.2 Rumor Theory and Social Media

Unlike lies, rumors are invalidated news that may include speculation and uncertainty of certain events (Buckner 1965). Rumors can be thought as a message or a communication between people about certain events or certain issues of concern that may or may not correspond to the truth (Mullen 1972). Rumors not only communicate messages but also emerge a form of interactions among people in a group of individuals when they face the threat of a crisis (Indrawan-Santiago et al. 2014; Peterson and Gist 1951; Rosnow 1988). Social media platforms, such as Twitter and microblog, sometimes spread wrong information or rumor because they are

citizen-based non-professional media. However, some information with reliable sources can minimize rumor propagation, suppressing the level of anxiety in the virtual community (Indrawan-Santiago et al. 2014; Oh et al. 2010; Okada et al. 2014; Spiro et al. 2012; Starbird et al. 2014). The social media messages containing rumors are different from those spreading news (Mendoza et al. 2010). Liao and Shi (2013) also reported the spreading of rumor during a nation-scale scandal using microblog.

7.2.3   The Outbreak of Ebola

To detect the disease outbreaks, a process was developed by the World Health Organization to action smartly in the big data context (Grein et al. 2000; Samaan et al. 2005). However, it seemed that this did not work well in West Africa because of the prevailing poverty in the region (Chan 2014). In September 2014, the US Center for Disease Control (CDC) predicted that 1.4 million people in Sierra Leone and Liberia would have succumbed to the disease if the epidemic continued its swift spread (Epstein 2014). Ebola has been frequently highlighted by newspaper, TV, and social media, after the virus has spread into many countries and has taken a lot of lives in 2014. More than 10 million tweets mentioned the word "Ebola" on Twitter within 20 days from 15 September 2014 in 170 countries (Luckerson 2014).

Of course, social media has also become new tools for us to analyze and predict how infectious diseases develop and evolve spatially and temporally. Along with the advancement of Internet technology, the spread of messages through social media had become much accelerated, more specific, and more short-lived than before. With constantly emerging social issues to prompt new rumors, the Internet actually provided an incubator for rumors to grow. This is understandable when considering that social media users came from all walks of life and participations in social media were very much unstructured and unorganized. As a results, social media, although not intentionally, nurture and spread rumors (Jin et al. 2014). Due to the widespread fear of Ebola and the lack of timely and transparent reporting of Ebola cases to the general public, rumors about Ebola infection possessed the two critical elements of Ebola infection being an important issue and the ambiguity in all news and reports about it. As a result, Ebola infection was widely discussed on social media.

Rumors about public health often bring out discontent in a society. Rumors also add unwanted noise in the use of social media information for any analysis or data exploration, creating road blocks to studies of public health issues (Ruths and Pfeffer 2014). This chapter uses the social media message regarding the 2014 Ebola infection news discussed by users of social media in China to study how rumors spread via social media. Specifically, this chapter addresses the following questions:

- What is the rumor diffusion process of Ebola in China (Ningbo and Guangzhou)?
- What are different topics regarding Ebola from social media in China (Ningbo and Guangzhou)?

- Who are those influential individuals or organizations involved in discussing the Ebola?
- What are the diffusion patterns of fear of Ebola?

This chapter is organized as the following. Section 2 reviews the research related social media, public health, and rumors. The paper discusses the method used to collect social media information and methods for unsupervised content classification, analysis, and the trends detection in Sect. 3. The following section summarizes the analysis and findings on the unique characteristics of how rumors spread in social media. Finally, we suggest future directions for additional research on this topic.

7.3   Method

In this article, we are interested in the outbreak of microblogging spatial patterns and variation-related rumors. We collected geotagged-Sina microblogging data, classified by topics using latent Dirichlet allocation (LDA) topic model (Blei et al. 2003) and analyzed by exploratory spatial data analysis.

7.3.1   Data Collection

Sina Weibo is the largest social network in China. By December 2013, the number of Weibo monthly active users reached 129.1 million, and daily active users reached 61.4 million. The posts in Sina Weibo reached the number of more than 2.8 billion in December, 2013. Sina Weibo provides developers with APIs (application programming interface) to facilitate users to extend application in Weibo. A limited sample of Sina Weibo can be collected through the official API. However, this will pose significant limitations on the detail of microblogs and the amount of data. Because web pages are unstructured, most crawlers are based on the simple definitions of the scope of web pages, which are subject to efficiency analysis (Plachouras et al. 2014). We then use a hybrid method, by integrating the API with a web crawler developed by our own to harvest the experimental data used in this study. In mid-October 2014, a rumor was spread on the Internet (Sohu 2014) about China's first Ebola-infected patient who was found in Ningbo, a coastal city in southeast China. This rumor also claimed the mortality rate of Ebola disease to be more than 90%. Almost in the same time, similar rumors about Ebola emerged in Guangzhou, another coastal city situated at the Pearl River delta (Xinhuanet 2014). Given the similarity of the characteristic of these two cities, we take the rumors in Ningbo and Guangzhou as our research objects. Considering the rumors may have some pre-warning stage, we began collecting data from the first day of October, 2014. To capture a full cycle of rumor diffusion process, we set the duration of our data collection as 60 days.

Previous study indicated that original posts on social media tend to have higher and more significant correlations than its retweets (Aslam et al. 2014); we thus solely collected original posts on Sina Weibo. After data cleaning and formatting, we obtained 1163 original microblogs about "Ningbo Ebola" and 1944 about "Guangzhou Ebola" in October and November 2014.

7.3.2   LDA

Text in each microblog contains the sentiment of user on Ebola rumors in different aspects. We use LDA model to classify the microblog content in order to explore topics which the rumors involved. LDA is an unsupervised machine learning techniques, used to identify the underlying topics from large-scale document collection and corpus, which has been already widely used for information extraction from tweets (Mehrotra et al. 2013).

LDA (Blei et al. 2003) is a widely used topic model based on the probabilistic theory. In this model, documents are composed of a series of random latent topics represented by the distribution of words. It has been used to automatically generate expertise profiles of online community (Liu et al. 2014), to quantify the distribution of probes between subcellular locations (Coelho et al. 2010), to mine opinion (Pang and Lee 2008), and to assign the annotation of large satellite images (Lienou et al. 2010).

First, we extract the microblog content from our processed dataset. After that the common stop words (such as "Ebola," because too often in the text sample, the keyword appears frequently) was removed. Then we consider a microblog as a document, using Gibbs sampling as the realization of LDA to carry topic extraction. There are two probabilistic parameters, α and β, describing the topic distribution. α represents the probabilities of topic distributions of per-document, and β is the probabilities of word distributions of per-topic.
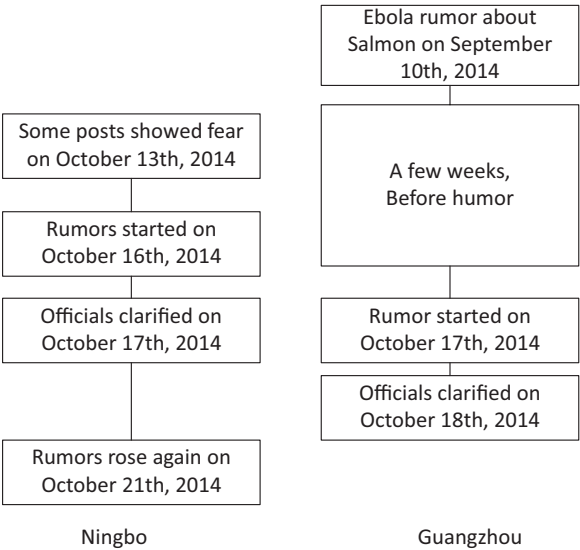
Recent years have witnessed significant interest in exploratory spatial data analysis (ESDA) techniques and application, such as crime (Grubesic and Mack 2008), regional economics (Ye and Carroll 2011; Ye and Rey 2013), transportation (Whalen et al. 2013), environment (Xie et al. 2012, 2013), and public health (Milinovich et al. 2014). ESDA should be considered as a descriptive step before suggesting dynamic factors to explain the spatial patterns under study and before estimating and testing more sophisticated regression models (Anselin 2004). ESDA can help detect the spatiotemporal complexity towards forming novel research questions.

Through deriving related Weibo information through LDA, we conduct ESDA to explore spatial, temporal, and user distribution of Ebola rumors.

7.4   Analysis and Interpretation

Ningbo is a seaport city in Zhejiang province in East China. Guangzhou is the largest city in South China and is the capital of Guangdong Province. In mid-October 2014, a rumor was spread on the Internet (Sohu 2014) saying China's first

Fig. 7.1 Rumor diffusion process of Ebola in Ningbo and Guangzhou

	Ebola rumor about Salmon on September 10th, 2014
Some posts showed fear on October 13th, 2014	A few weeks, Before humor
Rumors started on October 16th, 2014	
Officials clarified on October 17th, 2014	Rumor started on October 17th, 2014
	Officials clarified on October 18th, 2014
Rumors rose again on October 21th, 2014	
Ningbo	Guangzhou

Ebola-infected patient was found in Ningbo. This rumor also claimed the mortality rate of Ebola disease to be more than 90%. Almost in the same time, similar rumors about Ebola emerged in Guangzhou (Xinhuanet 2014) (Fig. 7.1). The fear of infectious disease such as dengue and SARS was the root for rumor to disperse among the mass media. It is also the case for Ebola. We notice that the first microblog that mentioned Ebola was posted in 13 October 2014. "Now the Ebola virus is very scary. There are a lot of Africans in Ningbo as well. Are they carrying the Ebola here?"

Our analysis shows that the rumor officially started on the night of 16 October 2014. A microblog post claimed that: "Is Ebola in China now? I am flying Taipei from Ningbo and the flight was held back for sanitation check. Someone on the plane appears to have Ebola symptom." Out of the attention and fear of Ebola, this microblog was reposted 181 times in a short while.

The former rumor microblog drew the attention from official department in Ningbo city. In the midnight of 17 October 2014, an official Weibo account named "Ningbo Health" was created. "There is a passenger from Nigeria who travelled from Taiwan to Ningbo on October 16th. He was diagnosed with certain fever symptoms. He is still under observation in a hospital." This microblog was then reposted by more than 2000 other Sina Weibo users.

Shortly after this official microblog, another official channel "Ningbo Announcement" posted a similar microblog and reposted the previous one from "Ningbo Health."

Eight hours after its initial announcement, "Ningbo Health" posted a follow-up microblog to clarify the Ebola rumor. "Update: the foreigner from Nigeria observed in the hospital who was detected slight fever yesterday has resumed to normal temperature now." Three hundred seventy-nine Sina Weibo users forwarded this post.

"Ningbo Announcement" also reposted the update microblog from "Ningbo Health" to clear the rumor. It restated that there is no clarified Ebola case in Ningbo city and advised people not to disperse and make up rumor on this event. On 21 October 2014, another similar rumor was dispersed in Sina Weibo: "the first case about Ebola was discovered in Ningbo. This disease has a mortality rate of more than 90%."

Rumor about Ebola was found in Guangzhou city 1 month before its emergence in Ningbo. On 10 September 2014, a post was spread on Sina Weibo. "By eating Salmon you will get infected with Ebola." Ever since the salmon rumor, there are occasional microblogs claiming that Ebola has emerged in Guangzhou. The Ebola rumors about Guangzhou first spread from WeChat (MicroMsg), a popular instant message tools, on the night, 17 October 2014. In a few hours later, some people sent a Weibo to request for clarifying. Similar to the official channel in Ningbo, an official channel "Sina Guangdong" got involved in the incident and stepped out to clarify this rumor. "According to Health Department of Guangdong Province, there is no Ebola case in Guangdong province so far." This clarification is against a previous rumor on Sina Weibo claiming that "a case of Ebola was identified in Canton fair. This patient is from Nigeria and he was sent to hospital for isolation." On 19 October, "Sina Guangdong" clarified against the previous rumor again.

As a transitional society, China is sensitive to many pressing issues, such as economic inequality, housing, food safety, and environmental pollution. Through forwarding and commenting negative news, people express their complaints to the virtual space, which easily lead to a variety of rumors.

7.4.1   Topics Regarding Ebola from Social Media

The content of microblog is often the description of certain event as well as a channel to disclose user's sentiment. In order to capture a comprehensive knowledge on what has been discussing on Sina Weibo, we used topic modeling method to dig out the topics among these Ebola-relevant microblogs. We treated each microblog as a document and classified them in different topics.

To extract the topics in Ningbo, we first delete the redundant microblogs as well as remove our thematic keyword "Ebola" and system keyword "Weibo." After the data preprocess of 1163 microblogs, we have 903 microblogs to run the LDA using the parameters (alpha = 2.5, beta = 0.01, and 1000 iteration) (Blei et al. 2003; Porteous et al. 2008). After generalization we have four topics (Table 7.1). The first category of topic captures the source of the rumor. It contains keywords such as "Nigeria," "men," "hospital," "fever," and "immigration." This also suggested that people know where Ebola might have come from or where Ebola infection was the most severe. Our second category of topic shows people fear towards Ebola. For example, keywords such as "intensive," "worry," "China," "prevention," and "aware" are related to the high populated area which is regarded more vulnerable to infectious disease. It also suggests that people were more concerned about their immediate

Table 7.1 Topics in Ningbo

Topic 1		Topic 2		Topic 3		Topic 4	
Word	Ratio	Word	Ratio	Word	Ratio	Word	Ratio
Nigeria	0.103	China	0.144	Eliminate	0.053	Virus	0.261
Men	0.085	Prevention	0.144	Already	0.051	Spread	0.026
Mild	0.071	Population	0.131	No	0.043	Attention	0.026
Hospital	0.066	Intensive	0.128	Epidemic	0.039	US	0.021
Fever	0.059	Aware	0.121	China	0.033	Clarify	0.020
Immigration	0.047	Worry	0.110	Scare	0.032	Mutate	0.017
Taiwan	0.036	Epidemic	0.045	Do not	0.029	Up to date	0.017
Examine	0.035	SARS	0.040	Beijing	0.025	Death	0.013
Afternoon	0.034	Afraid	0.036	Yesterday	0.023	Confirm	0.013
Transit	0.033	Possible	0.020	Nowadays	0.022	Prevention	0.013
Now	0.033	Sarsaparilla	0.007	Bye-bye	0.022	Tea leaf	0.012
Health bureau	0.028	Care	0.006	Confirm	0.020	False alarm	0.011
Report	0.021	Life	0.005	Rumor	0.019	Infectious	0.010
Epidemic	0.020	Fragile	0.004	Before	0.017	Overall	0.010
West Africa	0.020	Hurry	0.003	Center	0.014	Export	0.010
At present	0.019	Deliver	0.002	True	0.014	Occur	0.010
Infect	0.018	Container	0.001	Recently	0.014	Student	0.009
Prevention	0.018	Tencent.com	0.001	Foreigner	0.013	Said	0.009
Carry out	0.017	Sina.com	0.001	Risk	0.013	Global	0.009
Planning commission	0.016	A brand of cloth	0.001	Surprise	0.013	Decline	0.009

environment and their own well-being and people were conscious about what to do to prevent contracting the Ebola virus. In the third category of topic, we notice there are some keywords that show the negativity, such as "eliminate," "already," "no," "epidemic," "scare," etc. This gives a negative answer to whether Ebola has been emerging in China. The fourth category presents a global attention to Ebola. This includes keywords "virus," "spread," "attention," "US," and "clarify". The global spread of Ebola disease strengthens the fear of Ebola at regional and local areas, leveraging the root for rumor dispersion.

After the data preprocess, we have 1503 microblogs to run the LDA using the same parameters as Ningbo. Five topics are identified in Guangzhou Weibo messages (Table 7.2).

The first topic reflects the masses' fear towards Ebola. The keywords "black," "China," and "affected" show that due to the fact of large number African people living in Guangzhou, people tend to relate it to Ebola as it was outbreak in West Africa.

We notice the process of rumor clarification in the second topic. The keyword "news" shows a traditional channel to clear rumor, and "share" is how microblog got spread. "First case," "clarify," "official," and "rumor" show message from official source is crucial in clarification a rumor.

Table 7.2 Topics in Guangzhou

Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
Word	Ratio	Word	Ratio	Word	Ratio	Word	Ratio	Word	Ratio
Black	0.148	News	0.144	Virus	0.1	The China Import and Export Fair	0.111	Salmon	0.085
China	0.054	Share	0.067	Detect	0.037	Dengue	0.102	Friends	0.054
Affected	0.027	First case	0.039	Disease	0.033	Epidemic	0.061	Norway	0.044
Look	0.025	Clarify	0.038	Related	0.021	Strong	0.023	Safe	0.036
Hope	0.025	Official	0.036	Company	0.021	Enterprise	0.018	Network	0.025
Nowadays	0.024	Rumor	0.035	Asia	0.019	Influence	0.018	Post	0.024
Problem	0.022	NetEase.com	0.032	Local	0.019	Exhibiting	0.015	Eat	0.024
Cannot	0.018	Fear	0.022	Death	0.018	Today	0.014	Recently	0.024
True	0.018	Expert	0.022	Field	0.018	Opening	0.014	Repost	0.023
Control	0.018	Message	0.019	Infect	0.018	Yesterday	0.014	Sales	0.02
Illegal	0.017	Sina.com	0.013	Gene	0.017	Data	0.012	Halved	0.017
Should	0.016	Cause	0.013	A company	0.016	Pay attention	0.011	Rumor	0.014
Seen	0.016	Client	0.012	Protection	0.015	Decline	0.01	Never	0.014
African	0.015	Public	0.012	Dose	0.015	Current fair	0.01	Seafood	0.013
Worry	0.015	Up to date	0.01	Locate	0.013	Client	0.009	Like	0.013
Ill	0.014	ifeng.com	0.009	Success	0.012	Foreigner	0.009	Really	0.013
Lives	0.01	Tonight	0.009	In China	0.012	Attend	0.009	Reporter	0.012
Predict	0.009	Appear	0.009	Produce	0.011	Buyer	0.008	Import	0.011
Come in	0.009	Channel	0.008	Respirator	0.011	Product	0.008	Relief	0.011
Terrific	0.008	TV program	0.008	Market	0.01	Fair	0.006	Declare	0.011

Table 7.3 The post numbers
of users in Ningbo

Posts	Users
1	1089
2	31
3	2
4	3
6	1

The third topic shows how virus was detected and prevented. "Virus," "detect," "disease," and "related suggestion" indicate that detection and prevention towards virus are still the key process of public health. Topic 4 relates to China Import and Export Fair which is held every spring and autumn in Guangzhou. This event is a famous and top fair for overseas trade, attracting many international buyers and sellers. This represents general public's concern of Ebola on Import and Export Fair.

Topic 5 describes the issue of salmon. Salmon eaters in China have been frightened by titles of the popular articles in the social media like Weibo and WeChat as "China has stopped the import of a whole Norwegian salmon," "People are likely to get Ebola from eating salmon," and "Salmon carry with them the Ebola virus" in September 2014.

By comparing topics from Ebola rumors in Guangzhou and Ningbo, we notice two similar topics: fear towards Ebola and the clarification of rumor. Fear towards Ebola is the foundation of rumor, while the clarification of rumor terminates the rumor dispersion process.

There are different perspectives of two rumor events. First of all, the source of rumor is different. Rumor in Ningbo started from a scam microblog while rumor in Guangzhou was initially communicated in WeChat. Secondly, due to the fact that a large number of Africans are living in Guangzhou, there is delusion from some people that Ebola is close to them in Guangzhou. That explains people in Guangzhou pay more attention to the topic on detection and prevention of infectious disease. In addition, in the same period, an import and export fair was held in Guangzhou, and this brings in several businessmen from African. Again, people in Guangzhou show great speculation on whether the fair is related to Ebola. Last but not least, a rumor about eating salmon will lead to Ebola was pushing the rumor in Guangzhou as well. In a word, Guangzhou being a city in an economically prosperous region has more opportunity for being exposed to Western culture/influences. In Ningbo, the Ebola rumor is an accident breakout wherein a passenger from Nigeria was sent to a hospital in airport.

We summarize the post number of each Weibo user, and the result is presented in Table 7.3. As we can see, most users (1089) contributed just one post during the rumor process. The maximum number of posts a user publishes is six. All users who contribute more than four posts are individual type of users. This indicates that in the process of special event, official channel tends to be cautious.

Statistics on user post number in Guangzhou shows similar pattern except that there are more people who publish large number of microblogs (Table 7.4). This is due to the fact that Guangzhou has large population than Ningbo and also the mass media in Guangzhou is more active.

Table 7.4 The post numbers
of users in Guangzhou

Posts	Users
1	1426
2	121
3	26
4	9
5	9
6	4
7	2
8	1
11	2

Table 7.5 Top five reposted users in Ningbo

# of reposts	User	Type	City	Followers
7930	Jiangning Police Online	Official	Nanjing	834,668
1318	Top news	Official	Beijing	37,060,313
1043	CCTV News	Official	Beijing	25,825,867
959	Ningbo Health	Official	Ningbo	16,832
809	News in Realtime	Official	USA	107,728

Table 7.6 Top five reposted users in Ningbo

# of reposts	User		Type	City	Followers
1366	Sina Guangdong	Official	Guangzhou	1,227,876	1366
1029	People.com.cn	Official	Beijing	26,197,418	1029
908	huanqiu.com	Official	US	5,138,602	908
791	CCTV News	Official	Beijing	25,825,867	791
711	caijing.com.cn	Official	Beijing	11,287,244	711

Study of reposts shows that a total of 233 posts have been forwarded or reposted.
The top five users that receive reposts are official channels from the local government.
This indicates the trustiness from local citizens on their government (Tables 7.5
and 7.6).

Similar to Ningbo, the top five users that have been reposted are official users.
This again shows the authority of official account on Sina Weibo.

7.4.1.1 Users Asked for Clarification

In Ningbo, five users asked for verification from seven official accounts: two police
accounts, two news accounts, two other government accounts, and one Weibo
company account. Because the rumor in Guangzhou did not originate in Weibo, not
many verification requests were identified in Weibo. The main verification is on the
report from Zhujiang News.

7.4.1.2   Users Spread the Rumor

In Ningbo, 455 Weibo messages were rumor. Four hundred forty-five users (68 certified users) were involved. Thirty-one Weibo messages and 31 users (11 certified users) spread the rumor in Guangzhou. In both cities, most common users (uncertified users) form the main user group in spreading the rumor.

7.4.1.3   Users Helping the Official Source Confirm the Accurate Information

In Ningbo, 233 Weibo messages confirmed the accurate information. Among 229 users involved, there were 59 official accounts, 46 certified accounts, and 124 uncertified accounts. In Guangzhou, 194 Weibo messages confirmed the accurate information. Among 180 users involved, there were 52 official accounts, 94 certified accounts, and 34 uncertified accounts. In both cities, common users (uncertified users) are minority on helping the official source confirm the accurate information.

7.4.1.4   User and Levels of Postings

To explore the relationship between user characteristics and message repost, we present the scatter plot involving followers and repost (Fig. 7.2). Ningbo and Guangzhou have similar patterns: more followers, more reposts.

At the same time, we examine the relationship between number of friends and number of message reposts. Both cities witnessed the similar patterns again. As (Rapoport and Rebhun 1952) argued, a higher degree of connection in friendship will facilitate the spread of rumor (Fig. 7.3).
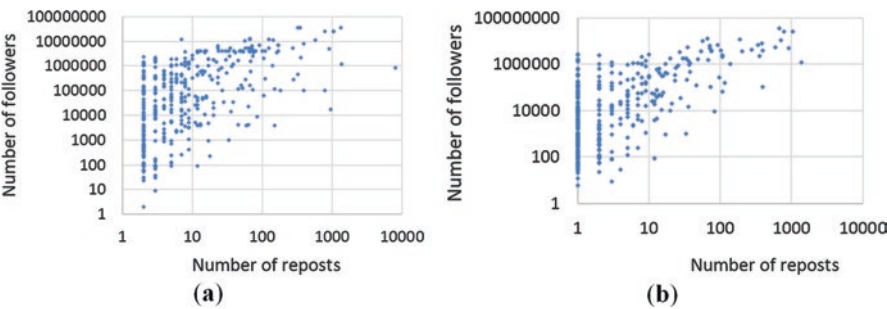

Fig. 7.2 Repost and followers (**a**) in Ningbo, (**b**) in Guangzhou
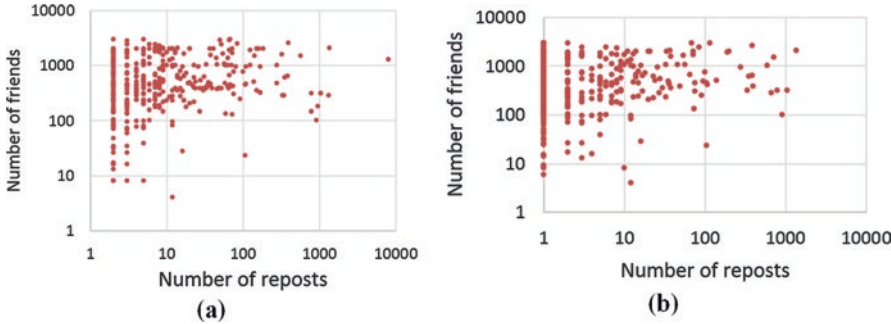
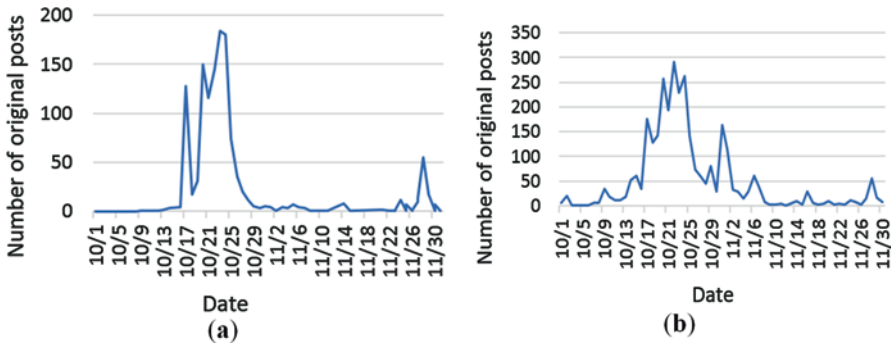**Fig. 7.3** Reposts and friends (**a**) in Ningbo, (**b**) in Guangzhou

Fig. 7.4. Trend of two months (**a**) in Ningbo, (**b**) in Guangzhou

7.4.2 Diffusion Patterns of Fear of Ebola in Space and Time

7.4.2.1 Trend in Time

1. By days

By plotting the number of microblogs related to Ebola by days, we can track the temporal trend of the evolvement of rumor process. Figure 7.4. (a) presents that from the rumor time 16 October, the number of microblogs started to surge, and there are two peaks in Ningbo's rumor. One indicates the fear to Ebola and another one corresponds to the prevalence of rumor.

Figure 7.4.b shows that there is one peak in terms of the number of microblog posts in Guangzhou by each date. This is majorly due to the "salmon" rumor related to Ebola. Another find is that the temporal peak of both cities appears to be in the same period, indicating that the same topic of rumor correlates in time
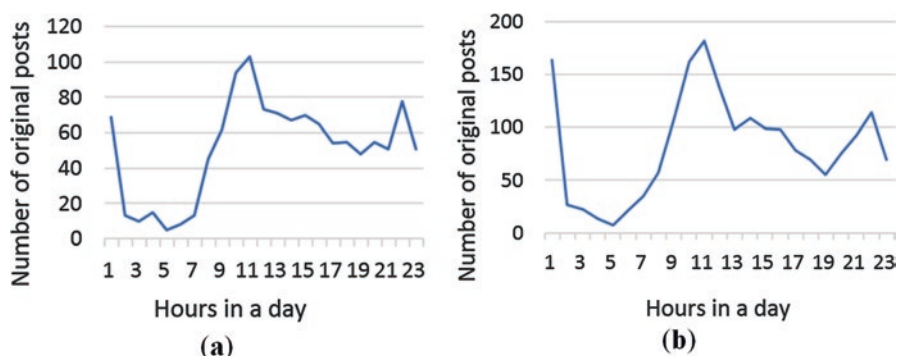
Fig. 7.5 Trend of a day (**a**) in Ningbo, (**b**) in Guangzhou

2. By hours

By plotting the time of each Ebola microblog of two cities in 24 h every day, we can see the temporal trend of Guangzhou and Ningbo corresponds well to the masses' daily activities (Fig. 7.5).

7.4.2.2 Trend in Space

The source of Ebola rumor in Ningbo originated from the Nigerian passenger who traveled to Ningbo international airport while the rumor in Guangzhou came from the local WeChat community with no further source. We then use geographic information system to map each microblog according to its location in order to explore the trend of rumor in space.

1. Distribution by locations

Figure 7.6 shows the geographic distribution of related microblog messages in China. As we can see, rumors about Ningbo's Ebola rumor were scattered beyond Zhejiang province. However, users who got involved in Guangzhou's Ebola rumor were mostly from Guangdong Province.

In order to further test the spatial cluster, we use Average Nearest Neighbor (Wong and Lee 2005) to test the degree of spatial cluster of points. The results show that the score of nearest neighbor ratios of users' location on Ningbo rumor is 0.2932 while the score of Guangzhou users is 0.4526. These numbers are less than the expect value 1, which indicates both point patterns appear to be cluster in space. By standardizing the two scores using Z-score, with Ningbo's Z-score being −10.1524 while Z-score in Guangzhou being −18.6868, both cities' NNR results are statistically significant.
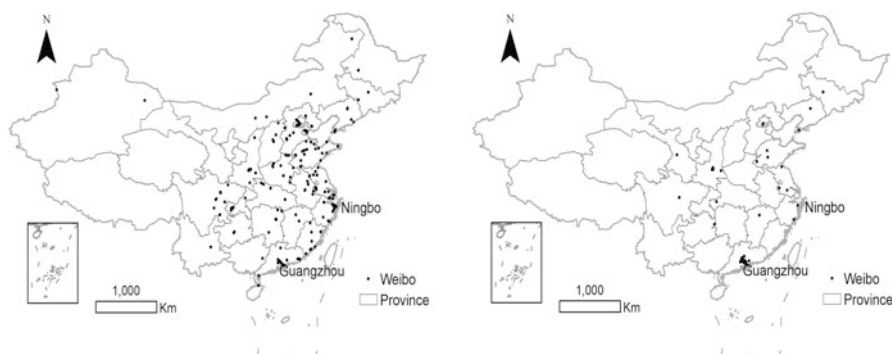
Fig. 7.6 The locations of Weibo related to (**a**) Ningbo, (**b**) Guangzhou

2. Distribution by profile

Users' profile can be used to identify a coarse spatial location (Hecht et al. 2011). In this study, we classify user location by province. In the Ningbo case, Zhejiang (where Ningbo is located) has the most users, followed by Beijing and Guangdong which also witnessed quite a few users. Regarding the Guangzhou case (where Guangdong is located), most users are identified in Guangdong, followed by Zhejiang. The rumors in these two places seem to be locally associated and also related to each other due to the same theme.

7.5 Discussion and Conclusion

During emergencies such as natural disasters or disease outbreak, it is often of critical importance to assess impact and response of people and the surrounding environment as an indicator for better evacuation planning and relief operations. Communication channels during times of emerging social crises and natural disasters are critical before, during, and after such events. Our study shows that social media messages are rich in content, which are ideal to capture and reflect a multitude of aspects of attitude, behaviors, and reactions of social media users to a specific topic or event. We thus argue that there is a pressing need to understand social media use during natural disasters and emergent social event from the perspective of social media users and affected citizens. This is also reflected in other study indicating that during a disaster, affected citizens are usually the most valuable sources who are capable of providing information before the public channel and agency. Therefore, it has been widely acknowledged that Humanitarian Assistance and Disaster Relief (HADR) responders can gain valuable insights and situational

awareness by monitoring social media and volunteered geographic information related to the emergent event (that study also examined a set of factors that can explain those uses, such as time of the tweets, the location and characteristics of the users, previous patterns of social media use, and the degree to which the crisis has directly or indirectly affected the users). Therefore, these messages can be used to monitor and track geopolitical and disaster events, support emergency response and coordination, and serve as a measure of public interest or concern about events (Huang and Xiao 2015).

In conclusion, this study is an exploratory research conducted for the 2014 Ebola season that uses social media as a possible method for identifying trends in fear of Ebola incidence. The specific Ebola rates and trends are not included in this paper because we want to address the rumors spread in social media and not the detection and prediction of the spread of influenza. There is no Ebola patient found in China in the incidence. In terms of temporal dimension, across the two cities that Weibo messages were collected from, Weibo rates and rumor rates both peaked around 23 November 2014, and both peaked on 11 o'clock in 1 day. Though Weibo is not the birthplace of the original rumor, it facilitates its spread because most Weibo users are anonymous. Driven by fear, the rumor spread demonstrates a fast outbreak within very short time period. LDA method illustrates that rumors relate to real events associated with Ebola, which represents the collective efforts in the context of emergency (Peterson and Gist 1951; Rosnow 1988). In addition, rumors have localized spatial pattern. The advantages to using social media to survey rumor incidence is that it not only would quicken response time for official, but also it would give a various aspects, which will help us to find the sources and reason, about the rumors. When a rumor spreads, it will spread to social network in a few hours or in seconds even. The topics discussed in Weibo about Ebola rumor show us the source of rumor, the event, and focus of people.

Although the attributes of social media messages, such as user name, text of post, time, location, as well as the number of followers and being followed, can be collected by crawler and API, age, gender, and race are not available in most social media platforms. Hence, it is difficult to detect who spread rumors from demographic information. According to a report in 2013, about 53% of users are under 23 years old, who might be significantly affected by rumors (Bai 2014). In the growing social media platform, millions of people create, share, and exchange information and thoughts. Further research can be investigated in other rumors with various spatiotemporal features. Social media platform can be enhanced through the development of an early warning system predicting the outbreaks of rumor diffusion, which can help prevent the spread of humors.

References

Achrekar, H., Gandhe, A., Lazarus, R., Yu, S. H., & Liu, B. (2011, April). Predicting flu trends using twitter data. In Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on (pp. 702–707). IEEE.

Anselin, L. (2004). Exploring spatial data with GeoDaTM: A workbook. *Urbana, 51*, 61801.

Aslam, A. A., Tsou, M., Spitzberg, H. B., An, L., Gawron, M. J., Gupta, K. D., Peddecord, M. K., Nagel, C. A., Allen, C., Yang, J., & Lindsay, S. (2014). The reliability of tweets as a supplementary method of seasonal influenza surveillance. *Journal of medical Internet research, 16*, e250.

Bai, S., Yuan, S., Hao, B., & Zhu, T. (2014). Predicting personality traits of microblog users. Web Intelligence and Agent Systems: An International Journal, 12(3), 249–265.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2012). GenBank. *Nucleic Acids Research, 41*, s1195.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research, 3*, 993–1022.

Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein, J. S., Hoen, A. G., Moyes, C. L., Farlow, A. W., Scott, T. W., & Hay, S. I. (2012). Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Neglected Tropical Diseases, 6*, e1760.

Brownstein, J. S., Freifeld, C. C., Reis, B. Y., & Mandl, K. D. (2008). Surveillance sans Frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Medicine, 5*, e151.

Buckner, H. T. (1965). A theory of rumor transmission. *Public Opinion Quarterly, 29*, 54–70.

Chan, M. (2014). Ebola virus disease in West Africa--no early end to the outbreak. *The New England Journal of Medicine, 371*, 1183–1185.

Coelho, L. P., Peng, T., & Robert. (2010). Quantifying the distribution of probes between subcellular locations using unsupervised pattern unmixing. *Bioinformatics, 26*, i7–i12.

Epstein, H. (2014). Ebola in Liberia: An epidemic of rumors. *New York Review of Books, 61*, 91.

Floris, R., & Zoppi, C. (2015). Social media-related geographic information in the context of strategic environmental assessment of municipal Masterplans: A case study concerning Sardinia (Italy). *Future Internet, 7*, 276–293.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*(4), 211–221.

Grein, T. W., Kamara, K. B., Rodier, G., Plant, A. J., Bovier, P., Ryan, M. J., Ohyama, T., & Heymann, D. L. (2000). Rumors of disease in the global village: Outbreak verification. *Emerging Infectious Diseases, 6*, 97–102.

Grubesic, T. H., & Mack, E. A. (2008). Spatio-temporal interaction of urban crime. *Journal of Quantitative Criminology, 24*, 285–306.

Hay, S. I., George, D. B., Moyes, C. L., & Brownstein, J. S. (2013). Big data opportunities for global infectious disease surveillance. *PLoS Medicine, 10*, e1001413.

Hecht, B., Hong, L., Suh, B., & Chi, E.H. (2011). Tweets from Justin Bieber's heart: The dynamics of the location field in user profiles. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 237–246). Vancouver: ACM.

Huang, Q., & Xiao, Y. (2015). Geographic situational awareness: Mining tweets for disaster preparedness, emergency response, impact, and recovery. *ISPRS International Journal of Geo-Information, 4*(3), 1549–1568.

Indrawan-Santiago, M., Miyabe, M., Nadamoto, A., & Aramaki, E. (2014). How do rumors spread during a crisis? Analysis of rumor expansion and disaffirmation on twitter after 3.11 in Japan. *International Journal of Web Information Systems, 10*, 394–412.

Jin, F., Wang, W., Zhao, L., Dougherty, E., Cao, Y., Lu, C., & Ramakrishnan, N. (2014). Misinformation propagation in the age of twitter. *Computer, 47*, 90–94.

Kay, S., Zhao, B., & Sui, D. (2014). Can social media clear the air? A case study of the air pollution problem in Chinese cities. *The Professional Geographer, 67*, 1–13.

King, D., Ramirez-Cano, D., Greaves, F., Vlaev, I., Beales, S., & Darzi, A. (2013). Twitter and the health reforms in the English National Health Service. *Health Policy, 110*, 291–297.

Liao, Q., & Shi, L. (2013). She gets a sports car from our donation: Rumor transmission in a Chinese microblogging community. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (pp. 587–598). San Antonio: ACM.

Lienou, M., Maitre, H., & Datcu, M. (2010). Semantic annotation of satellite images using latent Dirichlet allocation. *Geoscience and Remote Sensing Letters, IEEE, 7*, 28–32.

Liu, X., Wang, G. A., Johri, A., Zhou, M., & Fan, W. (2014). Harnessing global expertise: A comparative study of expertise profiling methods for online communities. *Information Systems Frontiers, 16*, 715–727.

Luckerson, V. (2014, October 7). Watch how word of Ebola exploded in America. *Time*.

MacEachren, A.M., Robinson, A.C., Jaiswal, A., Pezanowski, S., Savelyev, A., Blanford, J., & Mitra, P. (2011). Geo-twitter analytics: Applications in crisis management. In *25th International Cartographic Conference* (pp. 3–8).

Maynard, D., Gossen, G., Funk, A., & Fisichella, M. (2014). Should I care about your opinion? Detection of opinion interestingness and dynamics in social media. *Future Internet, 6*, 457–481.

Mehrotra, R., Sanner, S., Buntine, W., & Xie, L. (2013). Improving LDA topic models for microblogs via tweet pooling and automatic labeling. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 889–892). Dublin: ACM.

Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis: Can we trust what we RT? In *Proceedings of the First Workshop on Social Media Analytics* (pp. 71–79). Washington DC: ACM.

Messina, J.P., Brady, O.J., Pigott, D.M., Brownstein, J.S., Hoen, A.G., & Hay, S.I. (2014). A global compendium of human dengue virus occurrence. EP.

Milinovich, G. J., Williams, G. M., Clements, A. C. A., & Hu, W. (2014). Internet-based surveillance systems for monitoring emerging infectious diseases. *The Lancet Infectious Diseases, 14*, 160–168.

Mullen, P. B. (1972). Modern legend and rumor theory. *Journal of the Folklore Institute*, 95–109.

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. 1–14. *ICIS, 231*.

Okada, Y., Ikeda, K., Shinoda, K., Toriumi, F., Sakaki, T., Kazama, K., Numao, M., Noda, I., & Kurihara, S. (2014). SIR-extended information diffusion model of false rumor and its prevention strategy for twitter. *Journal Ref: Journal of Advanced Computational Intelligence and Intelligent Informatics, 18*, 598–607.

Padmanabhan, A., Wang, S., Cao, G., Hwang, M., Zhang, Z., Gao, Y., Soltani, K., & Liu, Y. (2014). FluMapper: A cyberGIS application for interactive analysis of massive location-based social media. *Concurrency and Computation: Practice and Experience, 26*, 2253–2265.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*, 1–135.

Peterson, W. A., & Gist, N. P. (1951). Rumor and public opinion. *American Journal of Sociology*, 159–167.

Plachouras, V., Carpentier, F., Faheem, M., Masanès, J., Risse, T., Senellart, P., Siehndel, P., & Stavrakas, Y. (2014). ARCOMEM crawling architecture. *Future Internet, 6*, 518–541.

Porteous, I., Newman, D., Ihler, A., Asuncion, A., Smyth, P. & Welling, M. (2008). Fast collapsed gibbs sampling for latent dirichlet allocation. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 569–577). ACM.

Rapoport, A., & Rebhun, L. I. (1952). On the mathematical theory of rumor spread. *The Bulletin of Mathematical Biophysics, 14*, 375–383.

Rogers, D. J., Wilson, A. J., Hay, S. I., & Graham, A. J. (2006). The global distribution of yellow fever and dengue. *Advances in Parasitology, 62*, 181–220.

Rosnow, R. L. (1988). Rumor as communication: A contextualist approach. *Journal of Communication, 38*, 12–28.

Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science, 346*, 1063–1064.

Sakaki, T., Okazaki, M., & Matsuo, Y. (2013). Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Transactions on Knowledge and Data Engineering, 25*, 919–931.

Samaan, G., Patel, M., Olowokure, B., Roces, M. C., & Oshitani, H. (2005). Rumor surveillance and avian influenza H5N1. *Emerging Infectious Diseases, 11*, 463–466.

Sohu. (2014). China's first Ebola-infected persons appeared Ningbo? Rumors!

Spiro, E.S., Fitzhugh, S., Sutton, J., Pierski, N., Greczek, M., & Butts, C.T. (2012). Rumoring during extreme events: a case study of deepwater horizon 2010. In *Proceedings of the 4th Annual ACM Web Science Conference* (pp. 275–283). Evanston: ACM.

Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R.M. (2014). Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 Boston marathon bombing. In *iConference 2014 Proceedings* (pp. 654–662). Boston.

Sui, D. Z. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment and Urban Systems, 32*(1), 1–5.

Velardi, P., Stilo, G., Tozzi, A. E., & Gesualdo, F. (2014). Twitter mining for fine-grained syndromic surveillance. *Artificial Intelligence in Medicine*, 153–163.

Whalen, K. E., Páez, A., & Carrasco, J. A. (2013). Mode choice of university students commuting to school and the role of active travel. *Journal of Transport Geography, 31*, 132–142.

Widener, M. J., & Li, W. (2014). Using geolocated twitter data to monitor the prevalence of healthy and unhealthy food references across the US. *Applied Geography, 54*, 189–197.

Wong, W., & Lee, J. (2005). *Statistical analysis of geographic information with ArcView GIS and ArcGIS*. Wiley: New Jersey.

Xie, H., Kung, C., & Zhao, Y. (2012). Spatial disparities of regional forest land change based on ESDA and GIS at the county level in Beijing-Tianjin-Hebei area. *Frontiers of Earth Science, 6*, 445–452.

Xie, H., Liu, Z., Wang, P., Liu, G., & Lu, F. (2013). Exploring the mechanisms of ecological land change based on the spatial autoregressive model: A case study of the Poyang lake eco-economic zone, China. *International Journal of Environmental Research and Public Health, 11*, 583–599.

Xinhuanet. (2014). Guangzhou now Ebola? Rumors! Guangdong Province Write appeared Ebola cases.

Yang, X., Ye, X., & Sui, D. Z. (2016). We know where you are: In space and place-enriching the geographical context through social media. *International Journal of Applied Geospatial Research (IJAGR), 7*(2), 61–75.

Yap, L. F., Bessho, M., Koshizuka, N., & Sakamura, K. (2012). User-generated content for location-based services: A review. In *Virtual communities, social networks and collaboration* (pp. 163–179). New York: Springer.

Ye, X., & Carroll, M. C. (2011). Exploratory space-time analysis of local economic development. *Applied Geography, 31*, 1049–1058.

Ye, X., & Rey, S. (2013). A framework for exploratory space-time analysis of economic data. *The Annals of Regional Science, 50*, 315–339.

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), 7–33.