Xu et al. EURASIP Journal on Wireless Communications and Networking (2016) 2016:44
DOI 10.1186/s13638-016-0553-0

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Participatory sensing-based semantic and spatial analysis of urban emergency events using mobile social media

Zheng Xu[1,2*], Hui Zhang[1], Vijayan Sugumaran[3], Kim-Kwang Raymond Choo[4], Lin Mei[2] and Yiwei Zhu[5]

Abstract

With the advances of information communication technologies, it is critical to improve the efficiency and accuracy of emergency management systems through modern data processing techniques. Geographic information system (GIS) models and simulation capabilities are used to exercise response and recovery plans during non-disaster times. They help the decision-makers understand near real-time possibilities during an event. In this paper, a participatory sensing-based model for mining spatial information of urban emergency events is introduced. Firstly, basic definitions of the proposed method are given. Secondly, positive samples are selected to mine the spatial information of urban emergency events. Thirdly, location and GIS information are extracted from positive samples. At last, the real spatial information is determined based on address and GIS information. Moreover, this study explores data mining, statistical analysis, and semantic analysis methods to obtain valuable information on public opinion and requirements based on Chinese microblog data. Typhoon Chan-hom is used as an example. Semantic analysis on microblog data is conducted and high-frequency keywords in different provinces are extracted for different stages of the event. With the geo-tagged and time-tagged data, the collected microblog data can be classified into different categories. Correspondingly, public opinion and requirements can be obtained from the spatial and temporal perspectives to enhance situation awareness and help the government offer more effective assistance.

Keywords: Participatory Sensing, Urban emergency events, Social media, Urban computing

1 Introduction

With the advances of information communication technologies, such as cloud computing [1, 2], Internet of Things [3, 4], and big data [5, 6], it is critical to improve the efficiency and accuracy of emergency management systems through modern data processing techniques. The past decade has witnessed the tremendous technical advances in sensor networks, Internet/Web of Things, cloud computing, mobile/embedded computing, spatial/temporal data processing, and big data, and these technologies have provided new opportunities and solutions to emergency management. Data processing/analysis in emergency management is a typical big data scenario. Numerous sensors and monitoring devices continuously sample the states of the physical world, while the web data processing techniques make the Internet a big data repository which can reflect the states of the cyber world and the human world. The efficient processing of these data imposes a challenge to the data management community. It is important to develop advanced data management and data processing mechanisms to support disaster detection, disaster response and control, rescue resource planning and scheduling, and emergency commanding. The objective conditions, such as the lack of information, timely changeable situation, short time for decision-making, and serious consequences, bring a great challenge for the government emergency response. The extraction of accurate information is very important in emergency response. Social media provide a platform for data mining and information extraction.

* Correspondence: juven_xz@163.com
This paper is the extended version (50 % new content) of the conference paper accepted by EM-GIS2015.
[1]Tsinghua University, Beijing, China
[2]The Third Research Institute of the Ministry of Public Security, Shanghai, China
Full list of author information is available at the end of the article

Emergency management aims to develop strategies and establish operations to decrease the potential impact of unexpected events (i.e., human or natural disasters). By quick response and rescue, it saves human lives from the secondary disasters and enhances the stability of communities after disasters. Emergency management involves four stages: planning and mitigation, preparedness, response, and recovery. Geospatial applications (including geographic information system (GIS)) have been extensively used in each stage of emergency management. Decision-makers can utilize the geospatial information to develop planning and mitigation strategies. GIS models and simulation capabilities are used to exercise response and recovery plans during non-disaster times. They help the decision-makers understand near real-time possibilities during an event.

The recent evolution of the Internet has permitted an unprecedented increase in content created by non-specialist users, thanks to a reduction in technical barriers [7]. Web users can provide their geographic information through social media, such as Twitter[1] and Weibo[2]; these are now commonly referred to as volunteered geographic information (VGI), having a huge potential to engage citizens in place-based issues and provide significant, timely, and cost-effective source for geographer's and other spatially related fields of research and management [8]. Usually, the Weibo users can be "social sensors [9]." A social sensor is defined as an agent that provides information about its environment on a social network after interaction with other agents [10]. The sensing message from social sensors can be used during a live fire emergency or traffic updates. Studies on emergency response using social media data have drawn much academic attention. Existing studies mainly focus on three perspectives: (1) information spreading on social network; (2) geolocation information subtraction; and (3) semantic analysis. The information dissemination of social media can enhance timely situation awareness in a crisis situation, because they can report on-the-ground information from general public and allow the government to spread information to a wide group of people. The geolocation information of social media data plays an important role in emergency detection and quick response. The geolocation related to social media data included content-based location, posting location, and registration location, and they were used to estimate the geographical location of a crisis. The semantic analysis of twitter messages is also of significance in situation awareness and emergency response. It can help extract information on damages, request for assistance, casualties, or warning to enhance situation awareness and receive assistance during crisis.

In this paper, a participatory sensing-based model for mining spatial information of urban emergency events is introduced. Firstly, basic definitions of the proposed method are given. Secondly, positive samples are selected to mine the spatial information of urban emergency events. Thirdly, location and GIS information are extracted from positive samples. At last, the real spatial information is determined based on address and GIS information. Moreover, this study explores data mining, statistical analysis, and semantic analysis methods to obtain valuable information on public opinion and requirements based on Chinese microblog data. Typhoon Chan-hom is used as an example. Semantic analysis on microblog data is conducted, and high-frequency keywords in different provinces are extracted for different stages of the event. With the geo-tagged and time-tagged data, the collected microblog data can be classified into different categories. Correspondingly, public opinion and requirements can be obtained from the spatial and temporal perspectives to enhance situation awareness and help government offer more effective assistance.

The rest of the paper is organized as follows. In the next section, the related work is given. Section 3 presents the proposed model. Case studies on real data sets are conducted in Section 4. The application on the proposed method is given in Section 5. The last section gives the conclusion of our work.

2 Related works

The early related work in this emerging research field has addressed the use of social media to generate map mashups to support collaborative real-time mapping [11] and give the overview of harvesting geospatial content [12]. Earle et al. [13] attempted assessing how fast tweeters reacted to the smaller (4.3 magnitude) and much more localized earthquake of Morgan Hill, CA, in March, 2009. Recently, the work of harvesting spatial information from the web has seen some activity in recent years. For example, it has been demonstrated that general purpose points of interest (POI) can be automatically derived from the users' map annotations [14] and vague geographic regions (e.g., Midlands or Middle West) delineated. Besides the texture content, georeferenced pictures from the photo-sharing website such as Flickr have been processed in terms of their density to show where the most famous landmarks are for a given location [15]. The Geospatial Exploratory Data Mining Web Agent that retrieves geographic information from web pages (related to outdoor activities) has also recently been discussed [16].

Recently, with the high-speed development of the social networks, such as Twitter and Weibo, many researchers have published their work of using the data from social networks including special events and localization of natural disasters [17]. Sakaki et al. [18] investigated the real-time nature of Twitter and put particular attention to event detection. The Twitter users are regarded as sensors. Their messages are used for detecting earthquake. A reporting system is developed for use in Japan by their proposed methods. Crooks et al. [19] thought Twitter as a distributed sensor system. They analyzed the spatial and temporal features of the Twitter feed activity responding to a 5.8-magnitude earthquake. Their experimental results argued that the Twitter users represent a hybrid form of a sensor system that allows for the identification and localization of the impact area of the event. Longueville et al. [20] used Twitter as a source of spatial-temporal information. By focusing on a real-life case of forest fire, they aimed to demonstrate its possible role to support emergency planning, risk, and damage assessment activities.

3 The proposed method

In this section, the proposed method is illustrated. Firstly, basic definitions of the proposed method are given. Secondly, positive samples are selected to mine the spatial information of urban emergency events. Thirdly, the semantic analysis of the social media is given. Fourthly, location and GIS information are extracted from positive samples. At last, the real spatial information is determined based on address and GIS information.

3.1 Basic definitions

In this section, basic definitions of the proposed method are given. An event is something that happens at some specific time, and often some specific places [15]. In the proposed method, the spatial information of an urban emergency event can be mined since some messages sensing by social users have exact location information. An urban emergency event is defined as follows.

3.1.1 Definition 1. Urban emergency event, e

An urban emergency event e is a tubal {T, L}, where T and L is the life course and spatial of e.

3.1.2 Definition 2. Words set of an urban emergency event e, W

Words set consists of keywords of an urban emergency event. Usually, these words can be used to search messages of an urban emergency event. It is noted that elements in the words set are synonyms since social users may use different words to describe the same urban emergency event.

Crowdsourcing or participatory sensing may be a potential solution mining the spatial information of urban emergency events. The proposed method aims at collecting and analyzing the spatial information from social users. The social network can be seen as a sensor receiver. Usually, the social network users can be seen as social sensors. The proposed method is set as a hierarchical data model including three different layers. The different layers of the proposed method are illustrated in Fig. 1.

(1) Social user layer. In this layer, the proposed method wants to collect the related data of urban emergency events. For example, if a user makes a message in Weibo about a fire occurrence, then the proposed method should collect this message.
(2) Crowdsourcing layer. In this layer, address and GIS information is mined from messages from social users. Of course, location and GIS information related to the same urban emergency events should be clustered.
(3) Spatial information layer. In this layer, the spatial information of the urban emergency event is mined.

3.2 Selecting positive samples

3.2.1 Definition 3. Positive samples of an urban emergency event e, PS

Positive sample is the set of messages posted by social users, which can be used to mine the spatial information of an urban emergency event. For example, the message “A fire is happened at Jiangsu Road.” can be seen as a positive sample. Based on common sense and our observations on real data, we have three heuristics to select positive samples from Weibo search results.

Heuristic 1: The Weibo message with the location or GIS information is prone to be a positive sample.
Since the location information is an important aspect for mining spatial information, the message with the location information can be thought as a positive sample. Different from the location information in Weibo message, the GIS information means the location of social users.
Heuristic 2: The Weibo message with the image or video information is prone to be a positive sample.
The image or video information is an important evidence to show the presence of social users. If a social user posts a message with the image or video of an urban emergency event, he is prone to be a witness of this event. Thus, he is also prone to post the accuracy spatial information of the urban emergency event.
Heuristic 3: The original Weibo message is prone to be a positive sample.
If a message is a forward message other than an original message, the social user is not a witness of the urban emergency event.

3.3 Semantic analysis

Semantic analysis aims at finding useful data from mass data to obtain valuable information or knowledge. Sina Weibo, one of the most popular microblogs in China,

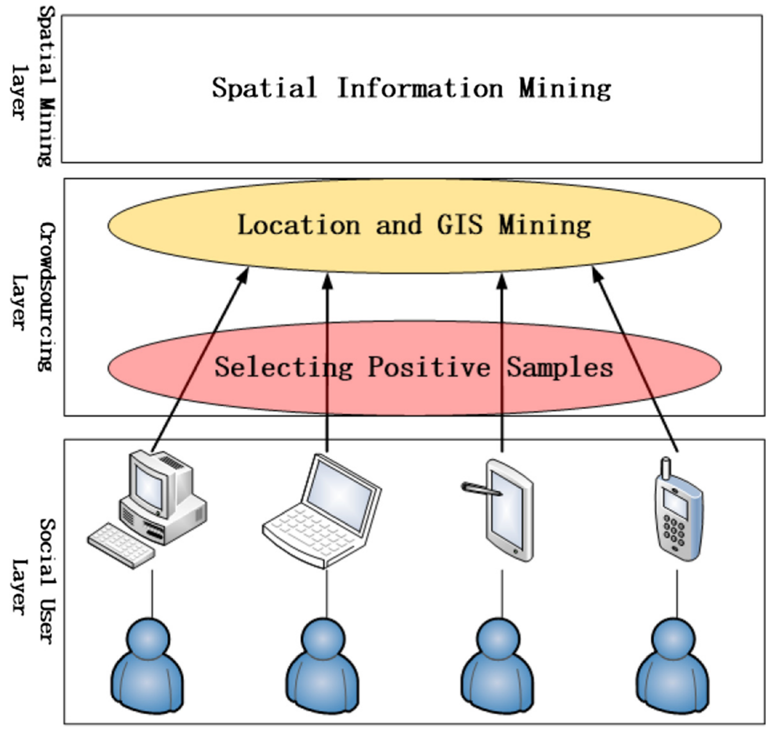

Fig. 1 The hierarchical structure of the proposed method

provides a platform for data mining. People express and record what they saw, heard, or experienced in the self-organized platform, and they also receive information from it. Sina Weibo is a huge information interaction center and the source of microblog data collection in this paper. The keyword-based search method was applied to collect microblog data through the Weibo API (an open interface of Sina Weibo). In the search, only microblog messages in Chinese language were collected and there were no geolocation constraints.

Text classification is to divide all the messages into several categories, and then semantic analysis is conducted in a certain category. The geo-tagged and time-tagged information of microblog messages was the standard for text classification. The collected microblog data were classified based on the geographic location of the posters. The number of microblog posts was counted in each province.

3.4 Mining location and GIS information

3.4.1 Definition 4. The location information of an urban emergency event e, LI

The location information can be extracted from the posted messages from positive samples. For example, in Fig. 2, a user post a message "I see a fire at Huaihai Road", the posted location information "Huaihai Road" can be extracted. In the proposed method, the location information detection is based on the Baidu Map, which can detect whether a word is a location name or not.

3.4.2 Definition 5. The GIS information of an urban emergency event e, GI

Unlike the location information, the GIS information is usually extracted from the check-in information from Weibo messages. The GIS information can be mined from the element of HTML page. For example, in Fig. 2, the GIS information of the message is "geo = 121.46, 31.22".

3.5 Determine spatial information

The spatial information of an urban emergency event usually consists of two aspects including the happening spatial information and the spread spatial information. The happening spatial information means the happening place of the urban emergency event. For example, in Fig. 2, the happening place of the fire is 358 Huaihai Road. The spread spatial information means the influential scope of the urban emergency event. For example, in Fig. 2, the influential place of the fire is Ruijing Road, which is shown in the check-in information. The influential place is usually equal to the check-in information of the social users. The check-in information means that the social users can see the happening emergency events in his place.

4 Simulations on real urban emergency events

4.1 The brief background of the Typhoon Chan-hom event

Typhoon Chan-hom hit Zhejiang province at 16:40 on July 11, 2015 as a category 2 storm. It was one of the

Fig. 2 The illustration example of location information. The message is "358 Huaihai Road is under fire". The check-in is "Ruijing 1 road, Luwan District, Shanghai"

strongest ones to hit the region this year and caused tremendous damages. More than one million people were evacuated in Zhejiang, and total economic losses in the province amounted to ¥8.86 billion (US$1.43 billion).

Typhoon Chan-hom formed in the northwest Pacific Ocean and was classified as a tropical depression on June 30, 2015. It was assigned the name Chan-hom when it upgraded to a tropical storm on July 1. With heavy storm and extreme rain, Chan-hom made landfall on Zhoushan, Zhejiang, on July 11, 2015 and the maximum sustained winds reached to 45 m/s. Thereafter, the storm accelerated north-northeast and moved over the Yellow Sea with degradation of its structure. Chan-hom was tracked each day based on the physical data mainly including GIS data, wind speed, rain rate, storm surge, and typhoon intensity.

4.2 The semantic analysis result

The keyword-based search method was applied to collect microblog data through the Weibo API (an open interface of Sina Weibo). The keywords included "Can Hong" (the Chinese of "Chan-hom"), "typhoon," and "Typhoon Chan-hom," and the period of data collection was from July 8, 2015 to July 15, 2015. In the search, only microblog messages in Chinese language were collected, and there were no geolocation constraints. After filtering the irrelevant data, the number of event-related microblog posts was 3321 with the geographic location of the posters and posting time attached. The number of microblog posts was counted in each province. The top four provinces were Zhejiang, Jiangsu, Shanghai, and Beijing, and their details were shown in Table 1. Based on the posting time, the collected data were classified into two categories: (1) before typhoon landfall (0:00 on July 8 to 16:40 on July 11) and (2) after typhoon landfall (16:40 on July 11 to 23:59 on July 15). The number of microblog posts at these two stages was 1154 and 2167, respectively, as shown in Table 2.

4.3 Risk perception and public opinion

Risk perception can be estimated by the number of social media messages to a certain degree. Risk perception was a driver to motivate a person to use social media during crisis. The scale that people use social media reflects people's risk perception during emergency. The number of microblog posts in different provinces was estimated based on the geolocation

Table 1 Classification by geolocation (the top four provinces)

Provinces	Zhejiang	Jiangsu	Shanghai	Beijing
Number of microblog posts	1320	284	257	206

Table 2 Classification by posting time

Stages	Before landfall	After landfall
Period	July 8, 0:00–July 11, 16:40	July 11, 16:40–July 15, 23:59
Number of microblog posts	1154	2176

information. The risk distribution map was drawn based on the number in each province, as shown in Fig. 3. It was clear that the citizens' risk perception in some areas was relatively high, such as Beijing and some coastal provinces, i.e., Zhejiang, Jiangsu, and Shanghai. The closer away from the landfall place was, the higher the citizens' risk perception was. Beijing, as the capital of China, citizens there tends to pay more attention on crisis than people in other provinces.

The semantic analysis on the geo-tagged microblog data helped obtain public opinion from the spatial perspective. The assistance could be offered to where it was required. Keyword-frequency analysis and comparative analysis were conducted in three provinces where risk was relatively high. These provinces were Zhejiang, Shanghai, and Beijing. They were the representatives of the landfall place, the adjacent province, and the capital, respectively.

In this study, public opinion was extracted from the keywords whose frequencies were higher than 10. The number after a keyword represented its frequency. The comparative analysis of keywords distribution of these three provinces was shown in Fig. 4. The extracted keywords were classified into five categories based on their meanings. They were current situation, damage information, response, future situation, and citizens' requirements. The number of keywords describing current situation was the biggest of the five categories in all provinces. The keywords of Zhejiang mainly clustered in four categories: current situation, damage information, response, and citizens' requirements while Shanghai and Beijing were current situation, response, and future situation. It was clear that the number of keywords was bigger and their frequencies were higher in Zhejiang than in other provinces.

The attention was paid more by the citizens in the place where a disaster happened than in other provinces. The citizens there concerned the damages and their needs while the citizens in other places wanted to know more on future situation, such as the possible track, intensity, tendency, and meteorology.

4.4 Spatial mining on the real explosion event

The *explosion* is also a major damage to the urban especially the big city. With the development of terrorism, the number and influence of the explosion event has

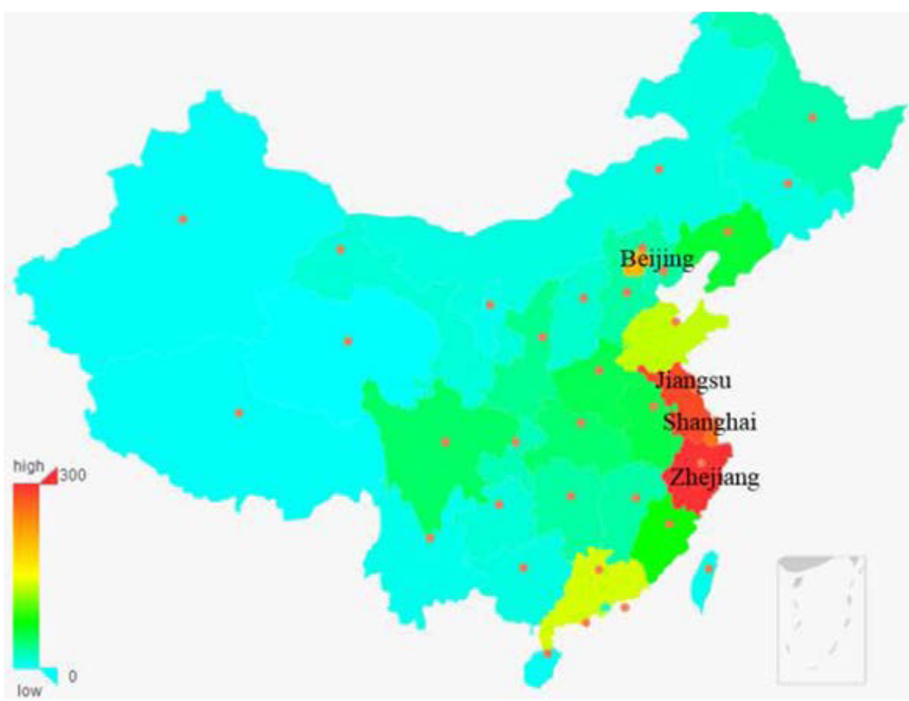

Fig. 3 Risk distribution map. In order to make the figure readable, the maximum of the range was 300 though the number of microblog posts in Zhejiang was 1320

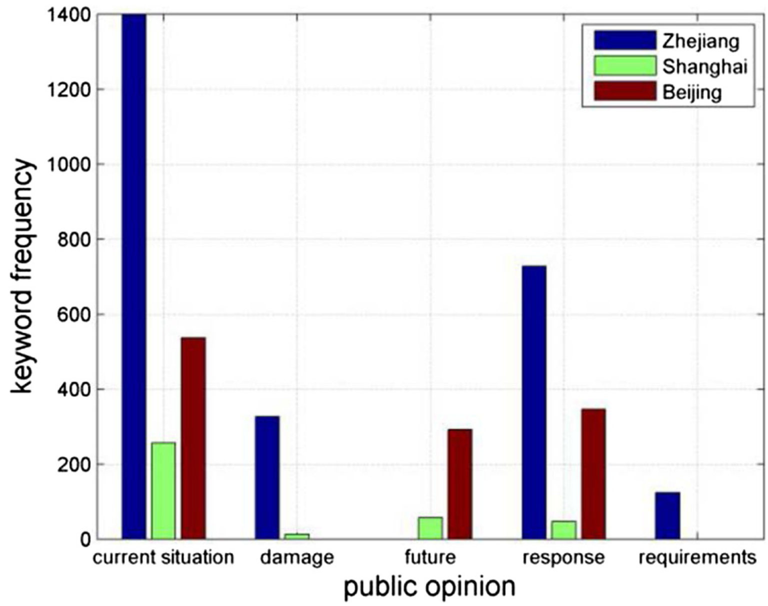

Fig. 4 Comparative analysis of keyword-frequency distribution in different provinces

become higher than before. For example, the "Boston Marathon" explosion event causes dozens of people dead. Thus, it is important to build the knowledge base of explosion emergency event. For example, the image or video of an explosion can be used by policeman to detect the suspect.

We select an explosion event that happened in 18:50 at "Taoyuan Apartment" on January 17, 2015 in Dalian. The proposed method is used to build the knowledge base of that event. The word "爆炸 (explosion)" is used to search in Weibo from 18:00 to 22:00. The search location is set as Dalian. Totally, 478 messages are returned. According to principles 1, 2, 3, and 4, 45 messages providing location information, check-in information, and image are selected as candidate messages. Figure 5 shows the timeline description of this case study. Only six timestamps are selected to build the timeline since they provide new keywords or locations. The red point in the map means the real location of that fire event. The red point with number means the GIS information of the different timestamps.

5 The applications on the proposed method

In this section, a natural gas leakage monitoring system is developed, which composes sensor network and data collection system based on cloud storage platform. The sensor network consists both mobile and stationary sensors, which help to provide large-scale city areas monitoring with less sensors and provide more rapid and accurate leakage locating with GIS mapping.

The functions of concentration monitoring are as follows. (1) Concentration monitoring includes the provision of a methane concentration heat map. The original data collected by the mobile detectors and stationary detectors are merely a cluster of discrete locations that describe the relationship between the GPS information and methane concentration. Therefore, these data locations must be processed using a classic Gaussian plume diffusion model in order to obtain the concentration distribution of the entire monitored area. Then, these data are shown by means of distribution on a GIS map. (2) Concentration monitoring also includes mobile detectors. The real-time locations (latitude, longitude, and height) and methane

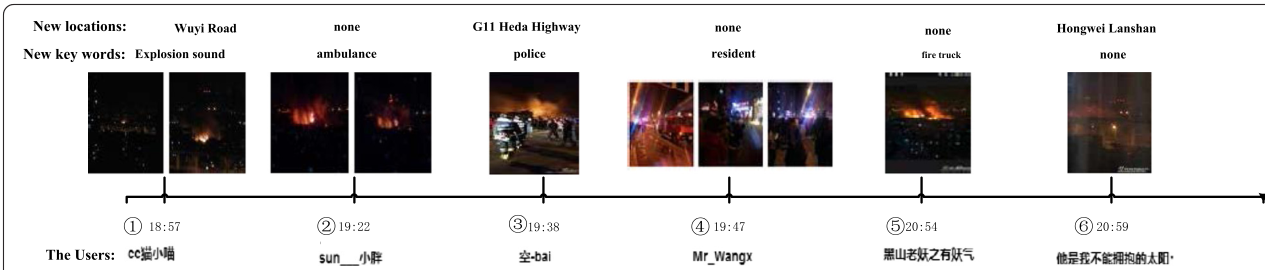

Fig. 5 The "explosion" case study using the proposed method

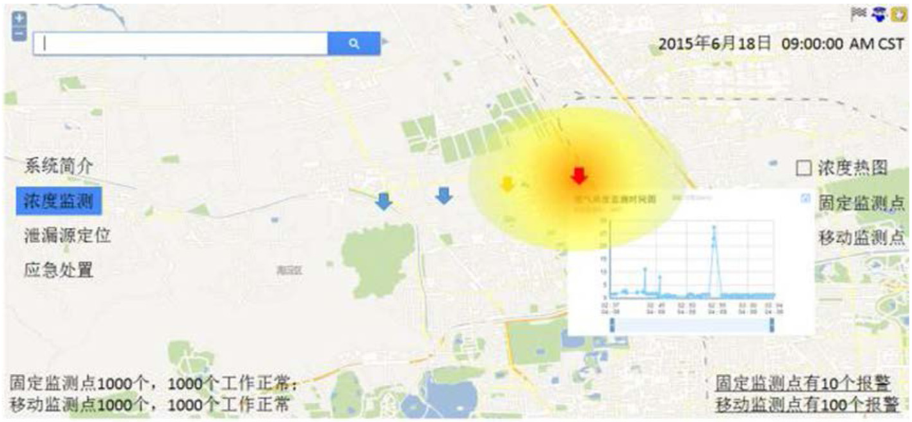

Fig. 6 Comparative analysis of keyword-frequency distribution in different provinces

concentrations are shown on the GIS map. In addition, every mobile detector is designated by a different color (red, orange, yellow, or green) according to the methane concentration. Red signifies the highest methane concentration and green the lowest. (3) Concentration monitoring includes stationary detectors. The real-time methane concentration and wind speed and direction obtained from each stationary detector are shown on the GIS map. The color designation method is similar to that for the mobile detectors. (4) The history data obtained from all the detectors can be queried via history methane concentration curves when the user clicks on a detector. The history data provide accuracy and reliability with regard to leakage warning. (5) Concentration monitoring includes the working state of the detectors. All the data for the citywide natural gas leakage monitoring system come from every detector, so the working state of the detectors determines whether the system is working normally. Therefore, it is necessary to monitor the working state of all the detectors. Each detector should periodically transfer a diagnostic report message to the control center. If the detector sends a faulty report or no report, then the detector is faulty and the maintenance staff must address such problems. (6) Concentration monitoring includes the level of warning. The real-time warning presents the security status of the monitored area. (7) Concentration monitoring includes a query detector. Users are able to query each detector according to the detector number and highlight it on the GIS map.

By clicking on the leakage location, the relative position information will appear on the GIS map, including the district name and street names, as well as the locations of significant subjects, such as maintenance center, firefighter offices, hospitals, schools, and chemical plants. A message will be sent automatically to relevant offices, such as management office of gas companies, maintenance center, and emergency response office. Figure 6 shows examples of the system.

6 Conclusions

In this paper, a participatory sensing-based model for mining spatial information of urban emergency events is introduced. Firstly, basic definitions of the proposed method are given. Secondly, positive samples are selected to mine the spatial information of urban emergency events. Thirdly, location and GIS information are extracted from positive samples. At last, the real spatial information is determined based on address and GIS information. Moreover, this study explores data mining, statistical analysis, and semantic analysis methods to obtain valuable information on public opinion and requirements based on Chinese microblog data. Typhoon Chan-hom is used as an example. Semantic analysis on microblog data is conducted and high-frequency keywords in different provinces are extracted for different stages of the event. With the geo-tagged and time-tagged data, the collected microblog data can be classified into different categories. Correspondingly, public opinion and requirements can be obtained from the spatial and temporal perspectives to enhance situation awareness and help government offer more effective assistance.

7 Endnotes

[1]www.twitter.com
[2]www.weibo.com

Competing interests
The authors declare that they have no competing interests.

Authors' information
Zheng Xu was born in Shanghai, China. He received Diploma and Ph.D. degrees from the School of Computing Engineering and Science, Shanghai University, Shanghai, in 2007 and 2012, respectively. He is currently working in the third research institute of ministry of public security and the postdoctoral in Tsinghua University, China. His current research interests

include topic detection and tracking, semantic Web, and Web mining. He has authored or co-authored more than 70 publications including IEEE Trans. On Fuzzy Systems, IEEE Trans. On Automation Science and Engineering, IEEE Trans. On Cloud Computing, IEEE Trans. On Emerging Topics in Computing, and IEEE Trans. on Systems, Man, and Cybernetics-Part C.
Hui Zhang is a professor at the Department of Engineering Physics, Tsinghua University. Hui Zhang's primary research lies in the area of thermal/fluid science, crystal growth, disaster prevention and mitigation, and thermal spray coating. Prof. Zhang is now the PI of a major research program of NSFC focused on "study on the city's comprehensive risk assessment and several key issues in emergency response". He is one of the directors of the UTC-Tsinghua Joint Institute on Building Energy, Safety and Control System. He focuses on developing next-generation building safety technology through integrated information and sense systems, effective situation awareness, and evacuation strategy.
Vijayan Sugumaran is a professor of Management Information Systems and Chair of the Department of Decision and Information Sciences at Oakland University, Rochester, MI, USA. He received his Ph.D. in Information Technology at George Mason University, Fairfax, VA, USA. His research interests are in the areas of Big Data Analytics, Intelligent Information Technologies, Ontologies, and Semantic Web. He has published over 150 peer-reviewed articles and edited 12 books. He has published in top-tier journals, such as "Information Systems Research," "ACM Transactions on Database Systems," "IEEE Transactions on Engineering Management," "Communications of the ACM," "IEEE Transactions on Education," and "IEEE Software". Dr. Sugumaran is the editor-in-chief of the *International Journal of Intelligent Information Technologies*. He is the Chair of the Intelligent Agent and Multi-Agent Systems mini-track for Americas Conference on Information Systems (AMCIS 1999–2016). He has served as the program co-chair for the International Conference on Applications of Natural Language to Information Systems (NLDB 2008, NLDB 2013 and NLDB 2016).
Kim-Kwang Raymond Choo received Ph.D. in Information Security in 2006 at Queensland University of Technology, Australia. He is currently a senior researcher at the University of South Australia. He was named one of the 10 Emerging Leaders in the Innovation category of The Weekend Australian Magazine/Microsoft's Next 100 series in 2009 and is the recipient of various awards including ESORICS 2015 Best Research Paper Award, Highly Commended Award in the 2014 "Best Chapter in a Book" Category by Australia New Zealand Policing Advisory Agency (ANZPAA), a Fulbright Scholarship in 2009, and a 2008 Australia Day Achievement Medallion.
Lin Mei received Ph.D. degree at Xi'an Jiaotong University, Xi'an, China, in 2000. He is a research fellow. From 2000 to 2006, he was a postdoctoral researcher at Fudan University, Shanghai, China; the University of Freiburg, Freiburg im Breisgau, Germany; and the German Research Center for Artificial Intelligence. He is currently the Director of the Technology R&D Center for the Internet of Things with the Third Research Institute of the Ministry of Public Security, China. He has published more than 40 papers. His research interests include computer vision, artificial intelligence, and big data processing.
Yiwei Zhu is an associate professor at Zhejiang Business Technology Institute, China. His current research interests include machine learning and data mining.

Acknowledgements
This work was supported in part by the National Science and Technology Major Project under Grant 2013ZX01033002-003, in part by the National High Technology Research and Development Program of China (863 Program) under Grant 2013AA014603, in part by the National Science Foundation of China under Grant 61300202, in part by the China Postdoctoral Science Foundation under Grant 2014 M560085, the project of Shanghai Municipal Commission of Economy and Information under Grant 12GA-19, and in part by the Science Foundation of Shanghai under Grant 13ZR1452900.

Author details
[1]Tsinghua University, Beijing, China. [2]The Third Research Institute of the Ministry of Public Security, Shanghai, China. [3]Oakland University, Rochester, Michigan 48309, USA. [4]University of South Australia, Adelaide, Australia. [5]Zhejiang Business Technology Institute, Ningbo, China.

Received: 23 November 2015 Accepted: 4 February 2016
Published online: 09 February 2016

References
1. Z Xu et al., Knowle: a semantic link network based system for organizing large scale online news events. Futur. Gener. Comput. Syst. **43–44**, 40–50 (2015)
2. L Wang, J Tao et al., G-Hadoop: MapReduce across distributed data centers for data-intensive computing. Futur. Gener. Comput. Syst. **29**(3), 739–750 (2013)
3. Z Xu et al., Semantic based representing and organizing surveillance big data using video structural description technology. J. Syst Software **102**, 217–225 (2015)
4. Z Xu, X Luo, S Zhang, X Wei, L Mei, C Hu, Mining temporal explicit and implicit semantic relations between entities using web search engines. Futur. Gener. Comput. Syst. **37**, 468–477 (2014)
5. C. Hu, Z. Xu, et al, Semantic Link Network Based Model for Organizing Multimedia Big Data. IEEE Trans. Emerg. Top. Comput. **2**(3), 376–387 (2014)
6. X Luo, Z Xu, J Yu, X Chen, Building association link network for semantic link on web resources. IEEE Trans. Autom. Sci. Eng. **8**(3), 482–494 (2011)
7. Business Models for the Next Generation of Software. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html (2005)
8. M.F. Goodchild, Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. Int. J. Spatial Data Infrastructures Res. **20**, 24–32 (2007)
9. T. Sakaki, M. Okazaki, and Y. Matsuo. in *Proceedings of the 19th international World Wide Web conference*. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors, (2010), p. 851–860
10. V Krishnamurthy, H Vincent Poor, A tutorial on interactive sensing in social networks. IEEE Transac. Comput. Soc. System **1**(1), 3–21 (2014)
11. K Field, J O'Brien, Cartoblography: experiments in using and organizing the spatial context of micro-blogging. Trans. GIS **15**, 5–23 (2010)
12. A. Stefanidis, A. Crooks, J. Radzikowski, Harvesting ambient geospatial information from social media feeds. GeoJournal. **77**, 50–62 (2012)
13. P Earle, M Guy, R Buckmaster, C Ostrum, S Horvath, A Vaughan, OMG earthquake! Can Twitter improve earthquake response? Seismol. Res. Lett. **81**, 246–251 (2010)
14. L Mummidi, J Krumm, Discovering points of interest from users' map annotations. GeoJournal **72**(3–4), 215–227 (2008)
15. D. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg, in *Proceedings of the 18th International World Wide Web Conference*. Mapping the World's Photos. (2009), p. 761–761
16. E. Pultar, M. Raubal, M.F. Goodchild, *in* Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. GEDMWA: Geospatial Exploratory Data Mining Web Agent, (2008)
17. M Trusov, V Bodapati, E Bucklin, Determining influential users in internet social networks. J. Market. Res. **47**, 643–658 (2010)
18. T Sakaki, M Okazaki, Y Matsuo, Tweet analysis for real-time event detection and earthquake reporting system development. IEEE Trans. Knowl. Data Eng. **25**(4), 919–931 (2013)
19. A. Crooks, A. Croitoru, A. Stefanidis, and J. Radzikowski, Earthquake: Twitter as a Distributed Sensor System. (Transaction in GIS, 2012), pp. 1–26
20. B. Longueville, R. Smith, and G. Luraschi, in *Proceedings of the International Workshop on Location-Based Social Networks*. OMG, from here I can see the flames, a use case of mining location based social networks to acquire spatio-temporal data on forest fires, (2009), p. 73–80