The Real-time Monitoring System of Social Big Data for Disaster Management

SEONHWA CHOI
Disaster Information Research Division
National Disaster Management Institute
136 Mapo-daero, Mapo-Gu, Seoul 121-719
Republic of Korea
shchoi333@gmail.com http://www.ndmi.go.kr

Abstract: In recent years, social media has become ubiquitous and important for social networking and content sharing. Especially, in the disaster area, social media supports backchannel communications, allowing for wide-scale interaction that can be collectively resourceful, self-policing, and generative of information that is otherwise hard to obtain. At the time of great Japan and Haiti earthquake, social media channels actively were utilizing to grasp the damage, to warn, and to exchange information. In this paper, we introduce the real-time monitoring system of social big data, named Social Big Board, for disaster management. This system crawls social big data, especially Twitter, analyses the disaster-related tweets, and displays disaster situations and trends in a map. We substantiated that there is the potential for utilization on our system to monitor disasters situations and trends in real time and grasp the point from big tweet data.

Key-Words: **Disaster Management, Social Big Data, Monitoring -System**

1 Introduction

During a crisis, social media supports backchannel communications, allowing for wide-scale interaction that can be collectively resourceful, self-policing, and generative of information that is otherwise hard to obtain[1]. User-generated contents through social media will often the first publicly provided material. The most immediate forms such as micro-blogging(Twitter) and social media(Facebook), will be prevalent in the earliest stages. For example, within one week of the 2010 earthquake in Haiti, “more than one in 10 Americans(13%) - including 24% of those younger than 30 - say that they’ve gotten or shard information about the Haiti earthquake through Facebook, Twitter, or another social networking site[2]. Social media should also be accessible during crisis from multitude of digital handheld devices.

At the time of great Japan earthquake, social media channels actively were utilizing to grasp the damage, to warn, and to exchange information. Consequently, social media has opened up the infinite potential of utilization for a prompt, accurate, and intelligent disaster management. During hurricane Sandy, New York was, unsurprisingly, the world headquarters of tweets about sandy. Its involvement in storm, coupled with its population, made the city the overwhelming social media voice of the event. The left map in Fig. 1 which geo-tagged tweets occurred in the days before and during the storm was created by Floating Sheep. Tweet locations are aggregated to census blocks. What they found was that Manhattan, with the greatest population and wealth, put out the most tweets. The right map in Fig. 1 show total number of tweets referencing floods per county during hurricane Sandy. Red dots in this map indicate locations of tweets. Tweets referencing flooding are almost exactly where you would expect them to be; in other words, the vast majority of tweets were located in the path of the hurricane. In this sense, the geography data shadows drawn from Twitter appear to be quite effective at reflecting experiences of the storm[3][4].

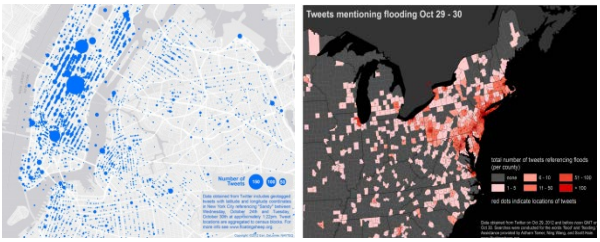

Fig. 1 Hurricane sandy and the geographies of flooding on twitter

In Korea as well as in the world, there were many cases related to disasters showing the power of social media. For instance, on Sep., 2010,

Gwanghwamun inundation damage occurred affected by a localized torrential rainfall in Seoul city on the first day of Chuseok holidays. At that time, there was no media promptly informing the damage situations. Only twitter reported inundation sings, status, evacuation tricks, and so on using pictures and messages related to inundation damage in real time. On July, 2011, Gangnam inundation damage was the same as Gwanghwamun damage.

The below Fig. 2 shows the number of tweets for a day at the time of Gangnam inundation damage. At the time of Gangnam inundation damage, 4 million tweets occurred for a day in Korea, and after filtering words related to inundation: flood, torrential rain, flowing backward, disaster, and so on, there remained 97 thousand tweets. Through extracting just tweets containing the location information of Gangnam, we could find out that 57 hundred tweets related to Gangnam inundation damage occurred.

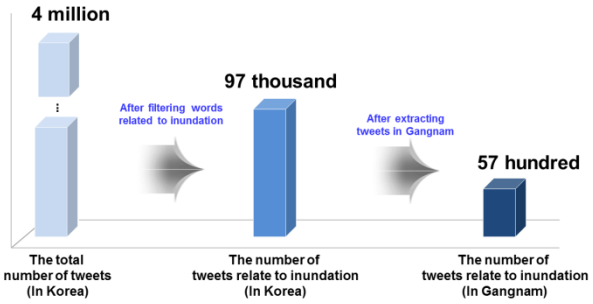

Fig. 2 The number of tweets for a day

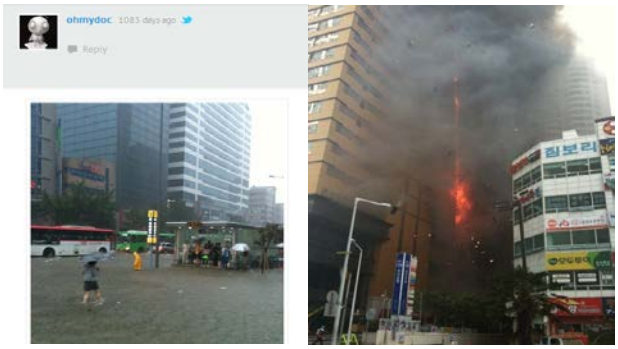

Fig. 3 Disaster scene photos within tweets

Since social media can also be constructed as a form of collective wisdom which is prompt, accurate, and well-formed information with a high propagation capacity, it has a big potential power on a disaster area as well as various categories.

Community information and other communication activities enabled by social media have increasingly gained prominence in the disaster area. The millions out of power in areas affected by super-storm Sandy have been struggling to find gas stations still open with resources. But one group of high school students in New Jersey have overnight devised a solution. Members of IMSOCIO at Franklin High School gathered Wednesday night to launch a crowd-sourced map that locates open gas stations in the New York-New Jersey area. Stations are identified by green, red, or yellow pins-each representing an open, sold out or charging station, where anyone could report on the availability of gas at stations there were visiting. The innovative use of social media and mapping technology to make live fuel supply data streams available filled a significant community information gap.

Fig. 4 Mappler service for reporting gas stations

With social media, everyone has the potential to be watchdogs, citizen journalists, photo journalists, and caring or nosy neighbours who can constantly survey the world around them and share what they find online. Like the above-mentioned, social media has the big potential for disaster management. In other words, disaster damage can be prevented or mitigated beforehand through the system and capacity to excavate and cope with disaster signs which can be detected by monitoring social media.

Therefore, in order to mitigate disaster damage, the enhancement of a scientific and intelligent disaster management system through the integrated analysis between unstructured data such as social media and structured data such as precipitation is necessary.

In this paper, we introduce the real-time monitoring system of social big data, named Social Big Board, for disaster management. This system crawls social big data, especially Twitter, analyses the disaster-related tweets, and displays disaster situations and trends in a map.

ISBN: 978-960-474-383-4

2. Social Big Board

2.1. Definition

Twitter which is an extremely popular online microblogging service has attracted lots of attention from a disaster management organization for the immense potential it provides for signs, situations, and feedback of disaster. So, we developed the real-time monitoring system of Twitter, named Social Big Board, for disaster management. This system can inform disaster signs and trends through social media at one view. To sense disaster signs and trends, several procedures described in Fig. 6(Appendix) are necessary. First of all, crawling and collecting the real-time tweets, about 360 thousands tweets occurring in Korea, by API functions provided by Twitter is performed, and then natural language processing: a morpheme analysis for elimination of stop and spam data, extracting main keywords from tweets, and filtering disaster-related tweets through comparing extracted keywords with pre-defined keywords by disaster experts will be performed. Finally, disaster-related tweets will be classified into 58 types of disaster involving social disasters as well as natural disasters.

2.2. Main Functions

There are two main functions in Social Big Board: real-time social monitoring and social trends for disaster management. A screenshot of Social Big Board shown in Fig. 5 is a service screen of real-time social monitoring.

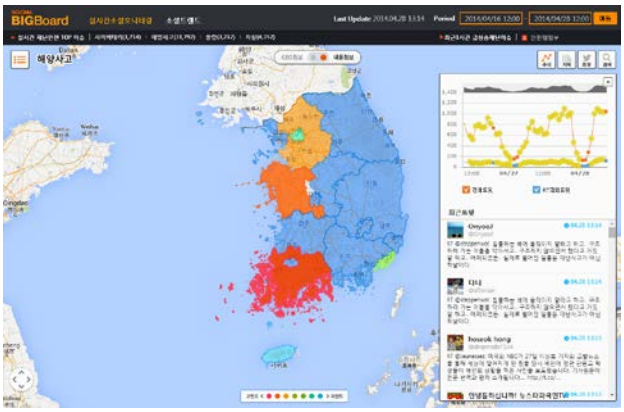

Fig. 5 A screenshot of Social Big Board

A Korean map in the middle of a screen represents tweet frequencies of any disaster during a specific period by regional groups such as administrative districts. Colour of a region represents tweet frequencies which a region name appeared in a tweet, and the redder is, the more tweet frequencies are. Just a tweet opening it's GPS (Global Positioning System) data, location and text of tweets also displayed in a map.

On the right hand side of screen, there are four buttons which provide following functions: tweet frequency trends hourly, tweet frequency of regional groups, tweet retrieval by regional groups, and search with in tweet texts. In addition, this system can detect skyrocketing disaster types beyond the normal frequency of probability model for one hour which built based on tweet frequencies in the past and inform this fact to a disaster monitor.

Second main function is social trends related to disasters. That provide related words set to any disaster. This could make easy to grasp the point without reading whole text of tweets related to a disaster. And also, provide emotional information which is analysed using pre-defined emotional words such as negative, neutral, and positive. Since this information represents people's emotion on a risk and a disaster, it will be useful information to plan and prepare the disaster policy reflecting people's demands according to disaster management levels.

3. Conclusion and future works

It cannot be emphasized enough that social media allows for wide-scale interaction that can be collectively resourceful, self-policing, and generative of information that is otherwise hard to obtain in a disaster area. In this paper, we introduce the real-time monitoring system of social big data, named Social Big Board, for disaster management. This system crawls social big data, especially Twitter, analyses the disaster-related tweets, and displays disaster situations and trends in a map. From Dec. 2013, Social Big Board has been test-operated at National Disaster Management Institute in Korea. Using Social Big Board, there were many monitoring cases: contagious disease, heavy snow fall, yellow dust, ferry sinking, and so on. By using monitoring information, we started to public the monthly disaster issues report, named Social Media & Disaster, from Mar. 2014.

ISBN: 978-960-474-383-4

By using our system to manage disasters, we could expect the changes of disaster management: not predicting a simple hazard based on the structured data but sensing and analysing signs using social data as well as the structured disaster data. We have a long way yet to put this system to practical use. For intelligence disaster management, we plan to research and develop the big data technique for disaster sign detection and response from 2013 by two step planning. First, for sensing disaster using unstructured data, we will develop the analysis and monitoring technology of unstructured data, mainly text, and then build the disaster sign sensing system. In the future, by developing the integrated analysis between structured data and unstructured data, namely big data, an intelligence disaster response system will come true.

References

[1] Jeannette S., Leysia P., and Irina S., "Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires", Proceedings of the 5th International ISCRAM Conference, Washington DC., USA May 2008.

[2] Centers for Disease Control and Prevention (CDD), Crisis Emergency Risk Communication, 2012.

[3] The Urban Geographies of Hurricane Sandy in New York City, floating sheep, http://www.floatingsheep.org/2012/10/the-urban-geographies-of-hurricane.html.

[4] Hurricane Sandy and the Geographies of Flooding on Twitter, floating sheep, http://www.floatingsheep.org/2012/10/hurricane-sandy-and-geographies-of.html.

[5] Lee Y.J., Park Y.S., and Choi S.H., "Developing an Intelligent System for the Analysis of Signs of Disaster", Proceedings of TIEMS Annual Conference, 2011.

[6] Seon-Hwa Choi, Woo-Jeong Choi, Jong-Kook Lee, "Monitoring and Sensing Disaster Issues from Social Media," Proc. of the 40th Conference of the KIPS, vol. 20 no. 1, 2013, pp. 332 ~ 335.

[7] Seung-Sig Kang, Morpheme Analysis and Information Retrieval, 2002.

[8] Sheng Yu, Subhash Kak, "A Survey of Prediction Using Social Media", The Computing Research Repository, 2012.

[9] Beob Mo Ryu, "Social Media Issue Detection & Monitoring based on Deep Language Analysis Techniques", Journal of Korean Institute of Information Science and Engineers, vol. 30 no. 6, 2012, pp.47 ~ 58.

[10] B. Liu, "Opinion mining & Summarization-Sentiment analysis", Proc. of WWW 2008 Tutorial, 2008.

[11] Plutchik, R, & Kelleman H., Emotion: Theory, research, and experience: Vol. 1. Theories of emotion, 1980, pp.3 ~ 33.

Appendix

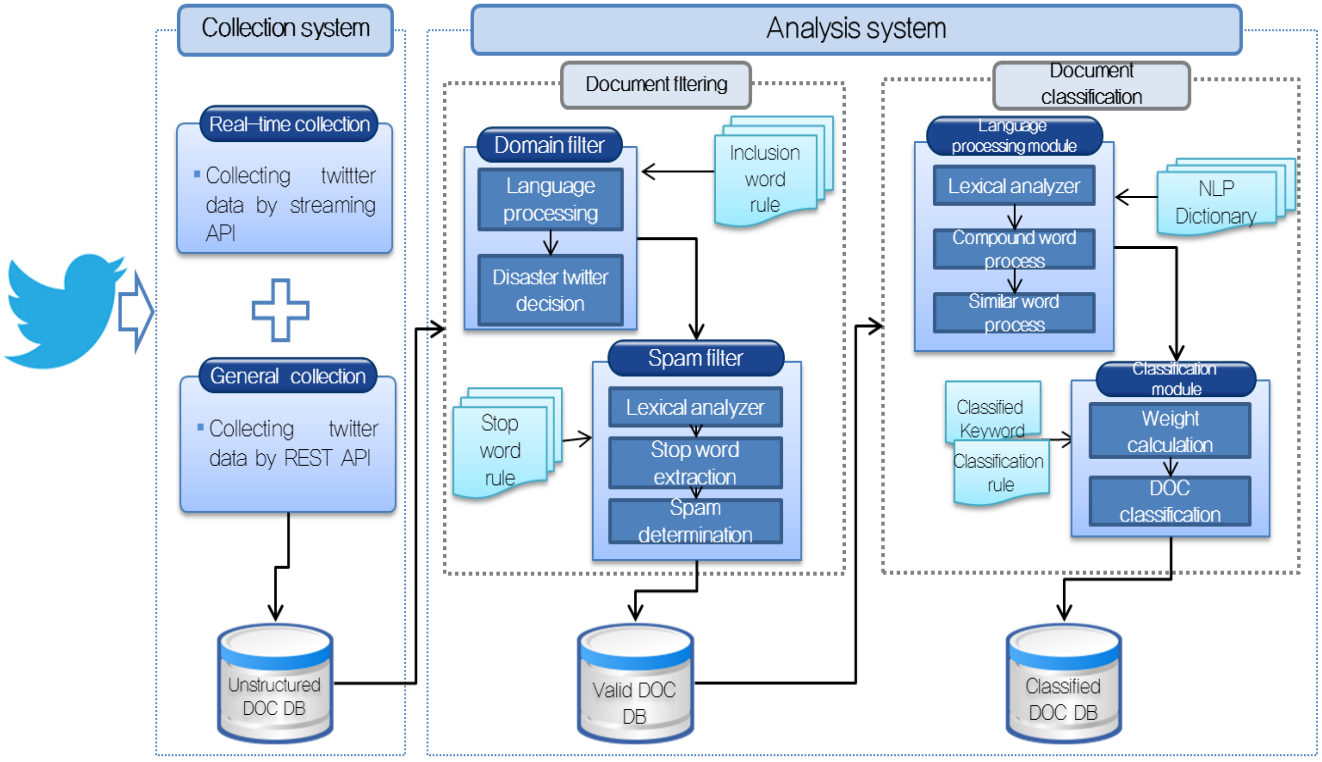

Fig. 6 Procedures of collection and analysis system

ISBN: 978-960-474-383-4