Typhoon Soudelor: Comparing remotely-sensed data and Tweets

Paul Haimes  
Tokyo Metropolitan University  
6-6 Asahigaoka, Hino, Tokyo Japan  
haimes@tmu.ac.jp

Tetsuaki Baba  
Tokyo Metropolitan University  
6-6 Asahigaoka, Hino, Tokyo Japan  
baba@tmu.ac.jp

Abstract
Typhoon Soudelor was the strongest typhoon recorded in 2015. It had a devastating impact on urban and regional areas of countries such as Taiwan, Japan and mainland China. To ascertain the effectiveness of online communication during this typhoon event, we are comparing near real-time remotely-sensed data with Tweets that are publicly available through Twitter’s platform. For the purpose of this study, we are looking at data points immediately before, during and after the point that Typhoon Soudelor crossed the area of the eastern Taiwanese city and region of Hualien. To do this, we are focusing on remotely-sensed readings and analysing content from Twitter from the same time period. Here, we consider and contrast these two data sources and discuss the strengths and weaknesses of both kinds of data as we look to accommodate both in an intuitive and easy to use mobile application for near real-time typhoon tracking.

Keywords: Crowdsourcing; Twitter; GIS; Disaster Management

1 Introduction
Typhoon Soudelor was the largest typhoon of 2015: the only super typhoon of the year. Affecting several territories including Taiwan, China and Japan, over a period of 10 days it left a path of destruction that resulted in the death of several people, dozens injured and hundreds of buildings destroyed [1, 2]. Like several recent disaster events (e.g., [3]), people in affected areas turned to Twitter in order to share information regarding the typhoon. Here, we provide examples of useful, localised information from the Typhoon Soudelor event as it made landfall in Taiwan, and discuss the benefits and challenges of integrating such data into a platform that relies primarily on remotely-sensed typhoon data from satellites.

2 Related work
Although there are several examples of systems that facilitate community-based crowdsourcing during and after emergencies (e.g. Verily [4]), we have found only one example in the Asia-Pacific region supported by a government authority that incorporates both official, scientifically-detected sources of data with crowd-sourced data in a disaster information scenario. GeoNet uses custom-built functionality to allow users to report earthquake tremors [5]. Users of the GeoNet application can provide information to the system and scientists can use crowdsourced data to verify their remotely-sensed data [5]. In Australia, another study compared near real-time remotely-sensed bushfire data with data sourced from Twitter [6]. Using data from a specific bushfire event, this study found that when compared to the remotely-sensed data, information sourced from Twitter was more timely, and provided more relevant, localised information [6]. We have adopted a similar approach here regarding Typhoon Soudelor. Crowdsourced data is something that emergency authorities need to consider accommodating [7].

3 Near real-time typhoon data
We are using a data feed provided to us by Japan’s National Institute of Informatics (NII). Their data feed covers much of the Pacific, including areas affected by Typhoon Soudelor such as Saipan, Taiwan and mainland China. The data — sourced from geostationary satellites — includes coordinates, category, time, wind speed and pressure of a typhoon.

We are making a prototype application (Fig. 1) that runs on iOS devices (iPhone, iPod Touch and iPad). We are developing the application in Swift [8] and using the data provided to us from NII. We are in the process of embedding Tweets in our application. For the purpose of this study, we are focusing on three readings in the data feed where the typhoon approached landfall near the Hualien region in east-central Taiwan. The first reading was taken prior to the typhoon crossing the coast. The second reading was when the typhoon had crossed into the Hualien region and the third is to the west of Hualien. These three readings are shown in Figure 1 and cover a time period of approximately six hours.

4 Twitter-sourced data
We manually searched Tweets using Twitter’s search functionality from 7th August 2015 where Soudelor approached Hualien and made landfall. Only Tweets referring to specific locales from individuals or small organisations (rather than official news agencies or emergency authorities) were of interest to us.
These tweets needed to refer to specific locales in eastern Taiwan and contain more detailed information than the remotely-sensed data described in the previous section. Here we have provided five examples from the search results that met our criteria:

- “Winds really ripped through Taroko just now, large piece of debris blocking road southbound highway 9 #typhoon #Soudelor” [9]
- “The lights began to die. it’s 5:05 am. #Taiwan #Soudelor Taichung City” [10]
- “This thing is ripping through Taroko Gorge. Worst was about 30 minutes ago. Stay safe Taiwan #Soudelor” [11]
- “Typhoon #Soudel has made landfall in the mountainous region north of Taroko, Taiwan... Taipei with very strong winds and torrential rains” [12]
- “Waked [sic] up by powerful 75 76 knots gust caused by #Soudelor at suburban Taipei... #Taiwan” [13].

5 Results

The Tweets we analysed contained information specific to locales such as Taroko National Park, Taipei and Taichung. Yet it was only through manually searching that we were able to find these Tweets related to specific areas. Although there were thousands of Tweets from the time period of the readings in Fig 1 (using the tag #Soudelor), the majority were not geo-tagged, and were too general to provide any meaningful information. The majority of Tweets were simply “retweeting” or posting information from third party news sources rather than providing specific, localised information. Much of this data would not be useful to our application. However, though the information in the examples above was not verifiable at the time, these Tweets clearly provide information that would be useful for people around the areas described, and would be useful in addition to remotely-sensed data.

6 Discussion and conclusion

Based on the two datasets, we found that while crowdsourced data clearly has the potential to provide rich information, much work is needed if data from Twitter is to be integrated into a system that wishes to remain an authoritative and credible source of information. Yet due to potentially rich data, crowdsourcing is something that agencies providing emergency information should learn to accommodate. A customised system, with various ways of verifying data, may be one way of doing this, rather than drawing directly on data from Twitter. However, a system capable of doing this in real-time disaster situations would require a strict verification system — through administrators or trusted users verifying or adding data manually — to ensure that only meaningful, useful data is added to the application.

7 Acknowledgements

This work is supported by a Kakenhi research grant (15F15052) from the Japan Society for the Promotion of Science.

8 References


Figure 1: Screenshot of our prototype, showing three points where Typhoon Soudelor was remotely-sensed as it approached and crossed into eastern Taiwan. Red icons indicate that the typhoon was category 5 (the highest category). Note that the green area above Hualien is Taroko National Park, which is referenced in several Tweets.