Geophysical Research Abstracts
Vol. 19, EGU2017-1102, 2017
EGU General Assembly 2017

Towards a global flood detection system using social media

Jens de Bruijn (1), Hans de Moel (1), Brenden Jongman (1,2), and Jeroen Aerts (1)

(1) IVM, Vrije Universiteit, Amsterdam, The Netherlands, (2) GFDRR, World Bank, Washington D.C., United States of America

It is widely recognized that an early warning is critical in improving international disaster response. Analysis of social media in real-time can provide valuable information about an event or help to detect unexpected events. For successful and reliable detection systems that work globally, it is important that sufficient data is available and that the algorithm works both in data-rich and data-poor environments. In this study, both a new geotagging system and multi-level event detection system for flood hazards was developed using Twitter data.

Geotagging algorithms that regard one tweet as a single document are well-studied. However, no algorithms exist that combine several sequential tweets mentioning keywords regarding a specific event type. Within the time frame of an event, multiple users use event related keywords that refer to the same place name. This notion allows us to treat several sequential tweets posted in the last 24 hours as one document. For all these tweets, we collect a series of spatial indicators given in the tweet metadata and extract additional topological indicators from the text. Using these indicators, we can reduce ambiguity and thus better estimate what locations are tweeted about.

Using these localized tweets, Bayesian change-point analysis is used to find significant increases of tweets mentioning countries, provinces or towns. In data-poor environments detection of events on a country level is possible, while in other, data-rich, environments detection on a city level is achieved. Additionally, on a city-level we analyse the spatial dependence of mentioned places. If multiple places within a limited spatial extent are mentioned, detection confidence increases.

We run the algorithm using 2 years of Twitter data with flood related keywords in 13 major languages and validate against a flood event database. We find that the geotagging algorithm yields significantly more data than previously developed algorithms and successfully deals with ambiguous place names. In addition, we show that our detection system can both quickly and reliably detect floods, even in countries where data is scarce, while achieving high detail in countries where more data is available.