Geophysical Research Abstracts
Vol. 20, EGU2018-404, 2018
EGU General Assembly 2018

Big Data of Urban Flooding: Dance with Social Media, Citizen Science, and Artificial Intelligence

Ruo-Qian Wang
University of Dundee, Civil Engineering, United Kingdom (r.u.wang@dundee.ac.uk)

This project is aimed at creating a new platform to live monitor urban floods with a high resolution in the scale of parcels and streets. To our knowledge, this project is the first to introduce computer vision to recognise crowd-sourcing flood photos.

Urban flooding is the most costly natural hazard in the world. One of the urgent issues of increasing the resilience and preparedness to urban flooding is the difficulty to collect sufficient data to identify, characterise and predict the variables that influence the occurrence, impact, severity and duration of the
flood event. This problem, especially the lack of high-resolution data, prevents a comprehensive scientific understanding of the flooding risk and an informed
decision-making of the emergency response. The traditional data collection method has difficulty to address the data demand at this high-resolution scale. For example, the updating frequency of remote sensing is limited by the orbit and period of satellites and the availability of aeroplanes that take aerial photos. In addition, sensor networks are still expensive to build and maintain in focused areas. Interviews of residents and damage reports by insurance companies are still the best available data collection methods, but such collected information can be incomplete and misleading in location and time due to the inaccuracy of witnesses' memories.

Data mining of social media and crowd-sourced data can potentially address this issue. In collaboration with mycoast.org, a mobile app company, we gathered thousands of photos from the general public. These photos are labelled with time and location information. We employed a computer-vision program and were able to determine whether a photo captures flooding scene. In addition, Natural Language Processing tool was used to extract flooding descriptions from Twitter, news reports, and traffic information.

The present study shows that social media and crowdsourcing can be used to complement the datasets developed based on traditional remote sensing and witness reports. Applying these methods in two case studies, we found these methods are generally informative in flood monitoring. Twitter data is found weakly correlated to precipitation departure. We determined a length scale of tweet volume pattern, at which the data points are most clustered. The computer vision processed crowdsourcing data is compared against the road closure data. The results show that computer vision still has room to improve, especially in coastal areas. These two methods are compared and a series of recommendation is given to improve the big data based flood monitoring in the future.