Mining social media can help improve disaster response efforts

January 20 2016

Leveraging publicly available social media posts could help disaster response agencies quickly identify impacted areas in need of assistance, according to a Penn State-led team of researchers. By analyzing the September 2013 Colorado floods, researchers showed that a combination of remote sensing, Twitter and Flickr data could be used to identify flooded areas.

"FEMA (the Federal Emergency Management Agency), the Red Cross and other response agencies use social media now to disseminate relevant information to the general public," said said Guido Cervone, associate professor of geography and associate director of the Penn State's Institute for CyberScience. "We have seen here that there is potential to use social media data from community members to help identify hotspots in need of aid, especially when it is paired with remote sensing imagery of the area."

After a disaster, response teams typically prioritize rescue and aid efforts with help from imagery and other data that show what regions are affected the most. Responders commonly use satellite imagery, but this on its own has drawbacks.

"Publicly available satellite imagery for a location isn't always available in a timely manner—sometimes it can take days before it becomes available," said Elena Sava, graduate student in geography, Penn State. "Our research focused on identifying data in non-traditional data streams that can prove mission critical for specific areas where there might be

damage. We wanted to see if social media could help filling the gaps in the satellite data."

The 2013 Colorado flooding was an unprecedented event. In nine days in September, Boulder, Colo., received more than 43 centimeters, or 17 inches of rain—nearly the amount of rainfall it normally receives in a year. Officials evacuated more than 10,000 people and had to rescue several thousand people and pets.

Because the flooding occurred in an urban setting, the researchers were able to access more than 150,000 tweets from people affected by the flooding. Using a tool called CarbonScanner, they identified clusters of posts suggesting possible locations of damage. Then, they analyzed more than 22,000 photos from the area obtained through satellites, Twitter, Flickr, the Civil Air Patrol, unmanned aerial vehicles and other sources.

Responders need information in real time during disasters, so the researchers developed an innovative approach to collect and analyze images from numerous sources in near real time. They developed a machine-learning algorithm to automatically analyze several thousand images, which allowed them to quickly identify individual pixels of images that contained water.

"We looked at a set of images and manually selected areas that we knew had water and areas that had no water," said Sava. "Then, we fed that information to the algorithm we had developed, and it allowed the computer to 'learn' what was and wasn't water."

The team's findings, published in the current issue of the *International Journal of Remote Sensing*, confirmed that Twitter data could serve to identify hotspots for which satellite imagery should be acquired.

The team also found that satellite imagery on its own was not always

reliable, and that social media can be fused with [remote sensing](#) imagery to help identify the extent of the flooding.

"If you look at [satellite imagery](#), downtown Boulder showed very little flooding," said Cervone. "However, by analyzing Flickr and Twitter data, we could find several cues that many areas were underwater. People would mention a 'river in the road' or mention how they can't get to their house because of the water. Those posts are indicators that we cannot rely on satellite data alone, and we must integrate [satellite data](#) with additional data from [unmanned aerial vehicles](#), airplanes and social media to identify the impacted areas."

The team is now fine-tuning their algorithms and assessing whether they can be applied to future flooding events. They are also investigating whether other [social media](#) platforms, such as Facebook or Instagram, could provide useful data to emergency responders. Collaborators on this study included Emily Schnebele and Nigel Waters, Penn State; Qunying Huang, University of Wisconsin, Madison; and Jeff Harrison, The Carbon Project. The team performed their analysis with supercomputers at the National Center for Atmospheric Research in Boulder, Colo.

Provided by Pennsylvania State University