

# Improving disaster response through Twitter data

July 11 2018, by Jordan Ford

---



“We are looking at the crisis as it happens,” said Prasenjit Mitra. “The best source to get timely information during a disaster is social media, particularly microblogs like Twitter. Credit: Thinkstock

Twitter data could give disaster relief teams real-time information to

provide aid and save lives, thanks to a new algorithm developed by an international team of researchers.

A team of researchers from Penn State, the Indian Institute of Technology Kharagpur, and the Qatar Computing Research Institute created an algorithm that analyzes Twitter data to identify smaller disaster-related events, known as sub-events, and generate highly accurate, real-time summaries that can be used to guide response activities.

The group presented their paper—"Identifying Sub-events and Summarizing Information from Microblogs during Disasters"—today (July 10) at the 41st International Association for Computing Machinery's Special Interest Group on Information Retrieval Conference on Research and Development in Information Retrieval in Ann Arbor, Michigan.

"We are looking at the crisis as it happens," said Prasenjit Mitra, associate dean for research in Penn State's College of Information Sciences and Technology and a contributor to the study.

"The best source to get timely [information](#) during a disaster is social media, particularly microblogs like Twitter," said Mitra. "Newspapers have yet to print and blogs have yet to publish, so Twitter allows for a near real-time view of an event from those impacted by it."

Analyzing this data and using it to generate reports related to a sub-topic of a disaster—such as infrastructure damage or shelter needs—could help humanitarian organizations better respond to the varying needs of individuals in an affected area.

Given the volume of data produced, manually managing this process in the immediate aftermath of a crisis is not always practical. There is also

often a need for unique updates related to particular topics within and across organizations.

"Several works on disaster-specific summarization in recent times proposed algorithms that mostly provide a general summary of the whole event," the researchers wrote in their paper. "However, different stakeholders like rescue workers, government agencies, field experts, [and] common people have different information needs."

In the study, the group collected more than 2.5 million [tweets](#) posted during three major global catastrophes—Typhoon Hagupit that hit the Philippines in 2014, the 2014 flood in Pakistan, and the 2015 earthquake in Nepal. Then, volunteers from the United Nations Office for the Coordination of Humanitarian Affairs trained a machine learning system by manually categorizing the tweets into different sub-events, such as food, medicine and infrastructure.

Once the system can identify tweets with a high level of accuracy, the researchers allow the system to categorize large amounts of data quickly and accurately without human intervention. As events develop, however, new categories of content appear that require the process to restart.

"At a certain point, there is a drift in topic. Topics shift from immediate response, such as people are trapped, to ongoing fallout, such as diseases or transportation issues," explained Mitra. "When the topic changes, we observe the machine's accuracy. If it falls below a certain threshold, the task force manually categorizes more tweets to further educate the machine."

Their "Dependency-Parser-based SUB-event detection" algorithm, known as DEPSUB, identified noun-verb pairs representing sub-topics—such as "bridge collapse" or "person trapped"—and ranked them based on how frequently they appear in tweets. Then, they created an

algorithm to write summaries on the broad event and the identified sub-events. Finally, human evaluators ranked the usefulness and accuracy of sub-events identified by DEPSUB and auto-generated summaries against those created by other existing methods.

The evaluators found both DEPSUB and their summary algorithm to be more relevant, useful and understandable compared to other leading algorithms. In the future, the researchers hope to apply their work to specialized situations, such as summarizing information on missing people, and pulling specific information from tweets that could create a more thorough description and visualization of an event.

"With a well-trained system, [human intervention](#) is not needed to categorize or summarize Twitter data," said Mitra. "This automated system is a first step in giving aid workers a scaffolding that they can refine to build a better overall summary of an event, as well as taking a more narrowly tailored view of some part of that larger event."

**More information:** Identifying Sub-events and Summarizing Information from Microblogs during Disasters.

[www.cnergres.iitkgp.ac.in/sube ... arizer/dataset.html#](http://www.cnergres.iitkgp.ac.in/sube...arizer/dataset.html#)

Provided by Pennsylvania State University

Citation: Improving disaster response through Twitter data (2018, July 11) retrieved 9 April 2024 from <https://phys.org/news/2018-07-disaster-response-twitter.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
---