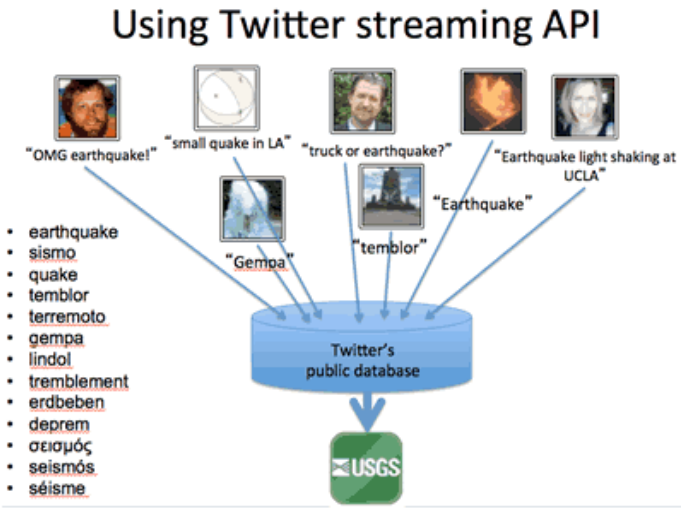
Tweet Earthquake Dispatch: Crowdsourcing Earthquake Detection

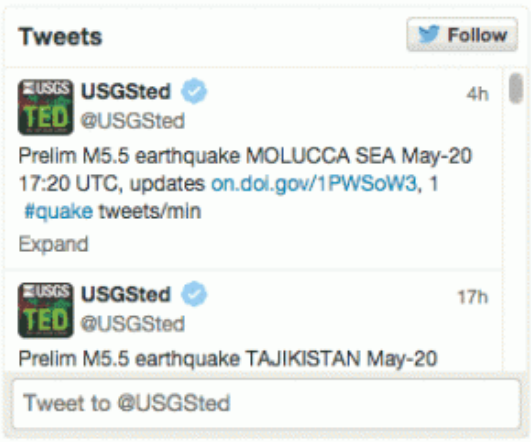
Case Study Overview



Within seconds of an earthquake, many people report what they felt through social networks such as Twitter. By contrast, scientific alerts can take up to 20 minutes, depending on the size and location of the earthquake.

Tweet Earthquake Dispatch uses social networks to support earthquake response and report information about related hazards. Through the *TED* system, the U.S. Geological Survey's National Earthquake Information Center analyzes data from social networks to detect earthquakes within seconds of their occurrence. It also uses *TED* to assess citizen reports of earthquakes alongside corresponding scientific reports and to tailor its own social networking about earthquakes to the corresponding level of public interest.

Project Description



TED harvests real-time tweets through a continuous connection to Twitter. The system applies a query parameter to reduce the stream to tweets that contain the keyword earthquake in several languages.

The system also applies other data-cleaning techniques. It merges tweets, ordering them, accounting for duplicates and filling any data gaps. Data from aggregators — users who redistribute secondhand earthquake information — are removed from the dataset.

For each tweet filtered by keyword, *TED* archives the creation time and text, the Twitter user location, the Twitter tweet ID, and the time the tweet appeared in the *TED* database. The system also uses the Yahoo Maps API Geocoding Service to estimate the latitude and longitude of the sender's location, if provided.

Around the clock, *TED* also ingests seismically derived earthquake information from the USGS's near-real-time internal global earthquake stream. *TED* archives the earthquake time, region, magnitude and hypocenter (latitude, longitude and depth). It also records the source of the scientifically derived earthquake information.

For earthquakes that have been verified using seismic instruments, *TED* subsequently sends out a tweet to followers with basic information about the earthquake.

Challenges

TED supports the mission of the USGS National Earthquake Information Center by providing rapid information about the occurrence of an earthquake. This information augments but does not replace the instrumental data collected by the USGS. *TED* provides no quantitative information about earthquake epicenter and magnitude, for example, and the geographic location it gives is not always reliable. Without verified quantitative data, the National Earthquake Information Center can neither issue public alerts nor trigger direct response measures.

Benefits and Outcomes

The main advantage of mining citizen reports through Twitter is speed. Rapid tweet-based earthquake detection can potentially fill the gap between the time when an earthquake occurs and the time when seismically derived information becomes available.

TED detects two to three earthquakes a day, on average. Especially in regions with few seismometers, *TED* reports often come in before traditional seismic networks detect an earthquake, giving seismologists early warning. *TED* sometimes detects earthquakes entirely missed by USGS's automatic processing system, thereby increasing the number of felt events known to the agency. In addition, the tweet text and attached images sometimes offer a rapid qualitative assessment of an earthquake's impact.

Tips

The *Tweet Earthquake Dispatch* case study illustrates the following step in the Federal Citizen Science and Crowdsourcing Toolkit:

- **Scope Out Your Problem — Know Where Your Project Fits**
  This project complements *Did You Feel It?,* delivering faster information about an earthquake occurrence and its general area. By contrast, *Did You Feel It?* produces quantitative maps showing the distribution of shaking.

Learn More

- Website: *Tweet Earthquake Dispatch*

Contact Information

Name: Paul Earle
Email: pearle@usgs.gov