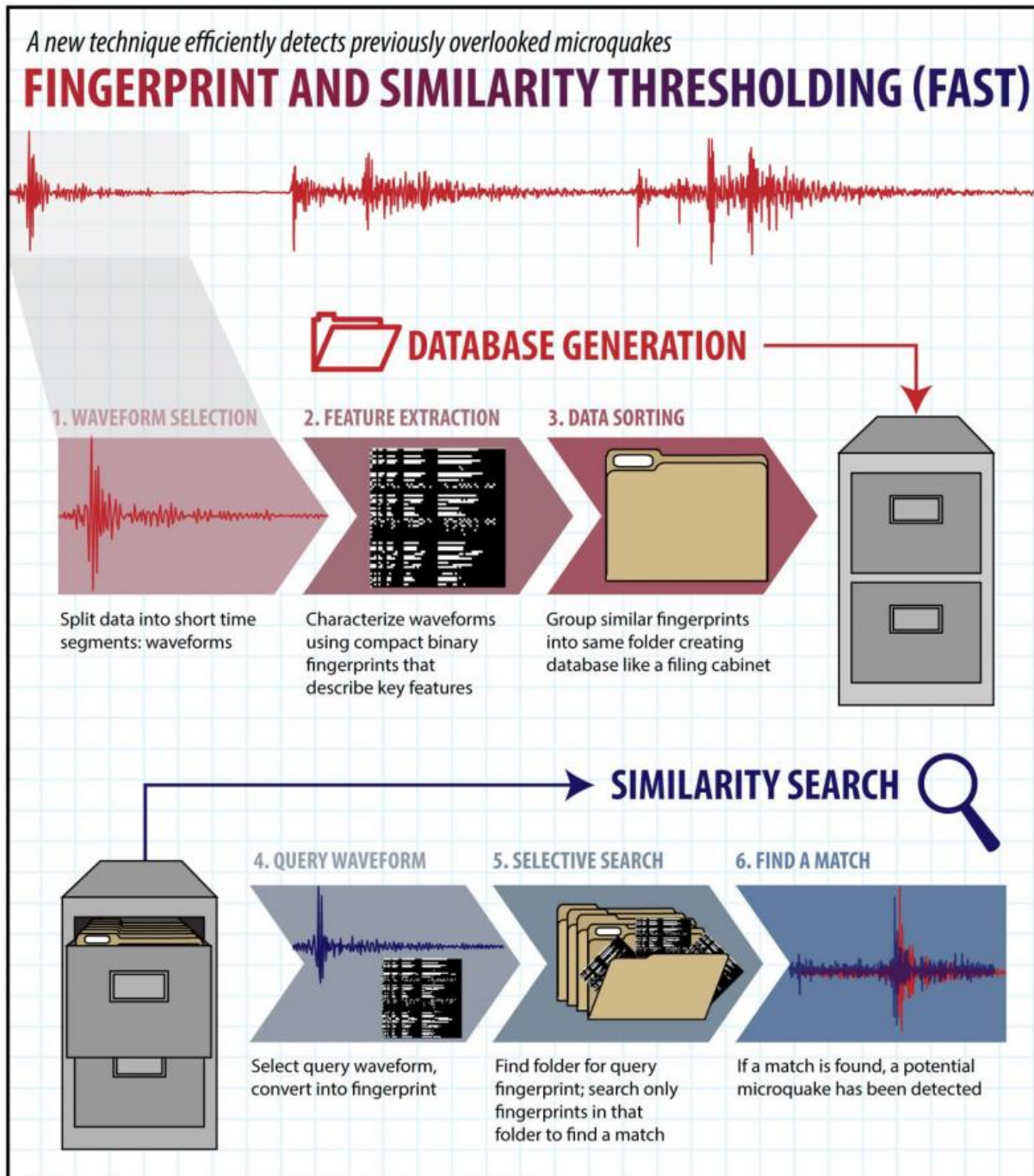


# Scientists develop 'Shazam for earthquakes'

December 4 2015

---



An illustration of how FAST works. Credit: Ilissa Ocko

An algorithm inspired by a popular song-matching app is helping Stanford scientists find previously overlooked earthquakes in large databases of ground motion measurements.

They call their algorithm Fingerprint And Similarity Thresholding, or FAST, and it could transform how [seismologists](#) detect microquakes - temblors that don't pack enough punch to register as earthquakes when analyzed by conventional methods. While microquakes don't threaten buildings or people, monitoring them could help scientists predict how frequently, and where, larger quakes are likely to occur.

"In the past decade or so, one of the major trends in seismology has been the use of waveform similarity to find weakly recorded earthquakes," said Greg Beroza, a professor of geophysics at Stanford School of Earth, Energy & Environmental Sciences.

The technique most commonly employed to do this, called template matching, functions by comparing an earthquake's seismic wave pattern against previously recorded wave signatures in a database. The downsides of template matching are that it can be time-consuming and that it requires seismologists to have a clear idea of the signal they are looking for ahead of time.

The FAST technique, which is detailed in the current issue of the journal *Science Advances*, circumvents both of these shortcomings by taking all of the recorded data from a seismic station and chopping the continuous signal into segments of a few seconds each. The signals are then compressed into compact representations, or "fingerprints," for rapid processing.

The fingerprints are then sorted into separate bins, or groups, based on

their similarities.

"We then search for pairs of fingerprints that are similar, and then map those back to the time windows that they came from," said study co-author Clara Yoon, a graduate student in Beroza's research group.

"That's how we identify the earthquakes."

Earthquakes occurring on the same section of a fault have similar fingerprints, regardless of their magnitudes, because the seismic waves they generate travel through the same underground structures to reach the surface.

"It doesn't matter if one earthquake happened 10 years ago and the other one happened yesterday. They're actually going to have waveforms that look very similar," Yoon said.

This sorting step, which the Stanford scientists compare to grouping similar documents in a filing cabinet, is why FAST is so efficient.

"Instead of comparing a signal to every other signal in the database, most of which are noise and not associated with any earthquakes at all, FAST compares like with like," said Beroza, who is the Wayne Loel Professor at Stanford. "Tests we have done on a 6-month data set show that FAST finds matches about 3,000 times faster than conventional techniques. Larger data sets should show an even greater advantage."

## **Eureka moment**

The idea for FAST occurred to Beroza several years ago. While perusing an electronics store, he heard a catchy song he didn't know playing over the speakers. Beroza pulled out his smartphone and opened Shazam, an app that could listen to and identify the song by name.

"Shazam did its thing and within 10 seconds it was trying to sell me the song," Beroza said. In a moment of insight, Beroza realized that Shazam wasn't simply comparing the digital file of the song against other files in a database. It was doing something more sophisticated, namely capturing the audio waveform of a short section of the song and comparing that snippet to other waveforms housed on an online server. Not only that, the app had to be able to quickly filter out irrelevant noise from the environment such as people's conversations.

"I thought, 'That's cool,' and then a moment later, 'That's what I want to do with seismology,'" Beroza said.

It took several years, but Beroza eventually assembled a team of computer-savvy researchers to help build upon his eureka moment. Drawing heavily upon recent advances in computer science, the group created a search algorithm capable of quickly scanning continuous ground motion data for similar matches.

"In the early stages, we thought that we were going to need high-performance supercomputers to tackle the problem in a brute force fashion by running thousands of comparisons at once," said study co-author Ossian O'Reilly, a graduate student in the Department of Geophysics. "But we soon realized that even they wouldn't be able to handle the amount of data we wanted to process. So we started learning about the ingenious algorithms devised by the computer science community for solving related problems."

In particular, the team borrowed techniques from data mining and machine learning to create FAST, said study co-author Karianne Bergen, a graduate student at Stanford's Institute for Computational and Mathematical Engineering (ICME).

"The scalability of FAST comes from the use of a data mining technique

called locality-sensitive hashing, or LSH," Bergen said. "LSH is a widely used technique for identifying similar items in large data sets. FAST is the first use of LSH in [earthquake](#) detection."

## Testing FAST

In the new study, the Stanford scientists used FAST to analyze a week's worth of data collected in 2011 by a seismic station on the Calaveras Fault in California's Bay Area. This same fault recently ruptured and set off a sequence of hundreds of small quakes.

Not only did FAST detect the known earthquakes, it also discovered several dozen weak quakes that had previously been overlooked.

"A lot of the newer earthquakes that we found were magnitude 1 or below, so that tells us our technique is really sensitive," Yoon said. "FAST was able to spot the missed quakes because it looks for similar wave patterns across the seismic data, regardless of their energy level."

The team thinks FAST could prove useful in places like Oklahoma and Arkansas, which have recently experienced spikes in suspected induced earthquakes due to the increased injection of wastewater from oil and gas development into the subsurface. "If you can detect the smaller quakes, you could identify the risk of a larger quake occurring from continued injection," Yoon said.

Similarly, an improved understanding of how often different magnitude earthquakes happen could help seismologists better predict how frequently large, natural quakes will occur.

The team is currently working on scaling up their FAST algorithm to analyze data collected across longer periods of time, from multiple seismic stations and in more challenging scenarios.

"That is very important if you want to be able to determine the location of these earthquakes," Yoon said. "If you have data from only one station, you can't accurately pinpoint the epicenter of a quake."

**More information:** "Earthquake detection through computationally efficient similarity search," *Science Advances*,  
[advances.sciencemag.org/content/1/11/e1501057](https://advances.sciencemag.org/content/1/11/e1501057)

Provided by Stanford University

Citation: Scientists develop 'Shazam for earthquakes' (2015, December 4) retrieved 16 August 2024 from <https://phys.org/news/2015-12-scientists-shazam-earthquakes.html>

<p>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.</p>
--