A FAST Data-Mining Approach for Similar Earthquake Detection

Z. Spica[1], C. Yoon[1], G. Beroza[1], K. Bergen[2], K. Rong[3], H. Elezabi[3], P. Bailis[3], P. Levis[3]

[1]Department of Geophysics, Stanford University
[2]Institute for Computational and Mathematical Engineering, Stanford University
[3]Department of Computer Science, Stanford University

The Fingerprint and Similarity Thresholding (FAST) earthquake detection algorithm finds small earthquakes hidden in continuous seismic data through an uninformed search for similar signals at all times. FAST does not assume prior knowledge of template waveforms, nor does it use labeled earthquake waveform examples to recognize new earthquakes, which are examples of informed search. FAST enables detection of similar-waveform earthquakes in long-duration continuous data by adapting data-mining techniques originally developed for audio and image search within massive databases. FAST converts seismic waveforms into compact binary features called "fingerprints", designed to be discriminative: similar earthquake waveforms have highly similar fingerprints, while fingerprints extracted from noise have low similarity. FAST then uses locality-sensitive hashing to organize the fingerprints into a database, and efficiently search for similar fingerprints with high probability by avoiding unnecessary comparison of dissimilar fingerprints. FAST is now able to detect similar earthquakes over network of multiple seismic stations. Instead of associating phase arrivals from a single earthquake, FAST uses pair-wise pseudo-association, which exploits the fact that the relative arrival time difference between a pair of similar earthquakes is the same across all stations. FAST is now an open-source software package, which can be applied to detect similar-waveform earthquakes in any seismic network, with varying duration of continuous data and number of stations. First results to show its strong potential on volcano and subduction zones will be presented.