Joint probabilistic detection, association, & localization I: Hierarchical modeling

Nimar S. Arora, Michael I. Jordan, Stuart Russell, and Erik B. Sudderth
Computer Science Division, University of California, Berkeley

As part of its Comprehensive Test Ban Treaty (CTBT) verification efforts, the International Data Centre (IDC) analyzes seismic and other signals collected from hundreds of stations around the world. Current processing at the IDC proceeds, like many other large-scale monitoring and surveillance systems, in a series of pipelined stages. From station processing to network processing, each decision is made on the basis of local information. This has the advantage of efficiency, and simplifies the structure of software implementations. However, this approach may reduce accuracy in the detection and phase classification of arrivals, association of detections to hypothesized events, and localization of small-magnitude events.

In our work, we approach such detection and association problems as ones of probabilistic inference—that is, finding the most likely explanation given evidence from observed waveforms. Inference is applied to a generative probability model that describes events, signal propagation, and signal detection by sensors. In simple terms, let X be a random variable ranging over all possible collections of events, with each event defined by time, location, magnitude, and type (natural or man-made). Let Y range over all possible waveform signal recordings at all detection stations. Then $P_\theta(X)$ describes a parameterized generative prior over events, and $P_\phi(Y \mid X)$ describes how the signal is propagated and measured (including travel time, selective absorption and scattering, noise, artifacts, sensor bias, sensor failures, etc.). Given observed recordings $Y = y$, we are interested in the posterior $P(X \mid Y = y)$, and perhaps in the value of X that maximizes it—i.e., the most likely explanation for all the sensor readings. As detailed below, an additional focus of our work is to robustly learn appropriate model parameters θ and ϕ from historical data.

Calculating and maximizing the likelihood of events is a difficult inference problem; our approach, which is based on Markov chain Monte Carlo (MCMC), is described in a companion abstract. The primary advantage we expect is that decisions about arrivals, phase classifications, and associations are made with the benefit of all available evidence, not just the local signal or the detections associated with a single hypothesized event. Important phenomena—such as the successful detection of sub-threshold signals, correction of phase classifications using arrival information at other stations, and removal of false events based on the *absence* of signals—should all fall out of our probabilistic framework without the need for special processing rules.

In our baseline model, natural events occur according to a spatially inhomogeneous Poisson process, with intensities modelled by a mixture of Gaussians estimated from historical data. Complex events (swarms and aftershocks) may then be captured via temporally inhomogeneous extensions. Man-made events have a uniform probability of occurring anywhere on the earth, with a tendency to occur closer to the surface. Phases are modelled via their amplitude, frequency distribution, and origin. In the simplest case, transmission times are characterized via the one-dimensional IASPEI-91 model, accounting for model errors with Gaussian uncertainty. Such homogeneous, approximate physical models can be further refined via historical data and previously developed corrections. Signal measurements are captured by station-specific models, based on sensor types and geometries, local frequency absorption characteristics, and time-varying noise models.

We have previously developed a formal modelling language with the necessary expressive power to describe probabilistic models of this kind, along with general inference algorithms for all expressible models (Milch *et al.*, 2005). When learning model parameters, we leverage the rich statistical literature on hierarchical, probabilistic graphical models (Jordan, 2004). This approach allows locally estimated historical statistics to be globally calibrated, and can flexibly incorporate complex features or nonparametric representations to better capture large historical datasets.

At the conference, we expect to be able to quantitatively demonstrate the advantages of our approach, at least for simulated data. When reporting their findings, such systems can easily flag low-confidence events, unexplained arrivals, and ambiguous classifications to focus the efforts of expert analysts.

References

Jordan, M. I. (2004). Graphical models. *Statistical Science*, *19*(1), 140–155.

Milch, B., Marthi, B., Sontag, D., Russell, S., Ong, D., and Kolobov, A. (2005). BLOG: Probabilistic models with unknown objects. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI)*, Edinburgh.