Joint probabilistic detection, association, & localization II: MCMC Inference

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

Methods for automatically associating detected waveform features with hypothesized seismic events, and localizing those events, are a critical component of efforts to verify the Comprehensive Test Ban Treaty (CTBT). As outlined in our companion abstract, we have developed a hierarchical model which views detection, association, and localization as an integrated probabilistic inference problem. In this abstract, we provide more details on the *Markov chain Monte Carlo* (MCMC) methods used to solve this inference task.

MCMC (Gilks *et al.*, 1996) generates samples from a posterior distribution $\pi(x)$ over possible worlds x by defining a Markov chain whose states are the worlds x, and whose *stationary distribution* is $\pi(x)$. In the *Metropolis–Hastings* (M-H) method, transitions in the Markov chain are constructed in two steps. First, given the current state x, a candidate next state x' is generated from a *proposal distribution* $q(x' \mid x)$, which may be (more or less) arbitrary. Second, the transition to x' is not automatic, but occurs with an *acceptance probability* defined as follows:

$$\alpha(x' \mid x) = \min\left(1, \frac{\pi(x')q(x \mid x')}{\pi(x)q(x' \mid x)}\right)$$

It is not necessary that all the variables composing state x be updated simultaneously, in a single transition function. For example, *single-component* M-H algorithms, such as the *Gibbs sampler*, alter individual variables in turn. More broadly, domain knowledge can be used to factor $q(\cdot \mid \cdot)$ into separate transition functions for various strongly coupled subsets of variables. Under easily verifiable conditions guaranteeing the Markov chain's ergodicity (Gilks *et al.*, 1996), the M-H acceptance probability defined above ensures convergence of the Markov chain to $\pi(x)$, the target distribution of interest.

The seismic event model outlined in our companion abstract is quite similar to those used in multitarget tracking, for which MCMC has proved very effective—see, for example, (Pasula *et al.*, 1999; Oh *et al.*, 2009). In this model, each world x is defined by a collection of events, a list of properties characterizing those events (times, locations, magnitudes, and types), and the association of each event to a set of observed detections. The target distribution $\pi(x) = P(x \mid y)$, the posterior distribution over worlds x given the observed waveform data y at all stations. Proposal distributions then implement several types of *moves* between worlds. For example, *birth* moves create new events; *death* moves delete existing events; *split* moves partition the detections for an event into two new events; *merge* moves combine event pairs; *swap* moves modify the properties and assocations for pairs of events. Importantly, the rules for accepting such complex moves need not be hand-designed. Instead, they are automatically determined by the underlying probabilistic model, which is in turn calibrated via historical data and scientific knowledge.

Consider a small seismic event which generates weak signals at several different stations, which might independently be mistaken for noise. A birth move may nevertheless hypothesize an event jointly explaining these detections. If the corresponding waveform data then aligns with the seismological knowledge encoded in the probabilistic model, the event may be detected even though no single station observes it unambiguously. Alternatively, if a large outlier reading is produced at a single station, moves which instantiate a corresponding (false) event would be rejected because of the absence of plausible detections at other sensors.

More broadly, one of the main advantages of our MCMC approach is its consistent handling of the relative uncertainties in different information sources. By avoiding low-level thresholds, we expect to improve accuracy and robustness. At the conference, we will present results quantitatively validating our approach, using ground-truth associations and locations provided either by simulation or human analysts.

References

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (Eds.). (1996). *Markov chain Monte Carlo in practice*. Chapman and Hall, London.

Oh, S., Russell, S., and Sastry, S. S. (2009). Markov chain Monte Carlo data association for multi-target tracking. *IEEE Transactions on Automatic Control*, *54*(3), 481–497.

Pasula, H., Russell, S. J., Ostland, M., and Ritov, Y. (1999). Tracking many objects with many sensors. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, Stockholm. Morgan Kaufmann.