HMM with non-emitting states for Map Matching

Wannes Meert, Mathias Verbeke

ECDA 2018, Paderborn, Germany

Abstract

When collecting GPS traces, the GPS points are often inaccurate and require cleaning before further analysis can be performed. Therefore, a common preprocessing step is *map matching*, i.e. mapping the GPS points to a sequence of actual roads. Simply mapping the points to the nearest street, however, will return inaccurate results.

A popular approach is to use a Hidden Markov Model and Viterbi decoding, with the actual road segments as hidden states and the GPS points as observations. This approach was popularized by Microsoft (Newson and Krumm, 2009) and implemented in various services such as Mapbox's Map Matching API for OpenStreetMap.

A typical problem for HMMs is that for every state transition (e.g. moving from one street to another) an observation is 'absorbed'. In settings where measurements are far apart, thus less frequent than transitioning segments, this model fails. Two common proposed solutions are: interpolation, and interweaving Viterbi with a search algorithm. Neither is ideal. The first typically requires a high sample rate with more GPS points than segments and will generate too many observations. In the second, the HMM model is not used for all computations which makes it impossible to introduce a more sophisticated transition function.

We propose a new map matching approach using HMMs with non-emitting states, inspired by Profile HMMs in bioinformatics. These are states that are not associated with an observation. They thus allow for dynamic interpolation based on the route. This model, however, loses the property that the Viterbi algorithm is executed on a lattice bounded by the number of observations. To cope with this, we propose a smart pruning strategy such that only non-emitting states with the highest relevance are visited.

We evaluate on two applications, tracking vehicles and runners, and show how we can now learn the transition probabilities from history.