Modeling frameworks for map-matching or path-inference algorithm

Asif Rehan
University of Connecticut

1. Hidden Markov Model (HMM)

The application of HMM in sequential data is very popular as in the Map-matching and path-inference algorithms. Transition probabilities are based on assigned values derived from some continuous function or some discrete values. The assumption that only the previous state influences the next state makes it difficult for map-matching purpose to effectively capture the true location at one instant given some information from a state that is much distant from current state.

In the paper by Hunter et al *(1)*, another limitation of HMM, the selection bias, is indicated. This bias originally comes from the Label bias in Dynamic Bayesian Networks (DBN) framework, where the state transition probabilities are normalized separately for each state. For such state-specific normalizing or scaling process, the feasible state transitions will be overruled by the state with single feasible transition case. The more is discussed by Hunter et al. in their Path Inference Filter approach *(1)*.

In HMM, observed variables are assumed conditionally independent.

$$\boxed{p(\vec{y},\vec{x}) = \prod_{i=0}^{n} p(y_i|y_{i-1})p(x_i|y_i)\,.} \quad (2)$$

CRF addresses this problem.

Solution: HMM model would need not much training for map-matching application if some predefined functions are used as state and transition probabilities, e.g. Gaussian distribution for emission probability and exponential distribution for state transition probability.

For finding the maximum likelihood path, employ dynamic programming methods. Viterbi algorithm can offer the necessary recursive function that offers the most likely path given the observed data sequence of the GPS points. Viterbi algorithm can be used to obtain the most likely path in real-time. But the forward recursion cannot look back into the sequence to correct the maximum likely path sequence. One approach to solving this problem is proposed by Goh et al *(3)*. The authors use Variable-length Moving Window to look back into a fraction of immediately previous sequence of observations. But the method sometimes provide sub-optimal path.

For smoothing, forward-backward algorithm can be used. Following this algorithm, the most likely state can be obtained by Viterbi algorithm.

2. Maximum Entropy Markov Model (MEMM)

This is a Markov model with transition probabilities obtained by logistic regression. So the limitations of Markov model also apply in MEMM. As in HMM, MEMM also can only consider the immediate previous state to determine the current one. Thus the selection bias is also present in MEMM as in the case of all Dynamic Bayesian Network approach. But the improvement over HMM is through the logistic regression that can take effect of multiple non-independent factors or feature functions. In HMM, the observations need to be independent of each other given the states of the respective observations.

In the MEMM, probability of the sequence of the states, given the observation sequence is given by a logistic regression.

$$p_{\text{MEMM}}(\mathbf{y}|\mathbf{x}) = \prod_{t=1}^{T} p(y_t|y_{t-1}, \mathbf{x})$$
$$p(y_t|y_{t-1}, \mathbf{x}) = \frac{1}{Z_t(y_{t-1}, \mathbf{x})} \exp\left\{\sum_{k=1}^{K} \theta_k f_k(y_t, y_{t-1}, \mathbf{x}_t)\right\}$$
$$Z_t(y_{t-1}, \mathbf{x}) = \sum_{y'} \exp\left\{\sum_{k=1}^{K} \theta_k f_k(y', y_{t-1}, \mathbf{x}_t)\right\} \tag{4}$$

The normalizer, Z is state-specific. As a result label bias exists in MEMM.

Solution: In training MEMM, the weights θ are obtained by several methods, usually the gradient descent method can work when using supervised learning. After obtaining the weights, the optimal path sequence is determined using Viterbi algorithm.

3. Conditional Random Field (CRF)

This is a discriminative approach depending on the conditional probabilities. It finds the sequence using Maximum Entropy model. The state transitions can take information from any non-independent correlated feature functions, providing more freedom than in HMM. The label bias, or selection bias, in Markov models, or DBNs in general, is avoided by generalizing the normalizing factor over the possible sequences.

To get the intuition of the model, these equations can be explored. The probability of the sequence of the states, given the observation sequence is:

$$\boxed{p_{\vec{\lambda}}(\vec{y}|\vec{x}) = \frac{1}{Z_{\vec{\lambda}}(\vec{x})} \cdot \exp\left(\sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_i f_i\left(y_{j-1}, y_j, \vec{x}, j\right)\right) .}$$

The global organizer is different from that in MEMM, as can be seen below:

$$\boxed{Z_{\vec{\lambda}}(\vec{x}) = \sum_{\vec{y}\in\mathcal{Y}} \exp\left(\sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_i f_i\left(y_{j-1}, y_j, \vec{x}, j\right)\right) .} \tag{2}$$

PIF paper by Hunter at al. *(1)* uses the CRF framework for inferring the path. Implementation of CRF in map-matching and path inference is formulated in the paper.

Solution: It requires a forward-backward algorithm to obtain the message from one end to the other in an observed sequence (smoothing). For quicker output, only forward algorithm can be used at the expense of accuracy. For The partition function needs to be estimated only while training model. The most likely sequence can be obtained using Viterbi algorithm.

References

1. Hunter, T., P. Abbeel, and A. Bayen, The Path Inference Filter: Model-Based Low-Latency Map Matching of Probe Vehicle Data. In *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, No. 2, April 2014, pp. 507-529
2. Klinger, R. and K. Tomanek, Classical Probabilistic Models and Conditional Random Fields, *Algorithm Engineering Report*, TR07-2-013, ISSN 1864-4503, Faculty of Computer Science Algorithm Engineering (Ls11), 44221 Dortmund / Germany, December 2007
3. Goh, C.Y., J. Dauwels, N. Mitrovic, M. T. Asif, A. Oran, P. Jaillet, Online map-matching based on Hidden Markov model for real-time traffic sensing applications, In Proceedings of the 15th International IEEE Annual Conference on Intelligent Transportation Systems, ITSC, 2012.
4. Sutton, C., and A. McCallum, An Introduction to Conditional Random Fields, arXiv preprint, arXiv:1011.4088, November 2010