Inertial Hidden Markov Models: Modeling Change in Multivariate Time Series

George D. Montañez (Carnegie Mellon University) | Saeed Amizadeh (Yahoo Labs) | Nikolay Laptev (Yahoo Labs)

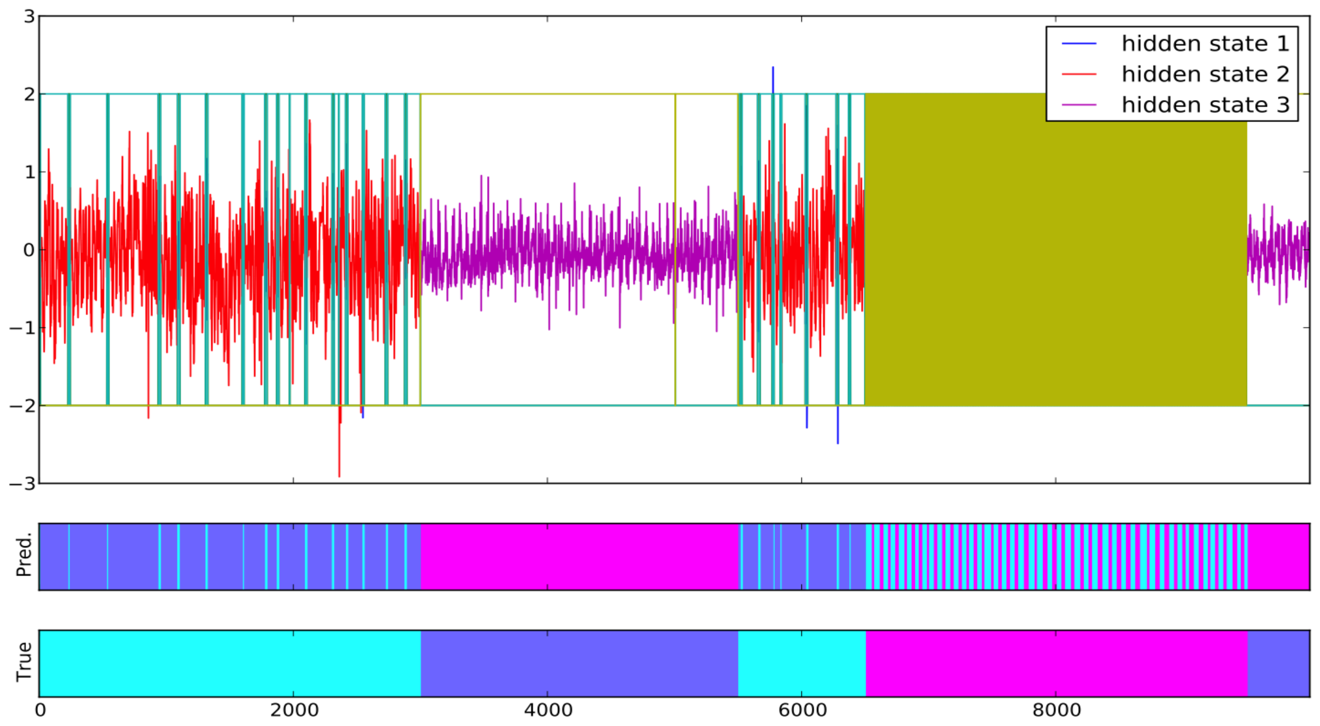
Motivation

Yahoo monitors millions of time series each day, looking for changes in the signals for advanced analytics.

Goal: Given a multivariate time series, find where systematic changes occur and map segments to a small number of states with *infrequent state transitions, aka state persistence*.

Hidden Markov Models

- Works well for segmenting sequential data.
- However, may over-segment leading to frequent state transitions.
- We need to impose state-persistence, i.e., few state changes over time.



Inertial HMMs

Two models which impose state persistence through a change to the likelihood model and corresponding expectation maximization (EM) update equations.

MAP Inertial HMM

- Include Dirichlet prior on transition matrix.
- Governed by a strength parameter, ς

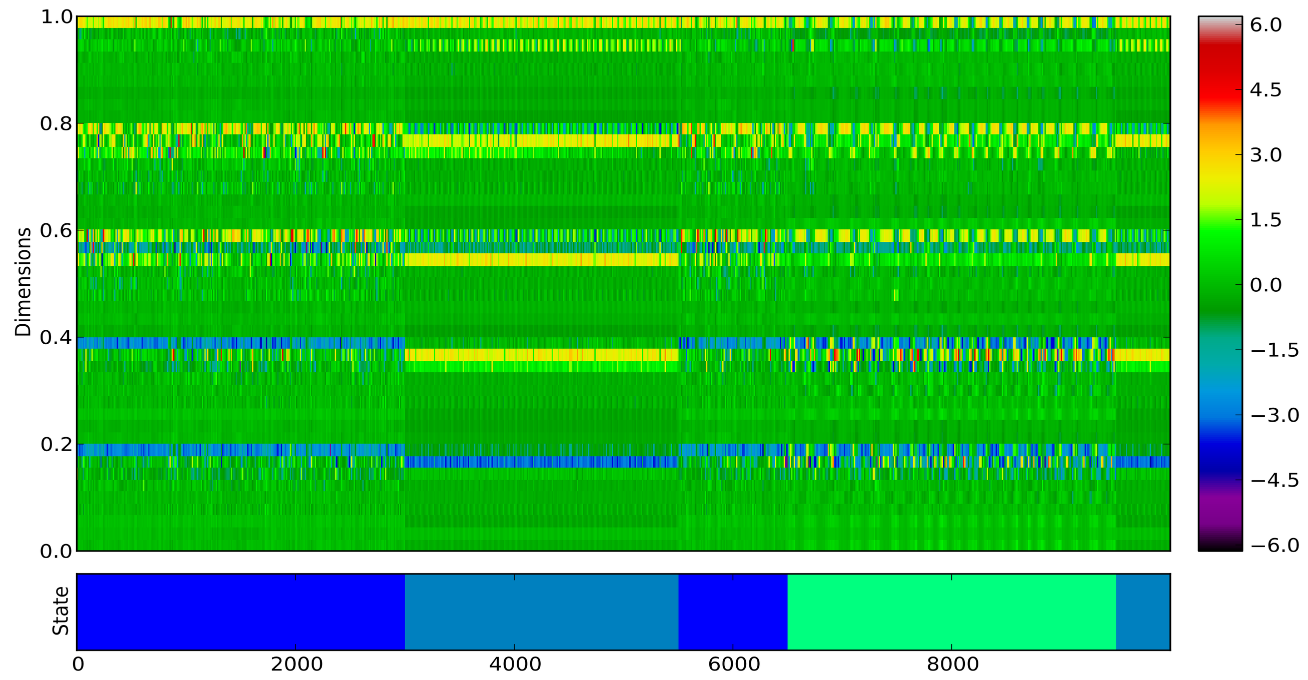
Pseudo-Observation Inertial HMM

- Alter the complete data-likelihood to include fictional self-transition observations.
- Governed by strength parameter, ς.

$$\mathbf{A}_{ij} \propto \begin{cases} ((T-1)^{\varsigma}-1)+B_T & i=j \\ B_T & i \neq j \end{cases}$$

Dataset

- 45D human activity accelerometer data
  - Activities included jumping, playing basketball, rowing, ascending stairs and walking.
  - Created 100 time series consisting on different combinations of activities and segmentations, 10K time steps each series.
  - Tested inertial methods performed vs standard HMM and Sticky HDP-HMM of Fox et al.



Evaluation Metrics

- Evaluated using:
  - Accuracy (for best permutation of labels)
  - Variation of Information
  - Number of Segments Difference (not shown)
  - Segment Number Ratio (not shown)

Sticky HDP-HMM

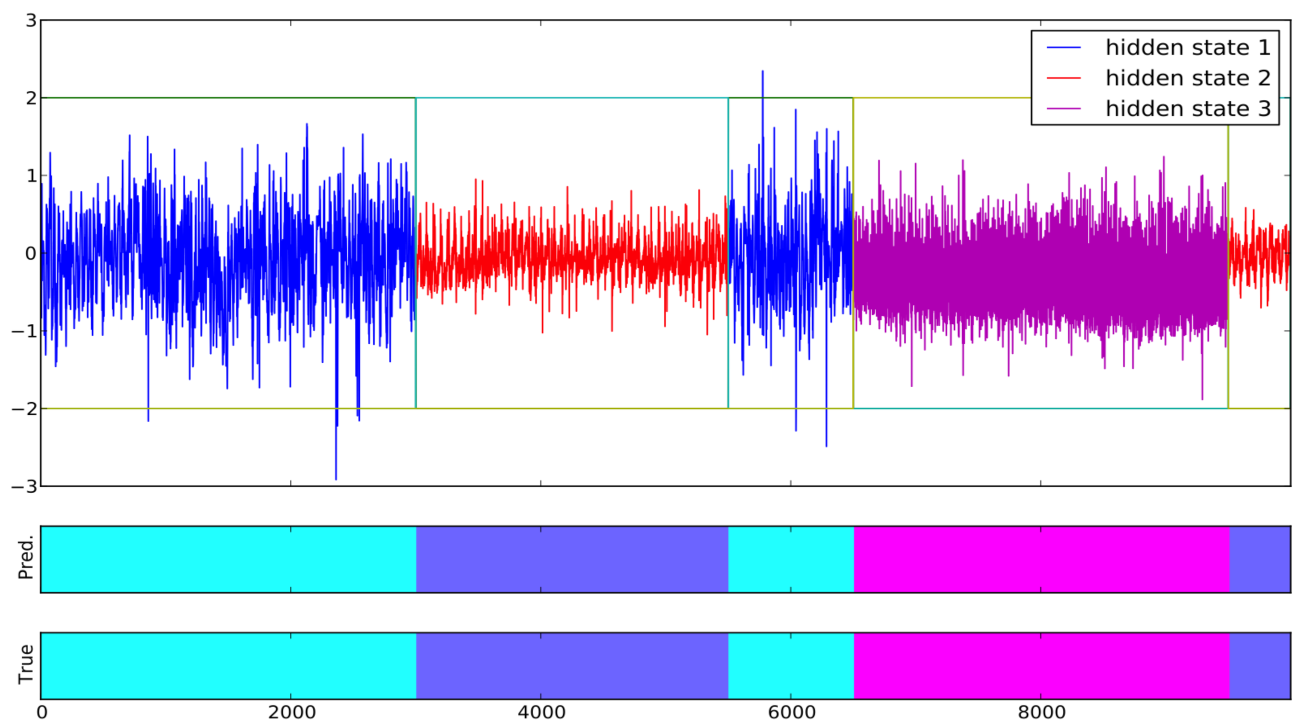
- State-of-the-art Bayesian hierarchical Dirichlet process hidden Markov model.
- Used publically available HDP-HMM toolbox, with default parameters for priors.
- κ – parameter, for "stickiness" of states.
- Truncation parameter *L* set to correct number of states.
- Sensitive to prior parameter values (see Results).

Results

Method	Accuracy	Var. of Info.
Standard HMM	0.79	0.38
Sticky HDP-HMM (κ = 100.0)	0.60	0.95
MAP Inertial HMM (ς = 33.5)	**0.94**	**0.14**
PsO Inertial HMM (ς = 49.0)	**0.94**	0.15

Results

MAP Inertial HMM – Example Segmentation



Extensions

- Automated parameter selection for inertial HMMs.
  - Used to select parameters in Results section.
- Online learning of inertial HMM model.

Main Advantages

- Works well on synthetic and real-world data.
- Very simple (change single update equation).
- Computationally efficient.
- Only two parameters, one automatically selected.
- Does not suffer from extreme sensitivity to parameter settings, as does sticky HDP-HMM.

Conclusion

- Simple modification of standard HMMs performs well on unsupervised segmentation task.
- Strongly outperforms state-of-the-art sticky HDP-HMM with default parameters.

References

- Emily B Fox, Erik B Sudderth, Michael I Jordan, Alan S Willsky, et al., *A sticky HDP-HMM with application to speaker diarization*, The Annals of Applied Statistics 5 (2011), no. 2A, 1020–1056.

YAHOO! LABS

ML MACHINE LEARNING DEPARTMENT Carnegie Mellon