Analysis of Temporal Clinical Patterns using Hidden Markov Models

CharmGil Hong and Milos Hauskrecht / Department of Computer Science, University of Pittsburgh

Intro & Motivation

- Increasing amount of clinical data in Electronic Health Record (EHR).
  - Modeling the dynamics of clinical data can be very complex.

Goals

- Building tractable temporal models of clinical data by finding a low-dimensional representation of patient states.
  - Represent patient case sequences using Hidden Markov Model (HMM).
  - Identify key patient states in a population of post-surgical cardiac patients.

Challenges

- Learning the dynamics of high-dimensional observation data in EHRs easily becomes intractable.
- Training HMMs needs to a priori decide how many hidden states will be used in the models.

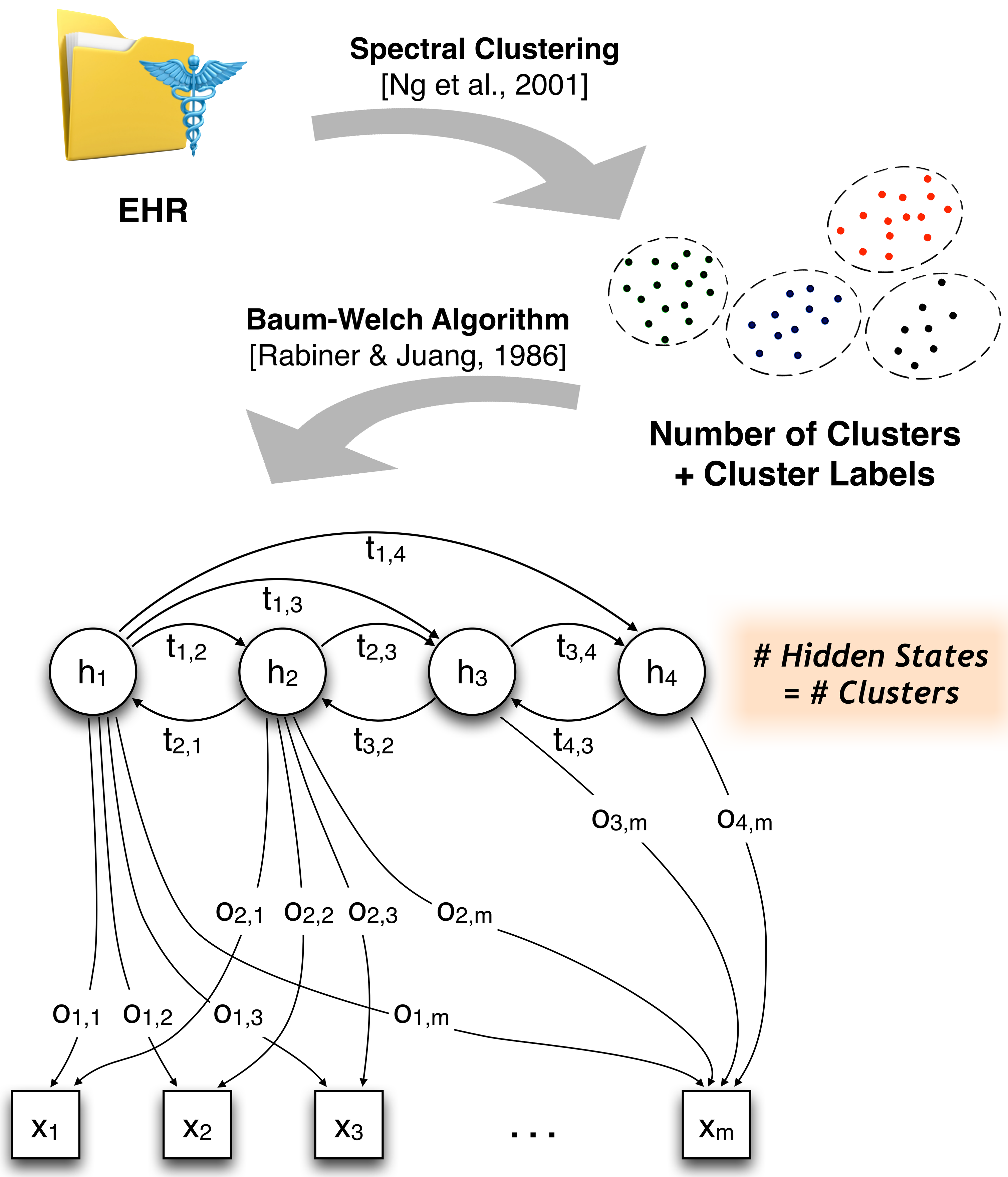
Appearing in the *ICML 2012 Workshop on Clinical Data Analysis*, Edinburgh, Scotland, UK, 2012.

Methods

Assumptions

- The observations defining the patient case at any point in time are independent of each other given the hidden state.
- The observations associated with the same hidden state are similar.

Procedure



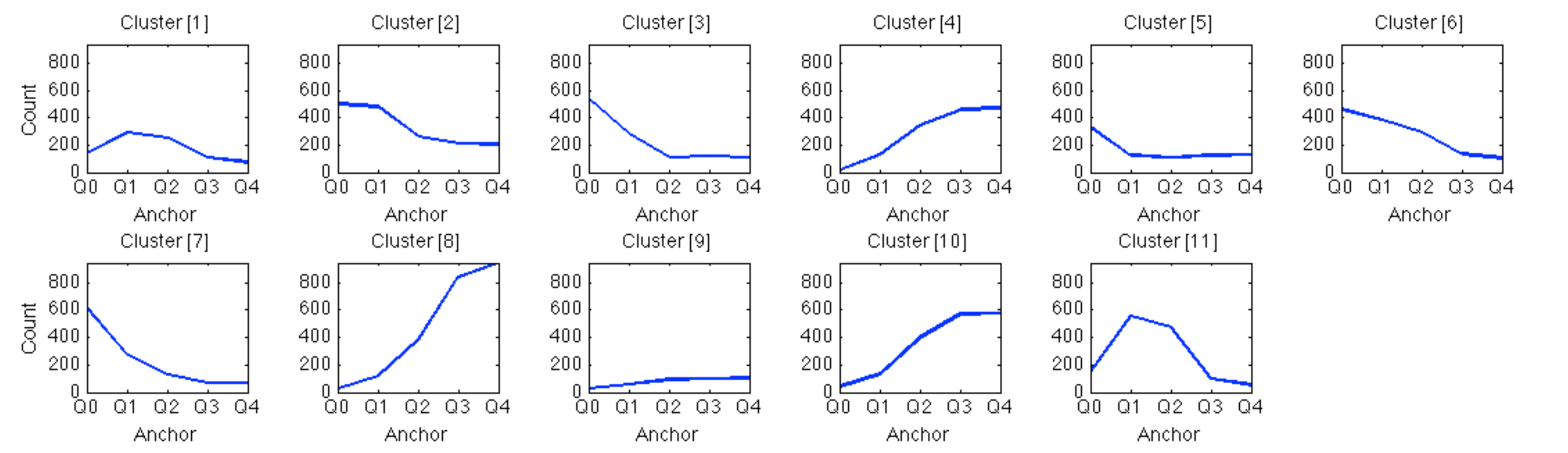
Hidden Markov Model

h_t: Hidden states
x_i: Observations
$t_{s,t}$: Transition prob. $p(h_t\,|\,h_s)$
$o_{t,i}$: Observation prob. $p(x_i\,|\,h_t)$

Experiments & Results

I. Analysis of Patient States using Spectral Clustering

- Goal: To identify and approximate important patient states.
- Data: Daily medication administration records that were given to 2,878 post-surgical cardiac patients (PCP) [Hauskrecht et al., 2010].
  - Binary vectors of 302 medications whether or not the corresponding medication was given to the patient during its time-window.
  - The length of patients' medication sequences ranged from 1 to 151 days.
  - Disregarding sequence structures, we have 30,828 binary medication administration vectors.
- Result: The support for clusters over five hospitalization periods (Q_i's).
  (Q_i denotes the 24-hr observation at the i/4 point of patient's hospitalization.)



II. Classification of a Surgery Type using HMMs

- Goal: To assess the modeling capability of our approach.
- Data: Daily medication records of 539 patients with prosthetic valve replacement surgery and 252 patients with tissue graft valve replacement surgery.
- Method:
  i) **Training**: For each of the 2 surgical groups, train an HMM with a training set.
  ii) **Classification**: Estimate the likelihood of each test sequence on both models. Pick the label of the model that maximizes the likelihood $p(x\,|\,M_k)$, where M_k is a model.
- Result: The confusion matrix for classification with hidden Markov Models.

		Predictions (in %)	
		Prosthetic Valve	Tissue Graft Valve
Labels	Prosthetic Valve	88%	12%
	Tissue Graft Valve	27%	73%