Dynamic Probabilistic Models for Latent Feature Propagation in Social Networks

Creighton Heaukulani, Zoubin Ghahramani

Reviewed by Zhao Song

March 21, 2014

Outline

Elements in Dynamic Network

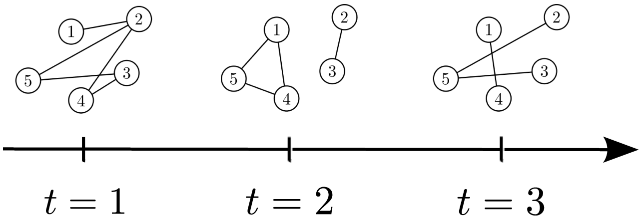


Figure 1 : A dynamic network shown in [Heaukulani 2013]

Link Adjacency Matrix $\boldsymbol{Y}^{(t)}$: $N \times N$ binary.

- $y_{ij}^{(t)} = 1$: There is a link between actor i and j at time t.
- A link can be interpreted as friendship or correspondence.
- $\boldsymbol{Y}^{(t)}$ is symmetric and its diagonal elements are meaningless.

Elements in Dynamic Network (Cont.)

$$\boldsymbol{H}^{(t)} =$$

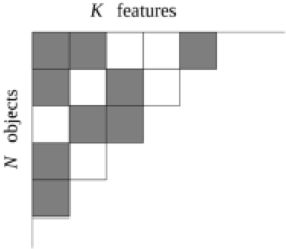


Figure 2 : Latent feature representation shown in [Heaukulani 2013]

Latent feature matrix $\boldsymbol{H}^{(t)}$: $N \times K$ binary.

- K features: For example, the feature k corresponds to “play tennis”.
- $h_{ik}^{(t)} = 1$: Actor i plays tennis at time t.

Hidden Markov Models

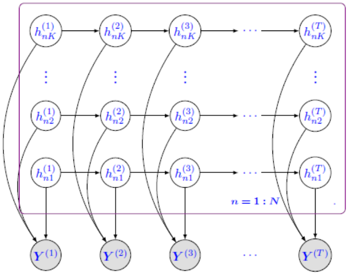


Figure 3 : hidden Markov model (HMM) shown in [Heaukulani 2013]

- $\boldsymbol{H}^{(1)}, \boldsymbol{H}^{(2)}, \ldots, \boldsymbol{H}^{(T)}$ comprise a latent Markov chain, i.e.

$$h_{ik}^{(t+1)} | h_{ik}^{(t)} \sim \boldsymbol{Q}(h_{ik}^{(t)}, h_{ik}^{(t+1)}).$$

- Transition probability does not depend on any observations.

Generative Process

Given latent structure $\boldsymbol{H}^{(t)}$ at time t:

$$
\begin{aligned}
y_{ij}^{(t)} | \boldsymbol{h}_i^{(t)}, \boldsymbol{h}_j^{(t)} &\sim \text{Bernoulli}\left(\pi_{ij}^{(t)}\right) \\
\pi_{ij}^{(t)} &= \sigma\left(\boldsymbol{h}_i^{(t)\,T}\, \boldsymbol{V}\, \boldsymbol{h}_j^{(t)} + s\right) \\
v_{kk'} &\sim \mathcal{N}(0,\, \sigma_v^2).
\end{aligned}
$$

where

- V: $K \times K$ feature-interaction weight matrix.
- $\sigma(x) = \dfrac{1}{1+e^{-x}}$: logistic sigmoid function.
- s: Link-bias parameter representing an underlying global probability of a link.

Outline

Outline

Motivation

In the context of social network, HMM assumes evolution of features over time does not depend on the past observations of social interactions.
But consider the followings:

- If my friend enjoy playing tennis, I am likely to start playing tennis.
- If a friend gets me to join the tennis team, I will likely befriend other tennis players.

This phenomenon is called as *Latent Feature Propagation*.
Goal: Capture the information propagation between network observation and latent structure over time.

Latent Feature Propagation Model

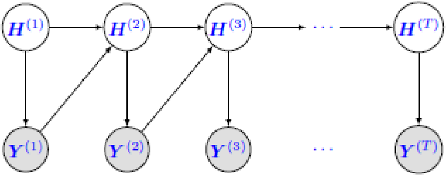


Figure 4 : Graphical representation for the latent feature propagation model

Latent Feature Propagation Model (Cont.)

Latent features evolve according to

$$h_{ik}^{(t+1)} | \mu_{ik}^{(t+1)} \sim \text{Bernoulli}\left[\sigma\left(c_k[\mu_{ik}^{(t+1)} - b_k]\right)\right] \tag{1}$$

$$\mu_{ik}^{(t+1)} = (1-\lambda_i)\, h_{ik}^{(t)} + \lambda_i \frac{h_{ik}^{(t)} + \sum_{j\in\epsilon(i,t)} w_j h_{jk}^{(t)}}{1 + \sum_{j'\in\epsilon(i,t)} w_{j'}} \tag{2}$$

where

- $\lambda_i \in [0,1]$: actor i's susceptibility to the influence of friend. Prior: Beta$(2,2)$.
- $w_i \in \mathbb{R}^+$: the weight of influence of actor i. Prior: Gamma$(1,1)$.
- $c_k \in \mathbb{R}^+$: the scale parameter for the persistence of feature k. Prior: Gamma$(1,1)$.
- $b_k \in \mathbb{R}^+$: the bias parameter for feature k. Prior: Gamma$(1,1)$.

Outline

Sample latent features $\boldsymbol{H}^{(1:T)}$

A forward-backward recursion algorithm [Scott 2002]:

- Forward step: Compute the 2×2 matrix $\boldsymbol{P}_t = (p_{trs})$ as

$$p_{trs} = P(h_{ik}^{(t-1)} = r,\, h_{ik}^{(t)} = s | \boldsymbol{Y}^{(1:t)}, \boldsymbol{H}_{-ik}^{(1:t)}, \Omega) \tag{3}$$

$$\propto \pi_{ik}^{(t-1)}(r)\, Q_{ik}^{(t-1,t)}(r,s)\, P(\boldsymbol{Y}^{(t)} | h_{ik}^{(t)} = s, \boldsymbol{H}_{-ik}^{(t)}, \Omega) \tag{4}$$

  where

  - * $\boldsymbol{H}_{-ik}^{(t)}$: the current states of all latent features at time t except $h_{ik}^{(t)}$.
  - * $Q_{ik}^{(t-1,t)}(r,s)$: Markov transition probability from state $h_{ik}^{(t-1)} = r$ to $h_{ik}^{(t)} = s$.

$$\begin{aligned} Q_{ik}^{(t-1,t)}(r,s) &= \left[\rho_{ik}^{(t)}\right]^s \left[1 - \rho_{ik}^{(t)}\right]^{(1-s)} \\ \rho_{ik}^{(t)} &= \sigma\left(c_k[\mu_{ik}^{(t)} - b_k]\right) \end{aligned}$$

  - *

$$\pi_{ik}^{(t)}(s) = P(h_{ik}^{(t)} = s | \boldsymbol{Y}^{(1:t)}, \boldsymbol{H}_{-ik}^{(1:t)}, \Omega) = \sum_r P_{trs} \tag{5}$$

Sample latent features $\boldsymbol{H}^{(1:T)}$ (Cont.)

- Stochastic backward step:
  * Sample $h_{ik}^{(T)} \sim \pi_{ik}^{(T)}(\cdot)$.
  * Sample remaining state backward for $t = T-1, T-2, \ldots, 1$ via

$$P(h_{ik}^{(T-t)} = r | h_{ik}^{(T-t+1)}, \boldsymbol{Y}^{(1:T-t+1)}, \Omega) \propto p_{T-t+1,\, r,\, h_{ik}^{(T-t+1)}}$$

Outline

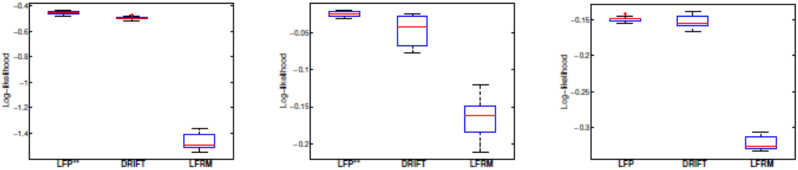
Experiments Setup

Compare the proposed LFP model with DRIFT model [Foulds et al. 2011] and LFRM model [Miller et al. 2009] on the following 3 datasets:

- Synthetic data: $N = 50$ actors, $T = 100$ time steps, and $K = 10$ latent features with parameters drawn from their prior distribution randomly.
- NIPS co-authorship: $N = 110$ researchers, $T = 17$ (from 1987 to 2003), and $K = 15$ set for training.
- INFOCOM $'06$: $N = 78$ students, $T = 50$, and $K = 10$ set for training.

Prediction of missing links

At each time point, different 20% of the interactions are held out as test set chosen uniformly at random.



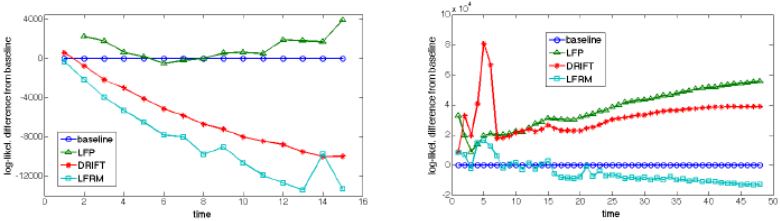
(a) Simulated data (b) NIPS dataset (c) INFOCOM dataset

Figure 5 : Log-likelihood of the test set.

Forecasting

For each time $t = 1, \ldots, T$, estimate the predictive distribution of the unseen network $\boldsymbol{Y}^{(t)}$ using

$$P(\boldsymbol{Y}^{(t)}|\boldsymbol{Y}^{(1:t-1)}) = \sum_{\boldsymbol{H}^{(t)}} \sum_{\boldsymbol{H}^{(1:t-1)}} P(\boldsymbol{Y}^{(t)}|\boldsymbol{H}^{(t)})$$
$$\times P(\boldsymbol{H}^{(t)}|\boldsymbol{H}^{(t-1)}, \boldsymbol{Y}^{(t-1)})\, P(\boldsymbol{H}^{(1:t-1)}|\boldsymbol{Y}^{(1:t-1)})$$



(a) NIPS dataset, $K = 15$ features. (b) INFOCOM dataset, $K = 10$ features.

Figure 6 : Forecasting a future unseen network.

Feature propagation

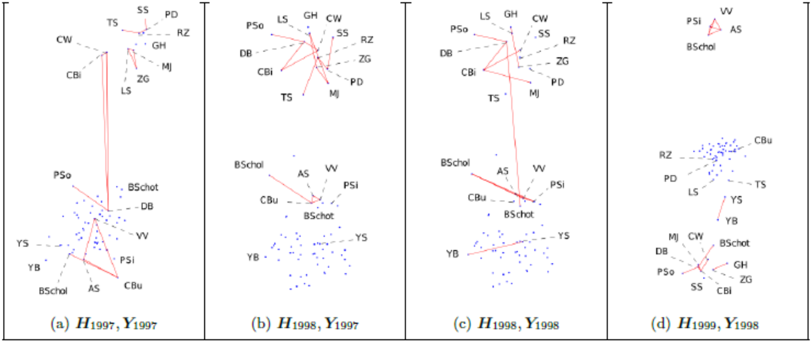


Figure 7 : Visualizing feature propagation in a subset of the NIPS dataset from 1997 to 1999.

References I

- James R Foulds, Christopher DuBois, Arthur U Asuncion, Carter T Butts, and Padhraic Smyth. "A dynamic relational infinite feature model for longitudinal social networks." : *International Conference on Artificial Intelligence and Statistics*. 2011, pp. 287–295.
- Creighton Heaukulani. *LFP Presentation*. June 2013. URL: http://mlg.eng.cam.ac.uk/heaukulani/LFP_Presentation.pdf.
- Kurt T Miller, Thomas L Griffiths, and Michael I Jordan. "Nonparametric latent feature models for link prediction." : *NIPS*. Vol. 9. 2009, pp. 1276–1284.
- Steven L Scott. "Bayesian methods for hidden Markov models." *Journal of the American Statistical Association* 97.457 (2002).