
Co-Evolving Mixed Membership BlockModels

Yoon-Sik Cho, Aram Galstyan, and Greg Ver Steeg
Information Science Institute
University of Southern California
Marina del Rey, CA 90292
{yoonsik, galstyan, greg}@isi.edu

Abstract

Many networks are complex dynamical systems, where both attributes of nodes and topology of the network (links structure) can change with time. We propose a model of co-evolving networks where both node attributes and network structure evolve under mutual influence. Specifically, we consider a mixed membership stochastic blockmodel, where the probability of observing a link between two nodes depends on their current membership vectors, while those membership vectors themselves evolve in the presence of a link between the nodes. Thus, the network is shaped by the interaction of stochastic processes describing the nodes, while the processes themselves are influenced by the changing network structure. We derive an efficient variational inference procedure for our model, and validate the model on both synthetic and real-world data.

1 Introduction

Recent surge in the online social media has made it possible to examine social networks at an unprecedented scale. Thus, it is important to have scalable approaches for modeling and understanding statistical properties of such systems. Many real-world networks are inherently complex dynamic systems, where both attributes of nodes and topology of the network can change with time. Traditionally, those two dynamical dimensions have been addressed separately. More recently, however, it has been realized that such separation is not adequate for capturing realistic behavior of networks. Different models that allow for interplay between evolution of attributes and the network structure have been proposed [5, 7].

To describe coupled individual and network dynamics, here we suggest a model of co-evolving network that is based on interacting stochastic processes. The idea behind this approach is the following: We represent the nodes as stochastic processes so that the network is shaped by the interaction of those processes. At the same time, those processes themselves are influenced by the changing network structure. This provides a feedback mechanism that is vital for capturing realistic behavior of complex real-world network.

Relational data is often described through stochastic block-structures, where each node is assigned to a block (roles), and the pattern of interactions between different nodes depend on the role assignment. Many situations, however, are better described by multi-faceted interactions, where nodes can bear multiple latent roles that influence their relationships to others. To account for such “mixed” interactions, Airoldi *et. al.* suggested Mixed Membership Stochastic BlockModels [1], where each node is described by a role membership vector, and where the interactions can be role-dependent. Here we propose CEMMB (Co-Evolving Mixed Membership Blockmodel), which provides a dynamic generalization of the mixed membership model in that it accounts for the coupled nature of node and network dynamics. In contrast to previously studied dynamics block-model [4], where the

dynamics was imposed *externally*, our model assumes that the membership evolution is driven by the interactions between the nodes through some influence mechanism.

2 Co-evolving Mixed Membership Blockmodel

Consider a set of N nodes, each of which can have K different roles, and let $\vec{\pi}_p^t$ be the mixed membership vector of node p at time t . Let Y_t be the network formed by those nodes at time t : $Y_t(p, q) = 1$ if the nodes p and q are connected at time t , and $Y_t(p, q) = 0$ otherwise. Further, let $Y_{0:T} = \{Y_0, Y_1, \dots, Y_T\}$ be a time sequence of such networks. The generative process that induces this sequence is described below.

- For each node p at time $t = 0$, employ a logistic normal distribution over a simplex sample $\pi_{p,k}^0 = \exp(\mu_{p,k}^0 - C(\vec{\mu}_p^0))$, $\vec{\mu}_p^0 \sim \mathcal{N}(\vec{\alpha}^0, A)$ where $C(\vec{\mu}) = \log(\sum_k \exp(\mu_k))$ is a normalization constant, and $\vec{\alpha}^0, A$ are prior mean, and covariance matrix.
- For each node p at time $t > 0$,
The mean of each normal distribution is updated due to *influence* from the neighbors at its previous step: $\vec{\alpha}_p^t = (1 - \beta_p)\vec{\mu}_p^{t-1} + \beta_p\vec{\mu}_{S(p,t-1)}$ where $\vec{\mu}_{S(p,t-1)}$ is average of $\vec{\mu}_s$ of the nodes which node p has met at time $t - 1$, and β_p describes how easily the node p is influenced by its neighbors. The membership vector at time t is $\pi_{p,k}^t = \exp(\mu_{p,k}^t - C(\vec{\mu}_p^t))$, $\vec{\mu}_p^t \sim \mathcal{N}(\vec{\alpha}_p^t, \Sigma_\mu)$ where Σ_μ is the noise term in the evolution process
- For each pair of nodes p and q at time t , draw role vectors:
– $\vec{z}_{p \rightarrow q}^t \sim \text{Multinomial}(\vec{z} | \vec{\pi}_p^t)$, sample a role indicator vector for donor p
– $\vec{z}_{p \leftarrow q}^t \sim \text{Multinomial}(\vec{z} | \vec{\pi}_q^t)$, sample a role indicator vector for receiver q .
Here $\vec{z}_{p \rightarrow q}$ is a unit indicator vector of dimension K , where $z_{p \rightarrow q, k} = 1$ means node p undertakes role k within the meeting with q .
- Sample a link between p and q according to Bernoulli trials:
– $Y_t(p, q) \sim \text{Bernoulli}(y | (1 - \rho)\vec{z}_{p \rightarrow q}^t B^t \vec{z}_{p \leftarrow q}^t)$ where B is a $K \times K$ role-compatibility matrix, so that B_{rs} described how likely is to form a link between two agents in roles r and s . When B is diagonal, the only possible interactions are among the nodes in the same role. Also, ρ is a parameter that accounts for the sparsity of the network.

Thus, the coupling between dynamics of different nodes is introduced by allowing the role vector of a node to be influenced by the role vectors of its neighbors. The influence is measured by a node-specific parameter β_p , which describes how easily the node p is influenced by its neighbors: $\beta_p = 0$ means it is not influenced at all, whereas $\beta_p = 1$ means the behavior is solely determined by the neighbors.

Inference and Learning Under the Co-Evolving MMSB, the joint probability of the data $Y_{0:T}$ and the latent variables $\{\vec{\mu}_{1:N}^t, \vec{z}_{p \rightarrow q}^t : p, q \in N, \vec{z}_{p \leftarrow q}^t : p, q \in N\}$ can be written in the following factored form.

$$p(Y_{0:T}, \vec{\mu}_{1:N}^{0:T}, \vec{Z}_{\rightarrow}^{0:T}, \vec{Z}_{\leftarrow}^{0:T} | \vec{\alpha}, A, B, \beta, \Sigma_\mu) = \prod_t \prod_{p,q} P(Y_t(p, q) | \vec{z}_{p \rightarrow q}^t, \vec{z}_{p \leftarrow q}^t, B^t) P(\vec{z}_{p \rightarrow q}^t | \vec{\mu}_p^t) \times P(\vec{z}_{p \leftarrow q}^t | \vec{\mu}_q^t) P(\vec{\mu}_p^{t+1} | \vec{\mu}_p^t, \vec{\mu}_{S(p,t)}, Y_t, \beta) \prod_p P(\vec{\mu}_p^0 | \vec{\alpha}, A) \quad (1)$$

Performing exact inference and learning under the exact model is not feasible. Thus, one needs to resort to approximate techniques. Here we use a variational EM [2, 3] approach. The main idea behind variational methods is to first posit a simpler distribution $q(X)$ over the latent variables with free (variational) parameters, and then fit those parameters so that the distribution is close to the true posterior in KL divergence.

$$D_{KL}(q||p) = \int_X q(X) \log \frac{q(X)}{p(X, Y)} dX \quad (2)$$

Algorithm 1 Variational EM

Input: data $Y_t(p, q)$, size N, T, K
Initialize all $\vec{\gamma}_p^t$
Start with initial guess of hyper parameters.
repeat
 repeat
 for $t = 0$ **to** T **do**
 repeat
 Update all $\{\sigma\}^t$ at time t
 Update all $\{\vec{\gamma}\}^t$ at time t
 until convergence in time single time t
 Update all $\{\zeta\}^t$ at time t
 Initialize $\phi_{p \rightarrow q}^t, \phi_{p \leftarrow q}^t$ to $\frac{1}{K}$ for all g, h
 Update all $\{\phi\}^t$ at time t
 end for
 until convergence across all time steps
 Update hyper parameters.
until convergence in haper parameters

Here we introduce the following factorized variational distribution:

$$q(\vec{\mu}_{1:N}^{0:T}, Z_{\rightarrow}^{0:T}, Z_{\leftarrow}^{0:T} | \vec{\gamma}_{1:N}^{0:T}, \Phi, \Phi) = \prod_{p,t} q_1(\vec{\mu}_p^t | \vec{\gamma}_p^t, \Sigma_p^t) \prod_{p,q,t} (q_2(\vec{z}_{p \rightarrow q}^t | \vec{\phi}_{p \rightarrow q}^t) q_2(\vec{z}_{p \leftarrow q}^t | \vec{\phi}_{p \leftarrow q}^t)) \quad (3)$$

where q_1 is the normal distribution, and q_2 is the multinomial distribution, and $\vec{\gamma}_p^t, \Sigma_p^t, \vec{\phi}_{p \rightarrow q}^t, \vec{\phi}_{p \leftarrow q}^t$ are the variational parameters. Variational EM algorithm works by iterating between the E-step of calculating the expectation value using the variational distribution, and the M-step of updating the model (hyper)parameters so that the data likelihood is locally maximized. The update equations for the parameters are rather cumbersome and will be provided elsewhere. The overall pseudo-algorithm is shown in Algorithm 1.

3 Results

We tested our model by generating a sequence of networks according to the process described above, for 50 nodes, and $K = 3$ latent roles across $T = 8$ time steps. We use a covariance matrix of $A = 3I$, and mean $\vec{\alpha}^0$ having homogeneous values for the prior, so that initially nodes have a well defined role (i.e., the membership vector is peaked around a single role). More precisely, the majority of nodes had around 90% of membership probability mass centered at a specific role, and on average 1/3 of those nodes will be having 90% on role 1, other 1/3s on $k = 2$, and $k = 3$. For the role-compatibility matrix, we gave high weight at the diagonal.

Starting from some initial parameter estimates, we performed variational EM and obtained a re-estimated parameters which were very close to the original values (ground truth). With those learned parameters, we inferred the hidden trajectory of agents as given by their mixed membership vector for each time step. The results are shown in Figure 1, where, for three nodes, we plot the projection of trajectories onto the simplex. One can see that for all three nodes, the inferred trajectories are very close to the actual ones.

We have also performed some preliminary experiments for testing our model against real-world data. In particular, we used senate co-sponsorship networks from the 97th to the 104th senate, by considering each senate as a separate time point in the dynamics. There were 43 senators who remained part of the senate during this period. For any pair of senators (p, q) in a given senate, we generated a directed link $p \rightarrow q$ if p co-sponsored at least 3 bills that q originally sponsored. The threshold of 3 bills was chosen not to have dense network. With this data, we wanted to test (a) to what extent senators tend to follow others who shares the political view (i.e., conservative vs. liberal) and (b) Whether some senators change their political creed easily, while others don't.

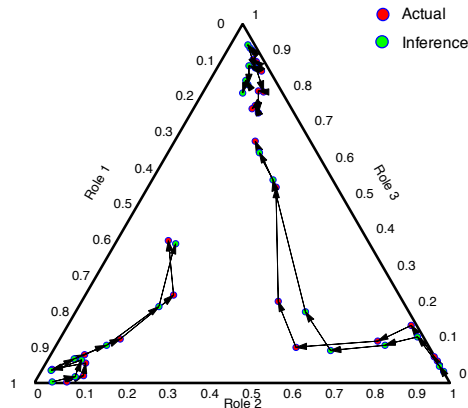


Figure 1: True and inferred mixed membership trajectories on a simplex.

The number of roles were chosen two to reflect the mostly bi-polar nature of the US Senate. The susceptibility of senators to influence is measured by the corresponding parameter β_p , which is learned using EM algorithm. High β means that a senator tend to change his/her role more easily. To initialize the EM procedure, we assigned the same β to all the senators, and start with a matrix which is weighted at the diagonal for B .

The role-compatibility matrix learned from the Variational EM was

$$B = \begin{bmatrix} 0.65 & 0.34 \\ 0.31 & 0.59 \end{bmatrix}$$

The high values at the diagonal allows us to verify that interaction is indeed more likely between senators that share the same role. Furthermore, the learned values of β showed that senators varied in their “susceptibility”. In particular, Sen. Arlen Specter was found to be the *most-influenceable* one, while Sen. Leahy was found to be the most *inert* one. Note that while there are no direct ways of estimating the “dynamism” of senators, our results seem to agree with our intuition about both senators (e.g., Sen. Specter switched the parties in 2009). To get some independent verification, we compared our results by the time-resolved ratings that ACU (American Conservative Union) assigns to senators every year¹. ACU rated every senators based on the selected votes which they believed to have a clear ideological distinction, so that high scores mean that they are truly conservative, while lower score suggests they are liberal. The ACU ratings are available for reach senator for every year from 1971 onwards. To compare the rating with our predictions (given by the membership vector) we had to properly normalized to former. And although those values cannot be used for quantitative agreement, we found that at least qualitatively, the inferred trajectories agree reasonably well with the ACU ratings. This agreement is rather remarkable since the ACU scores are based on selected votes rather than co-sponsorship network as in our data.

In conclusion, we have presented Co-Evolving Mixed Membership Blockmodel for modeling inter-coupled node and link dynamics in networks. We used variational EM approach for learning and inference with CEMMB, and were able to reproduce the hidden dynamics for synthetically generated data, both qualitatively and quantitatively. We also obtained reasonable agreement in our experiments with US Senate bill co-sponsorship data. As a future work, we intend to test our model against different real-world data, such as co-authorship network of publications, etc. We also plan to extend CEMMB in several ways. For instance, a bottleneck of the current model is that it explicitly considers links between all the pairs of nodes, resulting in a quadratic complexity in the network size. Most real world networks, however, are sparse, which is not accounted for in the current approach. Introducing sparsity into the model would greatly enhance its efficiency.

¹Accessible at <http://www.conservative.org/>

Acknowledgments

This research was supported in part by the National Science Foundation under grant No. 0916534 and the U.S. ARO MURI grant W911NF-06-1-0094.

References

- [1] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*, 9:1981–2014, 2008.
- [2] M. Beal and Z. Ghahramani. The variational bayesian em algorithm for incomplete data: with application to scoring graphical model structures. *Bayesian Statistics*, 7:453–464.
- [3] M. J. E. Xing and S. Russell. A generalized mean field algorithm for variational inference in exponential families. In *19th UAI*, 2003.
- [4] W. Fu, L. Song, and E. P. Xing. Dynamic mixed membership blockmodel for evolving networks. In *International Conference on Machine Learning*, 2009.
- [5] Holme, P., and Newman, M. E. J. 2006. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* 74(5):056108.
- [6] T. J. M. Jordan, Z. Ghahramani and L. Saul. *An Introduction to variational methods for graphical models*. Cambridge, MIT Press, 1999.
- [7] C. Steglich, T. A. B. Snijders, and M. Pearson. Interuniversity center for social science theory and methodology dynamic networks and behavior: Separating selection from influence, 2007.