Temporal Latent Space Network Model with VAR 1 Evolution of Latent Positions: Introduction

Anonymous Author(s)
Affiliation
Address
email

1 Motivation and Introduction

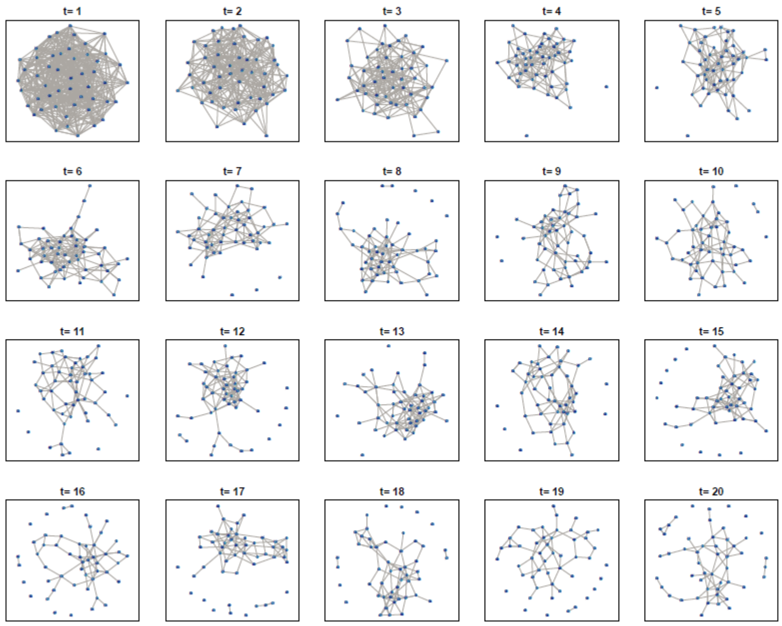
Methods that treat related networks to be independent over time are not well-suited for networks observed over multiple time points. By modeling the temporal networks as different instances of static networks, we lose important information about network evolution. Instead, modeling them as a continous process that accounts for time dependencies helps us understand the evolution of the underlying network structure. In this work, we propose a state space modeling approach for evolution of temporal networks using a class of latent variable models called latent space network models Hoff et al. (2002). We assume that network dynamics in time are direct functions of evolutions in the latent positions of the nodes. More specifically we focus on stationary vector autoregressive (VAR) representation of the evolution rather than random walk process to model a stable and stationary evolution process that is commonly seen in real world social networks.

We account for different possibilities of forming ties in temporal network data by extending the static latent space model (LSM). The LSM for social networks was introduced by Hoff et al. (2002); while Hoff and Ward (2004); Handcock et al. (2007); Raftery et al. (2012) have explored methodological and computational aspects of the static latent space models. LSM represents the underlying network structure by the positions of nodes in a continuous (Euclidean) latent space and thus allows for basic network properties like reciprocity and transitivity of the nodes, with possible extension to clusterability. We extend the ideas of reciprocity and transitivity in a static LSM to the temporal network setting. Nodes with ties at previous time points are more likely to have ties in the future, indicating that they will lie close to each other in the future latent space. Similary, if the nodes i and j, and the nodes i and k have ties at time t, the nodes i and k are more likely to have a tie at $t+1$. Furthermore, a network can evolve in time by the expansion of the latent positions making the network sparser, by the shrinkage of the latent positions which makes the network denser, or by no change in the latent positions which implies no substantive change in the network structure, after accounting for a common feature of the networks that do not change over time.

Existing temporal methods for latent space model focus on random walk evolution of latent positions (Sarkar and Moore, 2005; Sewell and Chen, 2014). While these models are a good starting point they do not necessarily account for many types of dynamics that we see in real world networks. Variance of the latent positions are increasing with time in random walk models thus implying that the network gets sparser with time, as demontrated in Figure 1. This assumption is not realistic in social networks, for example advice seeking network of teachers, where while the nodes move around in the space they do not necessarily keep moving away from each other. [[TODO:CITE]]

In this paper, we introduce an alternative way to specify the evolution of the latent positions as a stable and stationary process. We also present an MCMC algorithm to draw samples from the posterior distribution of the parameters. Latent space positions are identifiable in this model only upto the class of distance preserving transformations. This unidentifiability in the latent space positions also contributes to unidentifiability of the VAR parameter. Novel contribution of this work is in specifying the identifiable component of the VAR parameter and its significance in understanding the evolution of networks with time.

Recently, there has been some work on extending existing network methods to account for temporal networks, for example by Robins and Pattison (2001), Hanneke and Xing (2007),Hanneke et al. (2010), Westveld and Hoff (2011), Xing et al. (2010), Sarkar and Moore (2005) and Sewell and Chen (2014). Robins and Pattison (2001), Hanneke and Xing (2007) and Hanneke et al. (2010) have studied the networks observed over discrete time points in Exponential Random Graph Model (ERGM) settings, also known as ***temporal ERGMs or TERGMs***. TERGMs make standard Markov assumption on the evolution of a network graph such that Y_t is independent of $Y_1, \ldots, Y_{t-2}$ given Y_{t-1} (Hanneke and Xing, 2007), with an additional assumption that $P(Y_t|Y_{t-1})$ has an ERGM representation [[TODO:CITE ERGM PAPER HERE]]. Example of network statistics can include statistics representing stability, density, overall reciprocity, etc observed in the networks at time t and $t-1$. Similar to the static ERGM, assumptions on the dependence structure of ties for a network at time t conditional on the network at time $t-1$ will influence the potential network statistics that can be included in the model. As the assumption on the tie dependence deviates from independence, the model gets more and more complicated. Snijders (1996) have also developed stochastic actor oriented models using a continuous time Markov processes, the class of models that is very similar to ERGMs. Westveld and Hoff (2011) extended the static model for directed networks with sender and receiver random effects and fixed covariate effects, introduced by Gill and Swartz (2001) and later implemented by Hoff and Ward (2004), to account for temporal dependencies in ***mixed effects temporal model***. They assume autoregressive dependence structure on the sender-receiver effects and the overall residual, and hence account for additional correlation in random error (random effects) introduced by temporal dependencies. Xing et al. (2010) extended ***mixed membership stochastic block model (MMSBM)*** (Airoldi et al., 2006) to account for the temporal nature of networks, and called it the ***dynamic mixed***



$$
\begin{aligned}
Y_{ijt} &\sim \text{Bernoulli}(p_{ijt}) \\
p_{ijt} &= \beta_0 - ||Z_{it} - Z_{jt}|| \\
Z_{it} &\sim \text{Normal}(Z_{i(t-1)}, \Sigma) \\
\beta_0 &= 0.1 \\
\Sigma &= \begin{pmatrix} 0.75 & 0 \\ 0 & 0.75 \end{pmatrix} \\
T &= 20 \\
n_t &= 50, \forall T \\
dd &= 2
\end{aligned}
$$

Figure 1: Networks simulated under random walk plus noise evolution of the latent positions

membership stochastic block model (dMMSBM). Xing et al. (2010) developed a Bayesian state-space approach for modeling the evolution of the underlying roles of entities in a network, such that a network evolves in time through the random walk dependence structure on the hyperparameters of the prior distribution of the membership vectors and the block probabilities. Sarkar and Moore (2005) introduced a predictive latent space model for temporal networks. We also note that similar work is being developed independently by Sewell and Chen (2014). Both Sarkar and Moore (2005) and Sewell and Chen (2014) focus on random walk plus noise evolution.

Organization of the paper is as follows. We present the proposed model in section 2. We discuss the estimation of the model parameters using MCMC in section 3. We demonstrate the performance of the model in simulated data in section 4, and end with summary and conclusion in section 5.

2 Model

The ***latent space network model (LSM)*** introduced by Hoff et al. (2002) is characterized by the positions of the nodes in a low-dimensional latent space. Hoff et al. (2002) describe the latent space as a social space containing unobserved characteristics of the network, where nodes with similar latent characteristics will have nearby latent positions. First, it is assumed that conditional on the latent positions ties in the network are assumed to form independently. Secondly, probability of a tie between nodes i and j is inversely related to the interdistance between their latent positions. Implications of the assumption are: i. if two nodes share a tie, they lie close to each other in the latent space and,

ii. if two ties in a network share a common node, then the two remaining nodes will lie close to each other in the latent space hence increasing the probability of a tie between them. Thus, LSM accounts for the two basic network properties, ***reciprocity and transitivity***, which is described in more detail in Hoff et al. (2002).

Let Y denote a random variable representing a network graph with n nodes. It is usually represented by a $n \times n$ socio matrix, with entries Y_{ij}, where Y_{ij} measures the strength of a relationship from node i to node j, and can be either discrete or continuous. We use upper-case Y to denote a random variable, and lower-case y to denote its realization. For simplicity, we will consider a discrete and undirected network Y such that

$$Y_{ij} = \begin{cases} 1 \text{ if there is a tie between i and j} \\ 0 \text{ if there is no tie.} \end{cases}$$

However, these models can be easily extended to ordinal and continuous valued ties based on the techniques used for generalized linear models. We will use Z to denote a $n \times d$ matrix of the latent positions, such that its i^{th} row Z_i is a vector representing the position of a node i in a d dimensional latent space. Let p_{ij} denote the probability of forming a tie between nodes i and j and $d(Z_i, Z_j)$ denote the distance (for example, Euclidean) between the latent positions Z_i and Z_j.

LSM can then be written in notation as

$$\begin{aligned} Y_{ij} &\sim \text{Bernoulli}(p_{ij}) \\ \eta_{ij} &:= \text{logit}\,(p_{ij}) = \beta_0 - ||Z_i - Z_j|| \\ Z_i &\sim MVN(0, \Sigma). \end{aligned} \tag{1}$$

Further, likelihood of the observed network y conditional on the latent positions Z and the intercept β_0 is then given by

$$P(Y = y|Z, \beta_0) = \prod_{i \neq j} exp[\eta_{ij} Y_{ij} - \log{(1 + exp(\eta_{ij}))}]. \tag{2}$$

The intercept β_0 in the model can be seen as an overall fixed network effect, whereas the latent positions Z_is are the random effects. Further, it is evident from the model presented above that if two nodes i and j have the same latent position, then log-odds of forming a tie between i and j is β_0. LSM is a useful and appealing method for network analysis because it implicitly models different network features while making fewer assumptions about the dependence structure of the ties.

In this paper, we combine ideas from the latent space model of Hoff et al. (2002) and the state-space modeling approach [[TODO:CITE]] to model evolution of networks in time through the changes in latent positions. We allow stationary autoregressive dependence of order 1 in the nodal positions as a function of the previous latent positions.

Define Y_t as a sociomatrix of the network at time t with entries Y_{ijt} measuring a relationship from node i to j at that time point. Further, we will use $Y_{1:T} := [Y_1 Y_2 \ldots Y_T]$ to denote a block matrix of socio-matrices upto T time point, and $Z_{1:T} := [Z_1 Z_2 \ldots Z_T]$ to denote the block matrix of latent positions upto time T. Y_t is a $n \times n$ matrix of ties, with NA along the diagonal. Z_t is a $n \times d$ matrix of d dimensional latent space, with Z_{it} denoting the i^{th} row of Z_t. Finally, β_0 is an overall intercept of the model.

LSM for static model can be extended to account for temporal dependence as

$$\begin{aligned}
Y_{ijt} &\sim Bernoulli(p_{ijt}) \text{ for } i \neq j \\
logit(p_{ijt}) &= \beta_0 - ||Z_{it} - Z_{jt}|| \\
Z_{i,1} &\sim MVN(0, \Sigma_0), \text{ for } i = 1, \ldots, n \\
Z_{i,t} &= \Phi Z_{i,t-1} + \epsilon_t \text{ for } t = 2, \ldots, T \\
\epsilon_t &\sim MVN(0, \Sigma).
\end{aligned} \tag{3}$$

Further, assuming stationary VAR model we can compute the covariance matrix of each $Z_{i,t}$ as

$$vec(\Sigma_0) := vec(var(Z_{i,t})) = var(Z_{i,1}) = (I - \Phi * \Phi)^{-1} vec(\Sigma).$$

[[TODO:CITE]]

Here, $A * B$ is a Kronecker Delta product of two matrices A and B, I is an identity matrix of dimension $d^2 \times d^2$ and $vec(\Sigma)$ is a vector formed by stacking columns of Σ together.

Now lets look closely at the stationarity condition of centered VAR of order p, which is defined in Equation 4 as,

$$\begin{aligned}
Z_{it} &= \Phi_1 Z_{i(t-1)} + \Phi_2 Z_{i(t-2)} + \ldots + \Phi_p Z_{(i(t-p)} + \epsilon_{it} \\
Z_{it} - \sum_{j=1}^{p} \Phi_j Z_{i(t-j)} &= \epsilon_{it} \\
Z_{it} - \sum_{j=1}^{p} \Phi_j B^j Z_{it} &= \epsilon_{it} \\
(I - \sum_{j=1}^{p} \Phi_j B^j) Z_{it} &= \epsilon_{it}.
\end{aligned} \tag{4}$$

Here, B is the backshift operator and I is $d \times d$ identity matrix. Then we have the condition that Z_{it} is a stationary VAR(p) process if the roots of the $det\{I - \sum_{j=1}^{p} \Phi_j B^j\}$ are all outside the unit circle or equivalently all are greater than 1 in absolute value. For $VAR(1)$ process this condition is satisfied if the eigenvalues of Φ are less than 1 in absolute value.

3 Estimation

We use Metropolis Hastings within Gibbs algorithm to draw samples from the posterior distribution of the parameters, namely, β_0, $Z_{1:T}$, Φ and Σ. First, lets look at the joint likelihood of Y, Z, β_0 and Φ under the model illustrated in Equation 3. The joint likelihood of the data and the parameters in the model can be written as

$$\begin{aligned}
P(Y, Z, \beta_0, \Phi) &= P(Y|Z_1, \ldots, Z_T, \beta_0, \Phi)P(Z_1, \ldots, Z_T, \beta_0, \Phi) \\
&= \prod_{t=1}^{T} P(Y_t|Z_t, \beta_0) \times P(Z_1|0, \Sigma, \Phi) \times \prod_{t=2}^{T} P(Z_t|\Phi, Z_{t-1}, \Sigma) \times P(\beta_0) \times P(\Phi) \times P(\Sigma). \\
&= \prod_{t=1}^{T} \prod_{i \neq j} P(Y_{i,j,t}|Z_{i,t}, Z_{j,t}, \beta_0) \times \prod_{i=1}^{n} P(Z_{i,1}|0, \Sigma, \Phi) \\
&\times \prod_{t=2}^{T} \prod_{i=1}^{n} P(Z_{i,t}|\Phi, Z_{i,t-1}, \Sigma) \times P(\beta_0) \times P(\Phi) \times P(\Sigma)
\end{aligned} \tag{5}$$

The priors in the model for drawing samples from the poseterior distribution of the parameters using Metropolis Hastings within Gibbs algortihm can be specified as,

$$\begin{aligned}
\beta_0 &\sim \text{Normal}(\mu_0, \sigma_0^2) \\
\Sigma_{ii} &\sim \text{InverseGamma}(A, B) \\
\Phi_{ij} &\sim \text{Normal}(\mu, \tau)\mathbf{I}\{|\lambda_{ii}| < 1\}\forall ii
\end{aligned}$$

where, Φ_{ij} is the i, j^{th} entry of Φ and λ_{ii} is the ii^{th} eigen value of Φ. μ_z,Σ_0,μ_0,σ_0^2 are hyperparameters of the model specifying mean and variance of the prior distribution of latent positions at time $t = 1$ and intercept respectively. A and B are shape and rate parameters of the Inverse-Gamma prior distribution on the ii^{th} diagonal element of the variance of the error term in the latent space positions.

Likelihood of the network is related to the latent positions, Z_t's, only through inter-distance of the nodes in the latent space. Any isometric transformations in Z_t's will not change the first term representing the network likelihood in Equation 5. However, multivariate Normal density function, which is a prior density on latent positons, is not invariant to the isometric transformation unless the variance covariance matrix is diagonal representing a spherical distribution (Tong, 2012). These properties makes estimation of Z_ts and Φ non-identifiable. In the rest of the section, we discuss our approach to estimating identifiable component of Φ and its role in inference of the network dynamics.

We first show that the rotation in the latent positions is an issue only when drawing samples for the initial time point in Monte Carlo run in Theorem 3.1.

Theorem 3.1 *[[TODO:CITE/ Acknowledge Cosma for this proof]] Let Z_t and X_t be two isometric latent configurations. We may fix, within a single Monte Carlo run, the previous latent configuration Z_{t-1}, the evolution operator Φ, and the noise variance of the latent evolution Σ. Claim: If Σ^{-1} is strictly positive definite, then for Lebesgue-almost-all X_t isometric to Z_t, $Pr(X_t|Z_{t-1}, \Phi, \Sigma) \neq Pr(Z_t|Z_{t-1}, \Phi, \Sigma)$.*

Proof The two probabilities are equal iff their logarithms are equal. After canceling constant terms which are the same between the two probabilities, the logs are equal iff

$$(X_t - \Phi Z_{t-1})^T \Sigma^{-1} (X_t - \Phi Z_{t-1}) = (Z_t - \Phi Z_{t-1})^T \Sigma^{-1} (Z_t - \Phi Z_{t-1}).$$

Since (by assumption) Σ^{-1} is positive-definite, the quadratic forms appearing on either side of the equation are both ≥ 0. If only one is zero, there can't be equality, so either they're both zero or they're both positive. If they are both zero, then (again by positive-definiteness of the matrix) $X_t - \Phi Z_{t-1} = 0 = Z_t - \Phi Z_{t-1}$, implying $X_t = Z_t$, which contradicts the assumption that they are distinct.

Then, we are left with the case where both the quadratic forms equal the same positive number, lets call it c. The set of points in the latent space where the quadratic form is equal to c is an ellipsoid, centered at ΦZ_{t-1}, where the directions of the axes come from the eigenvectors of Σ and their widths from the eigenvalues. An important point here is the center, and the fact that, because the eigenvalues are all positive, the surface extends through all dimensions of the latent space.

Our concern is that X_t might be an -arbitrary- isometry of Z_t. But since X_t and Z_t must both lie on the ellipsoid, which is centered at ΦZ_{t-1}, they cannot be related through arbitrary translations. Thus we only need to concern ourselves with the case $X_t = RZ_t$, for an arbitrary rotation-reflection isometry R. But the set of such X_t consists of a sphere centered at the origin, which cannot coincide with, or even be a subset of, an ellipsoid centered elsewhere. There may be a set of intersection between the sphere and the ellipsoid, but (being the intersection of two low-dimensional surfaces) it will have Lebesgue measure 0. □

Thus, from the above proof we observe that while rotation of the latent positions at the first time point for each Monte Carlo draw is a potential concern for identification of Φ, the orientation of the positions for the subsequent time points at that draw will only depend on the orientation of the first one. We attempt to fix for the rotational effect by using Procustean transformation of the latent positions at the first time point to a fixed target for each MCMC draws.

Next we use results on how this tranformation affects the estimation of Φ parameter. Let us begin by rewriting the evolution model (state equation) that relates latent space positions at time t and time $t-1$:

$$Z_t = Z_{t-1}\Phi^T + \epsilon_t \tag{6}$$

We will frequently refer to 6 as an equation relating true underlying relationship between latent positions over time. Our goal in the estimation is to recover and understand this relationship. However, rotational invariance of the model while sampling the latent positions in this setup leads to unidentifiability in the Φ parameter. We now show that Φ can be identified only up to a class of similar matrix.

Let us consider the K^{th} draw of the latent positions for all T times points in MCMC estimation. We will denote it by $\hat{Z_{1:T}}^K$. The latent positions at $t = 1$ can only be estimated upto an isometry as the latent distance is invariant to such transformation. The orientation of $\hat{Z}_1$ also determines the orientation of $\hat{Z}_t$s for $t > 1$.

Denote $\hat{Z}_{1:T}^K$ as an isometric transformation of the true positions Z_t that satisfies the relationship in 6, upto Monte Carlo error. Let L_t^K denote the transformation operator at each time t during the estimation. Also, note that $L_t^K = L_{t-1}^K \forall t$. However, internodal distance is preserved during these transformations (upto MCMC error). If we use $D(.)$ to denote the Euclidean distance operator for a matrix . then we can write:

$$D(L_t^K(Z_t)) = D(Z_t) \forall t.$$

Further, since $Z_{1:T}$ are unknown (latent) parameters, $L_{1:T}^K$ are also unknown. We attempt to fix this problem by using procrustes transformation of the latent space positions, $\hat{Z}_1^K$, at each step of the Monte Carlo draw.

Let Z_{00} denote a fixed set of positions for n nodes in a d dimensional Euclidean space. We will call Z_{00} our ***target positions***. Then, at each step of MCMC we do procrustes transformation on $\hat{Z}_1^K$ such that they are as close to Z_{00} as possible while preserving the interdistances between the nodes. Lets denote this isometric transformation in the K^{th} MCMC draw by P_1^K. Note that $P_1(.)^K$ is a function of Z_{00} and $\hat{Z}_1^K$, as the transformation matrix for each time point depends on target Z_{00} and $\hat{Z}_1^K$.

In the next two Lemmas, we first show that doing procrustes transformation on $\hat{Z}_1^K$ produces the same set of positions as does the procrustes transformation on Z_1 directly. This fact justifies the use of procrustes transformation within our estimation method. Next, we show how the VAR parameter in the transformed positions are related to the true Φ.

Theorem 3.2 *Let Z_{00} be an arbitrary (centered) positions. For each $\hat{Z}_1^K$, let P_1^K denote a isometric transformation such that:*

$$P_1^K := argmin_T tr(Z_{00} - \hat{Z}_1^K T)(Z_{00} - \hat{Z}_1^K T)^T.$$

For each Z_1, let T_1^* denote a isometric transformation such that:

$$T_1^* = argmin_T tr(Z_{00} - Z_1 T)(Z_{00} - Z_1 T)^T.$$

Then, $T_1^* = L_1^K P_1^K$.

Proof Let $SVD(Z_{00}^T Z_1) = U\Lambda V^T$.

Then, $T_1^* := \text{argmin}_T tr(Z_{00} - Z_1 T)(Z_{00} - Z_1 T)^T = VU^T$.

Also,

$$\begin{aligned}
Z_{00}^T Z_1 L_1^K &= SVD(Z_{00}^T Z_1) L_1^K \\
&= U\Lambda V^T L_1^K \\
&= U\Lambda V_1^{*KT} \\
&\quad (V_t^{*KT} \text{ is an orthogonal matrix}) \\
&= SVD(Z_{00}^T Z_1 L_1^K).
\end{aligned}$$

Then, $P_1^K := \text{argmin}_T tr(Z_{00} - T Z_1 L_1^K) = V_1^{*K} U^T = L_1^{TK} V U^T$.

Finally note that: $L_1^K P_1^K = L_1^K L_1^{KT} V U^T = VU^T = T_1^*$. ∎

Theorem 3.3 *Let $SVD(\Phi) = U\Lambda V^T$ and $SVD(\Phi^{*K}) = U^{*K}\Lambda^{*K}V^{*TK}$ where, U, V, U^{*K} and V^{*K} are orthogonal matrices and Λ and Λ^{*K} are diagonal matrices of singular values.*

If $\hat{Z}_t^K = \hat{Z}_{t-1}^K(\Phi^{*K})^T + \epsilon_t^{*K}$, then Φ and Φ^{*K} are similar matrices. Further, $\Lambda = \Lambda^{*K}$.

Proof The new dependence equation between transformed latent space positions up to Monte Carlo error can be written as:

$$\begin{aligned}
P_t(W_t^K) &= P_{t-1}(W_{t-1}^K)\Phi^{*KT} + \epsilon_t^{*K} \\
P_t(L_t(Z_t)) &= P_{t-1}(L_{t-1}(Z_{t-1}))\Phi^{*T} + \epsilon_t^{*K} \\
Z_t(P_t^K L_t^K)^T &= Z_{t-1}(P_{t-1^K} L_{t-1}^K)^T \Phi^{*KT} + \epsilon_t^{*K} \\
Z_t &= Z_{t-1}(P_{t-1}^K L_{t-1}^K)^T \Phi^{*KT}(P_t^K L_t^K) + \epsilon_t^{*K} P_t^K L_t^K.
\end{aligned}$$

The above equations give us tools to relate Φ^* and Φ such that:

$$\Phi = (P_{t-1}^K L_{t-1}^K)^T \Phi^{*K}(P_t^K L_t^K). \tag{7}$$

First note that, $P_{t-1}L_{t-1}$ and $P_t L_t$ are both rotation matrices. Also,

$$\begin{aligned}
(P_{t-2}^K L_{t-2}^K)^T \Phi^{*K} P_{t-1}^K L_{t-1}^K &= (P_{t-1}^K L_{t-1}^K)^T \Phi^{*K}(P_t^K L_t^K) \\
\Rightarrow \Phi^{*K} &= (P_{t-2}^K L_{t-2}^K)(P_{t-1}^K L_{t-1}^K)^T \Phi^{*K}(P_t^K L_t^K)(P_{t-1}^K L_{t-1}^K)^T.
\end{aligned}$$

Thus we have that,

$$\begin{aligned}
(P_{t-2}^K L_{t-2}^K)(P_{t-1}^K L_{t-1}^K)^T &= I \\
\Rightarrow P_{t-2}^K L_{t-2}^K &= ((P_{t-1}^K L_{t-1}^K)^T)^{-1} \\
&= ((P_{t-1}^K L_{t-1}^K)^T)^T \\
&= (P_{t-1}^K L_{t-1}^K).
\end{aligned}$$

We can similarly show that $P_t^K L_t^K = P_{t-1}^K L_{t-1}^K = C^K$ for all $t = 1, \ldots, T$, where C^K is an orthogonal matrix. Putting this all together we can rewrite 7 as:

$$\Phi = C^{KT}\Phi^{*K}C^K = (C^K)^{-1}\Phi^{*K}C^K.$$

Thus, we showed that Φ and Φ^{*K} are similar matrices.

Next, denote $SVD(\Phi) = U\Lambda V^T$ where U and V are orthogonal matrices. Also, denote $SVD(\Phi^*)$ by $U^*\Lambda^* V^{*T}$. Since, $SVD(\Phi)$ is unique given some regulatory conditions on Φ we can show that $\Lambda = \Lambda^*$. Further,

$$(P_{t-1}L_{t-1})^T U^* = U$$

and

$$(P_t L_t) V^{*T} = V^T$$

Further, the variance covariance matrix of ϵ_t^* is also transformed such that $COV(\epsilon_{it}^* P_t L_t) = (P_t L_t)\Sigma^*(P_t L_t)^T = COV(\epsilon_{it})$. In some ways, this equation gives us a way to define a distribution of ϵ_t.

3.1 Estimation with Missing Nodes

The first term in the product in Equation 5 can be easily obtained from Equation 2 for each time point t since the networks are independent over time conditional on the latent positions and the intercept. However, we need to account for the changing number of the nodes over time while computing the likelihood of the latent positions. We will assume that nodes are missing at random. Further, once a node exits a network there is a very less chance that it will re-enter. Thus, if a node enters the network after $t = 1$ we will use the time point as its initial time and assume that its latent position has the same prior as Z_{i1}. Let $\{N_{t-1}\}$ denote set of nodes at time t that were also present at $t - 1$.

The likelihood in Equation 5 can be re-written as

$$\begin{aligned} P(Y, Z, \beta_0, \Phi) = &\prod_{t=1}^{T} \prod_{i \neq j} P(Y_{i,j,t} | z_{i,t}, z_{j,t}, \beta_0) \times \prod_{i=1}^{n_1} P(z_{i,1} | 0, \Sigma_0(\Phi)) \\ &\times \prod_{t=2}^{T} \prod_{i=1}^{n_t} [P(z_{i,t} | \Phi, z_{i,t-1}, \Sigma) \mathrm{I}(i \in \{N_{t-1}\}) + P(z_{i,t} | 0, \Sigma_0(\Phi)) \mathrm{I}(i \in \{N_{t-1}\})] \\ &\times P(\beta_0) \times P(\Phi) \times P(\Sigma). \end{aligned} \tag{8}$$

3.2 Prediction and Model Comparision

One of the goals of this work is to develop a systematic approach to predict future ties at T_1 given the networks upto time T. We can use the posterior MCMC draws of $Z_{1:T}$ to estimate $\hat{Z}_{1:T}$. Observe that, for VAR(1) evolution model:

$$
\begin{aligned}
P(Z_{T+1}|Y_{1:T} &= \int P(Z_{T+1}|Z_{1:T}, \beta_0, \Phi, \Sigma) P(Z_{1:T}|Y_{1:T}, \beta_0, \Phi, \Sigma) dZ_{1:T} d\beta_0 d\Phi d\Sigma \\
&\approx \frac{1}{L} \sum_{l=1}^{L} \prod_{i=1}^{n_{T+1}} [N(Z_{i(T+1)}|\Phi Z_{iT}^l, \Sigma^l) \mathsf{I}(i \in \{N_T\}) + N(Z_{i(T+1)}|\mu_z, \Sigma_0^l(\Phi^l)) \mathsf{I}(i \notin \{N_T\})]
\end{aligned} \tag{9}
$$

Then,

$$
\hat{Z}_{T+1} = \mathsf{E}(Z_{T+1}) = \begin{cases} \frac{1}{L} \sum_{l=1}^{L} (\Phi^l Z_{iT}^l) & \text{if}(i \in \{N_T\}) \\ \mu_z & \text{else} \end{cases}.
$$

And finally the predictive probability of tie between nodes i and j at time $T+1$ is

$$
\hat{p}_{ij(T+1)} = \hat{\beta}_0 - ||\hat{Z}_{i(T+1)} - \hat{Z}_{j(T+1)}||.
$$

We use similar method for prediction using estimates of latent space positions from random walk model.

To make latent space model comparable with other longitudinal data we use $Y_{1:T}$ to draw intercept and Σ in the following way:

4 Simulation

5 Summary and Discussion

References

Edoardo M. Airoldi, David M. Blei, Stephen E. Fienberg, Eric P. Xing, and Tommi Jaakkola. Mixed membership stochastic block models for relational data with application to protein-protein interactions. In *In Proceedings of the International Biometrics Society Annual Meeting*, 2006.

Paramjit S. Gill and Tim B. Swartz. Statistical analyses for round robin interaction data. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 29(2):pp. 321–331, 2001. ISSN 03195724. URL `http://www.jstor.org/stable/3316080`.

Mark S. Handcock, Adrian E. Raftery, and Jeremy M. Tantrum. Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(2):301–354, 2007.

Steve Hanneke and Eric P. Xing. Discrete temporal models of social networks. In *Proceedings of the 2006 Conference on Statistical Network Analysis*, ICML'06, pages 115–125, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 978-3-540-73132-0. URL `http://dl.acm.org/citation.cfm?id=1768341.1768352`.

Steve Hanneke, Wenjie Fu, Eric P. Xing, et al. Discrete temporal models of social networks. *Electronic Journal of Statistics*, 4:585–605, 2010.

Peter D. Hoff and Michael D. Ward. Modeling dependencies in international relations networks. *Political Analysis*, 12(2):160–175, 2004.

Peter D. Hoff, Adrian E. Raftery, and Mark S. Handcock. Latent space approaches to social network analysis. *JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION*, 97:1090–1098, 2002.

Adrian E. Raftery, Xiaoyue Niu, Peter D. Hoff, and Ka Yee Yeung. Fast inference for the latent space network model using a case-control approximate likelihood. *Journal of Computational and Graphical Statistics*, 21(4):901–919, 2012.

Garry Robins and Philippa Pattison. Random graph models for temporal processes in social networks. *The Journal of Mathematical Sociology*, 25(1):5–41, 2001. doi: 10.1080/0022250X.2001.9990243. URL `http://dx.doi.org/10.1080/0022250X.2001.9990243`.

Purnamrita Sarkar and Andrew W. Moore. Dynamic social network analysis using latent space models. *ACM SIGKDD EXPLORATIONS NEWSLETTER*, 7(2):31–40, 2005.

Daniel K. Sewell and Yuguo Chen. Latent space models for dynamic networks. *Journal of the American Statistical Association*, (ja):00–00, 2014. URL `http://dx.doi.org/10.1080/01621459.2014.988214`.

Tom A.B. Snijders. Stochastic actor-oriented models for network change. *Journal of mathematical sociology*, 21(1-2):149–172, 1996.

Yung Liang Tong. *The multivariate normal distribution*. Springer Science & Business Media, 2012.

Anton H. Westveld and Peter D. Hoff. A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict. *The Annals of Applied Statistics*, 5(2A):843–872, 2011.

Eric P. Xing, Wenjie Fu, and Le Song. A state-space mixed membership blockmodel for dynamic network tomography. *The Annals of Applied Statistics*, 4(2):535–566, 06 2010. doi: 10.1214/09-AOAS311. URL `http://dx.doi.org/10.1214/09-AOAS311`.