A review of dynamic network models with latent variables

Bomin Kim[1,†], **Kevin H. Lee**[2,†], **Lingzhou Xue**[1], **Xiaoyue Niu**[1,‡]

[1]Department of Statistics, The Pennsylvania State University

[2]Department of Statistics, Western Michigan University

Abstract

We present a selective review of statistical modeling of dynamic networks. We focus on models with latent variables, specifically, the latent space models and the latent class models (or stochastic blockmodels), which investigate both the observed features and the unobserved structure of networks. We begin with an overview of the static models, and then we introduce the dynamic extensions. For each dynamic model, we also discuss its applications that have been studied in the literature, with the data source listed in Appendix. Based on the review, we summarize a list of open problems and challenges in dynamic network modeling with latent variables.

Keywords

Dynamic networks; latent space model; stochastic blockmodel; latent variable model; social network analysis

1. Introduction

As statistical modeling of network data has been posited as a major topic of interest in diverse areas of study, there have been an increasing number of books [42, 57] and survey papers on random graphs and network models [12, 21, 73, 78, 83, 54]. Those existing surveys provide a comprehensive overview of the historical development of statistical network modeling, including the summary of network models that are not latent variable models—e.g., Exponential random graph models (ERGMs), the quadratic assignment procedure (QAP), stochastic actor oriented models (SAOMs)—as well as the latent variable models—e.g., statist latent space models (LSM) and the stochastic blockmodels (SBM). The key idea behind introducing latent variables into network analysis is to capture various forms of dependence between edges and get conditional independence in the error terms, which is one of the most challenging parts of network modeling. Considering the difference between latent variable models and the rest in the network literature, we aim to provide in-depth information on the unobserved or unmeasured structure of networks by presenting a selective review on dynamic network models with latent variables. This area has undergone significant developments in recent years, with emphasis on two classes of models, the latent space models (LSM) and the stochastic blockmodels (SBM).

‡Corresponding author: xiaoyue@psu.edu.
†These authors contributed equally to this work.

The latent space models (LSM) [35] assume that nodes are positioned in an R-dimensional latent space, and they tend to create edges to others that are closer in their latent positions. Due to this simple geometry-based assumption, the latent space models have an advantage in providing the useful visualization and interpretation of network or relational data, and thus have been widely used in numerous fields of study [10, 17, 25, 36, 43, 76, 77, 82]. The stochastic blockmodels (SBM) assume that the nodes of the network are partitioned into several unobserved (latent) classes (or blocks). The framework is first introduced by Holland et al. [37] which focuses on the case of *a priori* specified blocks, where the membership of nodes are known or assumed, and the goal is to estimate a matrix of edge probabilities. A statistical approach to *a posteriori* block modeling for networks is introduced by [74] and [59], where the objective is to estimate the matrix of edge probabilities and the memberships simultaneously. Since then, the communities found in the stochastic blockmodels have been interpreted meaningfully in many research fields. For example, in citation and collaboration networks, such communities could be interpreted as scientific disciplines [39, 58], while the communities in food web networks could be interpreted as ecological subsystems [20].

As dynamic network analysis—the study of networks that evolve over time—has become an emergent scientific field in the last decade, there has been a growing number of dynamic network models that incorporate a latent space model or stochastic blockmodel framework. Since they are relatively new, to the best of our knowledge, there has not been any attempt to describe the progress in dynamic latent variable models. In this work, therefore, we outline some of the prominent dynamic latent space models and summarize their interconnections, and also describe the recent approaches in dynamic stochastic blockmodels.

We organize the review in the following way. For each class of models (LSM and SBM), we first introduce a family of static models as the background information, then describe the dynamic models motivated by the static ones, and discuss the applications to real-world data. In Section 5, we present two diagrams summarizing the relationships of all the models and mention some open problems and remaining challenges in dynamic network models with latent variables. We also provide a list of available data sources that have been used in past literature in the Appendix.

2. Notations

Throughout the paper, we consider the set of nodes $\mathcal{N}$ labeled as $1, \ldots, N$, and assume the edges represented by an $N \times N$ adjacency matrix $Y = \left(y_{ij}\right)_{(i, j) \in \mathcal{N}}$, where y_{ij} being the edge value between i and j which could be binary, discrete, or continuous. To achieve the goal of statistical network models—to understand the dependence of the edges using the observed and unobserved structure—we define the observed pair-specific covariates as $X = \left(x_{ij}\right)_{(i, j) \in \mathcal{N}}$ and the unobserved latent variables as $Z = \left(z_i\right)_{i \in \mathcal{N}}$. Specifically, z_i represents an R-dimensional latent position (or vector) of node i in the latent space models (Section 3), and z_i represents the class of node i in the stochastic blockmodels where each z_i takes one value in the set of categories $\mathcal{C}$ labeled as $1, \ldots, C$ (Section 4). When extending to the dynamic models for discrete time points $t = 1, \ldots, T$, we use a t subscript on the variables (e.g., y_{ijt}, x_{ijt}, z_{it}) to denote "at time t". Generally, random variables are denoted by upper

case letters and fixed quantities or realizations of random variables are denoted by lower case letters. Other than the ones specified here, additional notation specific to each model is introduced later.

3. Latent space models

In this section, we first describe the original latent space model introduced by Hoff et al. [35]. Then we introduce two lines of research: (i) the latent position model [23], which is built upon the Euclidean distance space, and (ii) the latent factor model [28], which stems from the projection model. We present the dynamic extension of these static models and demonstrate how the two different frameworks have evolved in the dynamic network literature, in chronological order.

3.1. Static latent space models

3.1.1. Latent space approaches to social network analysis—The idea of social space is first introduced to network modeling by Hoff et al. [35], with the intuition that each node $i \in \mathcal{N}$ can be seen as a point z_i in an R-dimensional space that represents unobserved latent characteristics. The presence or absence of an edge between two nodes is independent of all other edges given the latent positions of the two nodes; thus the conditional probability of the adjacency matrix Y is

$$P(Y|Z, X, \theta) = \prod_{i \neq j} P\left(y_{ij}|z_i, z_j, x_{ij}, \theta\right),$$

where X are observed pair-specific covariates, $\theta = (\alpha, \beta)$ are the set of regression parameters. Here, the latent positions Z can be interpreted as the random effects in linear models.

There exist two different ways to model $P(y_{ij}|z_i, z_j, x_{ij}, \theta)$: the distance model and the projection model. The main assumption here is that nodes are positioned in a latent space, and they tend to create edges to others with shorter distance (distance model) or narrower angle (projection model) between their latent positions. Both distance and projection models can be applied to various types of edges (e.g., binary, discrete, continuous) via the linked mean parameter of a generalized linear model. Without loss of generality, we introduce the model formulations for the case of binary networks.

Latent distance model: For a binary network, the distance model takes the form of a logistic regression model with $\theta = (\alpha, \beta)$ and defines the log odds of an edge between nodes i and j as:

$$\eta_{ij} = \text{logodds}\left(y_{ij} = 1|z_i, z_j, x_{ij}, \alpha, \beta\right)$$
$$= \alpha + \beta' x_{ij} - \left\|z_i - z_j\right\|,$$

where α is an intercept term, β is a vector of coefficients for covariate effects, and $\|z_i - z_j\|$ is the Euclidean distance between nodes in the latent space. In other words, the existence of an

edge in the adjacency matrix (i.e., $Y_{ij} = 1$) is determined by the dyad attributes x_{ij} and the distance between the pair of nodes $\|z_i - z_j\|$.

Krivitsky et al. [47] proposes the latent cluster random effects model that includes the distance between latent space positions, model-based clustering of the latent positions, and random sender and receiver effects (discussed below), and introduces Bayesian estimation methods for both binary and non-binary network data.

Latent projection model: The projection model posits a different assumption:

$$\eta_{ij} = \text{logodds}\left(y_{ij} = 1 | z_i, z_j, x_{ij}, \alpha, \beta\right)$$
$$= \alpha + \beta' x_{ij} + \frac{z'_i z_j}{\|z_j\|},$$

which changes the interpretation so that i and j are more likely to form an edge if z_i and z_j are in the same direction $\left(\text{i.e., } z'_i z_j > 0\right)$, while less likely to form an edge if they are in opposite directions $\left(\text{i.e., } z'_i z_j < 0\right)$. Due to the denominator term $\|z_j\|$, this also allows asymmetric edge probabilities (i.e., $\eta_{ij} \quad \eta_{ji}$) even when the dyadic covariates x_{ij} are symmetric.

3.1.2. Latent position cluster model—As identifying groups of similar nodes is often of interest to network researchers, Handcock et al. [23] extends the latent space models to allow clustering of nodes in networks. Referred to as the "latent position cluster model", this approach combines the latent space models and the model-based clustering method [7]. The modeling framework is built upon the latent distance model

$$\text{logodds}\left(y_{ij} = 1 | z_i, z_j, x_{ij}, \alpha, \beta\right) = \beta'_0 x_{ij} - \beta_1 \|z_i - z_j\|,$$

where each position $z_i \in \mathbb{R}^K$ is drawn from a finite mixture of multivariate normal distributions from G groups. Assuming different means and covariance matrices for each group $g = 1, \ldots, G$, the model proposes

$$z_i \sim \sum_{g=1}^{G} \lambda_g N_d\left(\mu_g, \sigma_g^2 I_R\right),$$

where λ_g is the probability that an actor belongs to g, and I_R is the $R \times R$ identity matrix.

Inference of the latent space model and the clustering model can be made with either the maximum likelihood method (MLE) or the Bayesian approach via Monte Carlo Markov chain (MCMC), and those methods are implemented in the R package 'latentnet' [44, 46].

3.1.3. Bilinear mixed-effects model—Providing some computational and conceptual advantages to the latent space models, Hoff [28] develops the generalized bilinear mixed-

effects models (GBME), a network regression model that is in line with the latent space models, especially the latent projection models. Starting from the traditional regression model on dyadic data

$$y_{ij} = \beta' x_{ij} + \epsilon_{ij},$$

this new approach includes additive and multiplicative random effects in the error terms, i.e.,

$$\epsilon_{ij} = a_i + b_j + z_i' z_j + \gamma_{ij}$$

$$(a_i, b_j)' \sim N(0, \Sigma_{ab}),$$

$$z_i \sim N(0, \sigma_z^2 I_R),$$

$$(\gamma_{ij}, \gamma_{j,i})' \sim N(0, \Sigma_\gamma), \text{ where }$$

$$\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \text{ and } \Sigma_\gamma = \sigma_\gamma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

This representation allows us to explain the second-order dependence often seen in dyadic data: a sender effect a_i, a receiver effect b_j, and a within dyad effect γ_{ij}. Variance parameters σ_a^2 and σ_b^2 represent the dependence of observations having a common sender and receiver, respectively, with the correlation between sender and receiver effects described by σ_{ab}. Moreover, reciprocity and within-actor correlation are measured by σ_γ^2 and the correlation parameter ρ. Finally, the bilinear effect in the product term $z_i' z_j$ is known to capture the third-order dependence in dyadic data such as transitivity (i.e., if $y_{ij} = 1$ and $y_{jk} = 1$, then we tend to have $y_{ik} = 1$), balance (i.e., $y_{ij} \times y_{jk} \times y_{ki} > 0$ for signed unordered edges), and clusterability (i.e., a triad can be divided into groups).

The GBME framework has been further studied and modified in [29, 30, 55], and has been referred to as the "latent factor models", the "eigenmodel", and the "additive and multiplicative effects (AME) models" [27, 32, 55]. Throughout this paper, we use the name "latent factor models" to emphasize the objective of the paper in understanding latent variable models. Hoff [29] shows that the latent factor model generalizes the latent distance model and the latent class model. For the effective use of the latent factor models, the R package "AMEN" [32] provides estimation and inference for a class of AME models for ordinal, continuous, binary and other types of dyadic data.

3.2. Dynamic latent distance models

In this section, we summarize some dynamic latent distance models which extend the latent distance models with Markovian properties.

3.2.1. Dynamic social network in latent space model

The first dynamic latent distance model appeared in the machine learning literature by Sakar and Moore [67]. They named it the "dynamic social network in latent space (DSNL) model". This model assumes that Y_t (i.e., the observed pairwise adjacency matrix at time t) is only dependent on the current latent positions Z_t ("observational model") while allowing the latent positions to move over time with standard Markovian assumptions ("transition model").

The observation model for the graph can be written as

$$P\!\left(Y_t \mid Z_t\right) = \prod_{i \sim j} p_{ijt} \prod_{i \not\sim j} \left(1 - p_{ijt}\right),$$

where $i \sim j$ and $i \not\sim j$ denote the existence and absence of a link, respectively, and p_{ijt} is the probability of a link between i and j at time t. Specifically, it has the following form:

$$p_{ijt} = \frac{1}{1 + e^{\left(d_{ijt} - r_{ijt}\right)}} K\!\left(d_{ijt}\right) + \rho\left(1 - K\!\left(d_{ijt}\right)\right),$$

where $d_{ijt} = \|z_{it} - z_{jt}\|$ is the Euclidean distance between i and j in the latent space as in the original latent distance model at time t, r_{ijt} is the maximum of the radii of i and j at time t, defined as $r_{ijt} \propto (max(\delta_{it}, \delta_{jt}) + 1)$ with δ_{it} being the degree of node i at time t, K is a biquadratic kernel $K(d_{ijt}) = (1 - (d_{ijt}/r_{ijt})^2)^2$, and ρ is a constant noise.

The transition model assumes a Gaussian random walk:

$$Z_t \mid Z_{t-1} \sim N\!\left(Z_{t-1}, \sigma^2 I\right),$$

and the resulting $\log P(Z_t \mid Z_{t-1})$ becomes

$$\log P\!\left(Z_t \mid Z_{t-1}\right) \propto - \sum_{i=1}^{n} \left| z_{it} - z_{i(t-1)} \right|^2 / 2\sigma^2,$$

such that we want to estimate the positions using

$$P\!\left(Z_t \mid Y_t, Z_{t-1}\right) \propto P\!\left(Y_t \mid Z_t\right) P\!\left(Z_t \mid Z_{t-1}\right).$$

For efficient optimization of the likelihood, the authors develop the two-stage learning algorithms: (i) generalized multidimensional scaling (MDS) to find the initial latent positions across time and (ii) nonlinear conjugate gradient (CG) optimization starting from

the initial estimates. Additionally, the Procrustean transformation is applied to the coordinates from MDS so that Z_t maintains the same orientation as Z_{t-1}.

The DSNL model is applied to NIPS conference paper co-authorship dataset (network size up to 11,000) by separating the 12 years of data into three discrete time points: 1987–1990, 1991–1994, and 1995–1998. The analysis is mainly focused on investigating the dynamics of some well-connected researchers in the machine learning community, and authors find out some noticeable changes in the network embeddings via the examination of latent positions. Sakar et al. [68] further extends the model to be applicable for dynamic bipartite or two-mode networks, such as author-word data from the same NIPS data using text corpora.

3.2.2. Dynamic latent distance model with popularity and activity effects—

Sewell and Chen [69] propose a network model for longitudinal networks which allows each node to have a temporal trajectory in an R-dimensional latent Euclidean space, with additional features to capture the nodes' popularity and activity effects.

For a binary network Y_t, the model formulation relies on the log-odds η_{ij}:

$$P\big(Y_t | Z_t, \beta_{IN}, \beta_{OUT}, r_{1:N}\big) = \prod_{i \neq j} \frac{\exp\big(y_{ijt}\eta_{ijt}\big)}{1 + \exp\big(y_{ijt}\eta_{ijt}\big)}, \text{ where}$$

$$\eta_{ijt} = \beta_{IN}\left(1 - \frac{d_{ijt}}{r_j}\right) + \beta_{OUT}\left(1 - \frac{d_{ijt}}{r_i}\right),$$

where the distance between i and j at time t is defined as $d_{ijt} = \|z_{it} - z_{jt}\|$ and r_i is a vector of positive actor-specific parameters with constraints $\sum_{i=1}^{N} r_i = 1$. Specifically, r_i can be thought of as the ith actor's social reach or radius which does not vary over time. In addition, this formulation separates the two parameters β_{IN} and β_{OUT} to measure the global popularity and activity effects, respectively. Note that this model deals with directed networks (i.e., y_{ij} y_{ji}) as well as undirected ones (i.e., $y_{ij} = y_{ji}$), while the DSNL model in Section 3.2.1 only allows undirected edges.

Similar to the DSNL model, the latent positions Z_t at time t are modeled by a Markov process with a transition equation for $t = 2, 3, \ldots, T$:

$$\pi\big(Z_t | Z_{t-1}, \phi\big) = \prod_{i=1}^{n} N\big(z_{it} | z_{i(t-1)}, \sigma^2 I_R\big),$$

where the initial positions have independent $z_{i1} \sim N(0, \tau^2 I_R)$ for $i = 1, \ldots, N$ and $\phi = (\tau^2, \sigma^2, \beta_{IN}, \beta_{OUT}, r_{1:n})$ is the set of parameters of interest. To estimate the parameters ϕ, the authors adopt a Bayesan approach and implement MCMC algorithm using Metropolis-Hastings within Gibbs. They also perform a Procrustes transformation in order to re-orient the sampled trajectories.

This dynamic latent distance model is applied to two datasets. The Dutch classroom dataset includes directed relationships among 25 students, surveyed over four time points. The bill cosponsorship dataset consists of the cosponsorship history of 644 members of Congress who served during the 97th to 101st Congresses in the U.S. House of Representatives (five time points). The model provides useful insights into the networks, such as the overall gender and ethnic separation in the Dutch classroom example and the reflection of the political ideology of the latent positions in the bill cosponsorship example. In their follow-up work, Sewell and Chen [70] adjust the model for weighted edges including rank-ordered count data and non-negative continuous edges, via link functions and data augmentation. The applicability and interpretability of this extended model are demonstrated with mobile phone call log data and world exports/imports data, which are examples of count edges and continuous edges, respectively.

3.2.3.   Dynamic latent distance model for bipartite network—A bipartite graph is a graph in which vertices are divided into two groups and links only exist across groups. Friel et al. [19] develop a statistical model for bipartite temporal networks by extending the DSNL model, with different assumptions from the bipartite network model in Sakar et al. [68]. While Sakar et al. [68] model the co-occurrence of counts of authors and words from their empirical distribution, Friel et al. [19] directly model the evolution of edges through three Markov processes: one on the parameters, one on the latent positions, and one on the edges.

Define the binary edge $y_{ijt} = 1$ if node i in the first group ($i = 1, \ldots, N$) is connected to node j in the second group ($j = 1, \ldots, M$) at time t (e.g., director i sitting on board j in year t), and $y_{ijt} = 0$ otherwise. The chance of forming an edge at time t depends on the previous state $y_{ij(t-1)}$, and the latent distance between the two nodes in two different groups. The model assumes that

$$\text{logit}\left(p_{ijt}\right) = y_{ij(t-1)}\gamma_t + \left(1 - y_{ij(t-1)}\right)\beta_t - \left\| z_{it} - w_{jt} \right\|,$$

where γ_t is an edge persistence parameter and β_t is a non-edge persistence parameter, which are separately defined to capture the difference in the persistence of edges and non-edges, and z_{it} and w_{jt} are the R-dimensional latent positions of node i in the first group and j in the second group, respectively, at time t. Note that while the general framework assumes time-varying latent positions z_{it}, Friel et al. [19] fixed the first group's latent positions (i.e., z_i) for their application.

The latent positions in the first group are assumed to have independent Gaussian prior, while the latent positions in the second group are modeled using random walk processes (given the initial positions at $t = 1$):

$$P(Z) = \prod_{t=1}^{T} \prod_{i=1}^{N} \prod_{r=1}^{R} f\left(z_{irt}; 0, \tau_z\right),$$

$$P\left(W \middle| \tau_w, \tau_w^0\right) = \prod_{j=1}^{M} \prod_{r=1}^{R} f\left(w_{jr1}; 0, \tau_w^0\right) \times \prod_{t=2}^{T} \prod_{j=1}^{M} \prod_{r=1}^{R} f\left(w_{jrt}; w_{jr(t-1)}, \tau_w\right),$$

Where $\left(\tau_w, \tau_w^0, \tau_z\right)$ are the precision parameters. Similar to the latent positions of the second group, this model further assumes that the parameters γ_t and β_t also follow the Markov processes:

$$P\left(\gamma | \tau_\gamma, \tau_\gamma^0\right) = f\left(\gamma_1; 0, \tau_\gamma^0\right) \times \prod_{t=2}^{T} f\left(\gamma_t; \gamma_{t-1}, \tau_\gamma\right),$$

$$P\left(\beta | \tau_\beta, \tau_\beta^0\right) = f\left(\beta_1; 0, \tau_\beta^0\right) \times \prod_{t=2}^{T} f\left(\beta_t; \beta_{t-1}, \tau_\beta\right),$$

where $\left(\tau_\gamma, T_\gamma^0, T_\beta, \tau_\beta^0\right)$ are the precision parameters. Same as the inference procedures in other dynamic latent distance models, MCMC sampling is used to sample from the joint posterior distribution, and Procrustes transformation is applied for fixed orientation.

Using the model, the authors analyze the dynamic evolution of the leading Irish companies and their directors from 2003 to 2013. The network data contain $N = 761$ directors (first group), $M = 59$ companies (second group), and 3,855 edges. Mainly focused on understanding the persistence of links and the heterogeneity in the latent positions, the analysis reveals an increasing level of interlocking board behavior before and during the financial crisis, and stabilization thereafter.

3.3. Dynamic latent factor models

As mentioned before, the multiplicative form in the projection model [35] has been continuously studied under the name of "latent factor models". As shown in Hoff [29], latent factor models have some advantages over the latent distance model when both homophily—a type of pattern in which similar nodes may be more likely to attach to each other than dissimilar ones—and stochastic equivalence—a type of pattern in which the nodes can be divided into groups such that members of the same group have similar patterns of relationships—are present in the network. In this section, we demonstrate the development of dynamic latent factor models along with the motivating examples that highlight the models' capability of capturing higher-order dependence such as reciprocity—the tendency of one node to create a return link in a directed network—and stochastic equivalence.

3.3.1. Dynamic gravity model—Gravity models [49] in social sciences are used to model the bilateral relationships that can be predicted by the mass and distance of the pair. In modeling international trade and conflicts, Ward and Hoff [81] and Ward et al. [82] are the earlier works that combine the bilinear mixed-effects model (GBME) in Section 3.1.3 with the gravity models. However, their applications to temporal networks are done by

separately fitting the static model to each time point's network. Later, Ward et al. [80] improve their approach by including the previous year's fitted positions. The resulting model has the following form:

$$\log\left(y_{ijt}\right) = \beta' x_{ijt} + \beta_{uv} \hat{u}'_{i(t-1)} \hat{v}_{j(t-1)} + a_{it} + b_{jt} + u'_{it} v_{jt} + \epsilon_{ij},$$

where x_{ijt} consists of three covariates in the gravity models—the log of gross domestic product (GDP) of country i and j at time t, respectively, and the log of geographic distance between country i and j at time t. Different from the GBME, this model separates the latent positions to sending activity u and receiving activity v, and also includes the lagged position term $\beta_{uv} \hat{u}'_{i(t-1)} \hat{v}_{j(t-1)}$ to estimate the effect of positions at time $t-1$ on the edges at time t

This model is applied to bilateral trade data from 1990 to 2008. After several goodness of fit tests including the R^2 checks and out-of-sample predictions, the authors confirm the presence of strong second-order (i.e., reciprocity) and third-order dependencies in the world trade network. They demonstrate that the GBME model outperforms conventional ones in international trade literature, regarding predicting the observed values and understanding unobserved dyadic relationships.

3.3.2. Hierarchical multilinear model—By treating dynamic network data as an example, Hoff [31, 34] develops a general modeling framework for array data via reduced-rank decompositions, in which the array can be expressed as products of low-dimensional latent factors. Specifically for symmetric dynamic network data Y_t, the model takes the following form:

$$Y_t = X_t \beta + \Gamma_t,$$

$$\Gamma_t = Z \Lambda_t Z' + E_t,$$

where $Z = (z_1, ..., z_N)'$ is a matrix consisting of the time-invariant eigenvectors, assumed to have a matrix normal prior distribution, and $\Lambda_t = \mathrm{diag}(\lambda_{1t}, ..., \lambda_{Rt})$ is the time-varying eigenvalue matrix. This can be seen as a generalization of the latent factor models, using the reduced-rank approximations to matrices. Further extensions to arrays can be found in Hoff [33], with the longitudinal network serving as an example of the general model. This model is implemented as the R package "AMEN" [32].

In Hoff [34], the model is applied to analyze the international cooperation and conflict data in Ward et al. [82]. The dataset includes the records of militarized conflict and cooperation of 66 countries in every five years from 1950 to 1985, along with some economic and political characteristics of the countries. The model discovers that a majority of the conflicts over the cold war period involves economically large countries, and also reveals different conflict and cooperation patterns across countries over time.

3.3.3. Dynamic bilinear effects model

—Gaussian processes (GP's) have long been used in modeling temporally or spatially dependent data. Durante and Dunson [13, 15] propose a model that extends the bilinear component in Section 3.1.3 to the dynamic case by assuming the mean parameters and the latent factors are evolving in time via Gaussian processes. In particular, the model assumes

$$\eta_{ijt} = \mu_t + \beta'_t x_{ijt} + z'_{it} z_{jt},$$

where η is the linked mean parameter of a generalized linear model, μ_t is the mean process, and z_{it} and z_{jt} are the R-dimensional latent positions at time t. Here, the parameters μ, β, and z_{ir} (for $i = 1, \ldots, n$, and $r = 1, \ldots, k$) vary over time given their GP priors

$$\left(\mu_1, \ldots, \mu_T\right)' \sim N\left(0, c_\mu\right),$$

$$\left(z_{ir1}, \ldots, z_{irT}\right)' \sim N\left(0, \tau_r^{-1} c_z\right),$$

where c_μ and c_z are both $T \times T$ matrices constructed using the squared exponential covariance function—as a function of the distance between two time points t_1 and t_2. In other words,

$$c_\mu\left(t_1, t_2\right) = \exp\left\{-\kappa_\mu\left(t_1 - t_2\right)^2\right\},$$

$$c_z\left(t_1, t_2\right) = \exp\left\{-\kappa_z\left(t_1 - t_2\right)^2\right\},$$

$$\tau_r = \prod_{l=1}^{r} v_l, \quad v_1 \sim Ga\left(a_1, 1\right), \quad v_l \sim Ga\left(a_2, 1\right), \quad l \geq 2,$$

where κ_μ and κ_z are the length-scale parameters in GP, and τ_r is the variance parameter with hierarchical Gamma priors. The main advantage of this model is that it relaxes the Markovian assumption in most of the previously mentioned work and allows an unequal spacing of the observed time points.

Durante and Dunson [15] apply the method to analyze the co-movements of 23 National Stock Market Indices from 2004 to 2013. The results successfully reflect the global financial crisis and Greek debt crisis periods. The time-varying predictors help to discover the existence of international financial contagion effects and the opposite effects of verbal and material cooperation efforts on financial co-movements.

4. Stochastic blockmodels

In this section, we briefly summarize a statistical approach to *a posteriori* block modeling for networks introduced by [74] and [59] where the membership of nodes are unknown as well as the two relevant models, the mixed membership stochastic blockmodel [2] and the degree-corrected stochastic blockmodel [41], and then introduce several dynamic extensions of these static models.

4.1. Static stochastic blockmodels

The main assumption of the stochastic blockmodels is that the nodes of the network are partitioned into several latent classes (or blocks). To be specific, Snijders and Nowicki [74] and Nowicki and Snijders [59] assume the set of nodes $i \in \mathcal{N}$ is partitioned into C categories labeled $1, \ldots, C$. The class of node i is denoted by z_i and the classes are obtained in the vector $z = (z_i)_{i=1}^{N}$. The set of classes is denoted by $\mathcal{C} = \{1, \ldots, C\}$. The models further assume that the probability distribution of the edge between two nodes depends only on the classes to which they belong. Thus the edges are conditionally independent given the class vector.

Stochastic blockmodels inherit the philosophy of finite mixture models, and assume that the unobserved classes Z_i are *i.i.d.* random variables with the probability

$$P(Z_i = c) = \theta_c,$$

for class $c \in \mathcal{C}$. Therefore, the joint distribution of $Z = \{Z_1, \ldots, Z_N\}$ is defined by

$$P(Z_1 = z_1, \ldots, Z_N = z_N) = \theta_1^{m_1} \cdots \theta_C^{m_C},$$

where

$$m_c = \sum_{i=1}^{N} I(z_i = c)$$

denotes the number of vertices with class c.

The model for edges between the nodes depends on the node classes in the following ways. Given the vector of node classes $Z = z$, the random vectors Y_{ij} for $(i, j) \in \mathcal{N}$ with $i < j$ are independent, and the probabilities are

$$P(Y_{ij} = y | Z = z) = \eta_y(z_i, z_j),$$

where $y = (y_{i \to j}, y_{j \to i}) \in \mathcal{Y}$ is a vector of edge values and $\eta_y(z_i, z_j)$ is the class dependent edge probabilities which satisfy the condition

$$\sum_{y \in \mathscr{Y}} \eta_y(g, h) = 1 \quad \text{for all} \quad g, h \in \mathscr{C}.$$

The conditional distribution of Y given the vector of classes z is given in the following form

$$P(Y = y|z, \theta, \eta) = \left(\prod_{y \in \mathscr{Y}} \prod_{1 \le g < h \le C} \left(\eta_y(g, h)\right)^{e_y(g, h)}\right) \times \left(\prod_{y \in \mathscr{Y}} \prod_{c = 1}^{C} \left(\eta_y(c, c)\right)^{e_y(c, c)}\right),$$

where $e_y(g, h), y \in \mathscr{Y}, 1 \le g \le h \le C$, are the edge counts for block (g, h).

For the inference of the stochastic blockmodels, both Bayesian and frequentist approaches are proposed. In the Bayesian approach, the prior distribution for parameters is taken to be a product of Dirichlet distributions for the class distribution and the edges between and within classes of the given memberships. The posterior distribution is estimated using the Gibbs sampler. On the other hand, in the frequentist approach, several different algorithms are proposed to estimate the classes: Rohe et al. [65] propose the spectral clustering to estimate classes in the model, Guédon and Vershynin [22] and Amini and Levina [4] write the class estimation problem as a semidefinite optimization problem and find the solution, and Amini et al. [3] propose a pseudo-likelihood algorithm which provides consistent estimates of classes.

Despite the simplicity in model formulation, stochastic blockmodels provide a powerful tool in modeling networks. They allow one to represent the effect of unobserved heterogeneity of individual positions or preferences on the pattern of pairwise relations. Since the heterogeneity is modeled by the stochastic membership of the c classes, regarding cluster analysis, it is more similar to a mixture model rather than a discrete classification model.

4.1.1.   Mixed membership stochastic blockmodel—Many real-world networks are multi-faceted. However, stochastic blockmodels suffer from a limitation that each node can only belong to one group, or in other words, play a single latent role. To overcome this issue, Airoldi et al. [2] relax the assumption of a single latent role for nodes and develop the mixed membership stochastic blockmodel.

In this paper [2], the authors focus on directed networks and assume the observed network is generated according to node-specific distributions of community membership and edge-specific indicator vectors denoting the membership in one of the C communities. Each node is associated with a randomly drawn vector $\overrightarrow{\pi}_i$ for node i, where $\pi_{i,c}$ denotes the probability of node i belonging to group c. That is, each node can simultaneously belong to multiple groups with different degrees of affiliation. The probabilities of edges between different groups are defined by the matrix of Bernoulli rates $B_{C \times C}$, where $B(g, h)$ rep-resents the probability of having an edge between a node from group g and a node form group h. The mixed membership stochastic blockmodel posits that the $\{Y_{ij}\}_{1 < i, j < N}$ are drawn from the following generative process.

•   For each node $i = 1, \ldots, N$:

- Draw a C dimensional mixed membership vector $\overrightarrow{\pi}_i \sim \text{Dirichlet}(\overrightarrow{\alpha})$.

- For each possible edge variable Y_{ij}:

  - Draw membership indicator for the initiator $\overrightarrow{z}_{i \to j} \sim \text{Multinomial}(\overrightarrow{\pi}_i)$.

  - Draw membership indicator for the receiver $\overrightarrow{z}_{i \leftarrow j} \sim \text{Multinomial}(\overrightarrow{\pi}_j)$.

  - Sample the edge $Y_{ij} \sim \text{Bernoulli}(\overrightarrow{z}'_{i \to j} B \overrightarrow{z}_{i \leftarrow j})$.

Here, the indicator vector $\overrightarrow{z}_{i \to j}$ denotes the group membership of node i when node i has out-going edge to node j, and $\overrightarrow{z}_{j \to i}$ denotes the group membership of node j when node j has out-going edge to node i.

Under the mixed membership stochastic blockmodel assumptions, the joint probability of the data Y and the latent variables $\{\overrightarrow{\pi}_{1:N}, Z_{\to}, Z_{\leftarrow}\}$ can be written the in following form:

$$P(Y, \overrightarrow{\pi}_{1:N}, Z_{\to}, Z_{\leftarrow} | \overrightarrow{\alpha}, B)$$

$$= \prod_{(i,j) \in \mathcal{N}} P(Y_{ij} | \overrightarrow{z}_{i \to j}, \overrightarrow{z}_{i \leftarrow j}, B) P(\overrightarrow{z}_{i \to j} | \overrightarrow{\pi}_i) P(\overrightarrow{z}_{i+j} | \overrightarrow{\pi}_j) \times \prod_{i=1}^{N} P(\overrightarrow{\pi}_i | \overrightarrow{\alpha})$$

where $\overrightarrow{\pi}_{1:N}$ is the set of mixed membership vectors, and $Z_{\to}$ and $Z_{\leftarrow}$ are the sets of membership indicator vectors.

A nested variational inference algorithm is used for posterior inference on the per-node mixed membership vectors and per-pair roles. To compute the empirical Bayes estimates of the model parameters, variational expectation-maximization (EM) algorithm is used. Moreover, in recent years, Mao et al. [52] and Jin et al. [40] propose a consistent algorithm for inferring mixed membership of nodes in the mixed membership stochastic blockmodel.

4.1.2. Degree-corrected stochastic blockmodel—Although the stochastic blockmodels have been popularly used as a tool for detecting community structure in networks, they fail to capture the heterogeneity of node degrees (i.e., number of edges the node has to other nodes) within communities which often observed in real-world networks. To solve this problem, Karrer and Newman [41] relax the assumption that the stochastic blockmodels treat all nodes within a community as stochastically equivalent, and propose the degree-corrected stochastic blockmodel that can consider node covariates.

This model focuses on undirected networks and allows networks to contain both multi-edges and self-edges, even though many real-world networks have no self-connecting edges. Let Y be an undirected multigraph on n nodes, possibly including self-edges, and let Y_{ij} be an element of the adjacency matrix of the multigraph. Also, there is a new set of parameters θ_i controlling the expected degrees of nodes i. The model assumes that the number of edges between each pair of nodes is independently Poisson distributed, and defines $\theta_i \theta_j \omega_{gh}$ to be

the expected value of the adjacency matrix element Y_{ij} for nodes i and j lying in groups $g \in \mathscr{C}$ and $h \in \mathscr{C}$, respectively. Then, Y has the probability

$$P(Y|\theta, \omega, z) = \prod_{i < j} \frac{\left(\theta_i \theta_j \omega_{z_i z_j}\right)^{Y_{ij}}}{Y_{ij}!} \exp\left(-\theta_i \theta_j \omega_{z_i z_j}\right)$$

$$\times \prod_i \frac{\left(\frac{1}{2}\theta_i^2 \omega_{z_i z_j}\right)^{Y_{ij}/2}}{\left(Y_{ij}/2\right)!} \exp\left(-\frac{1}{2}\theta_i^2 \omega_{z_i z_j}\right).$$

To estimate the parameters and infer the membership of nodes, the authors suggest a novel method where the log-likelihood is maximized in two stages. First, they find maximum likelihood values of model parameters θ_i and ω_{gh}. Then, they propose information-theoretic quantities for community detection or clustering. Similar to Rohe et al. [65], Qin and Rohe [61] propose spectral clustering under the degree-corrected stochastic blockmodel. Also, Zhao et al. [87] generalize the consistency framework of the stochastic blockmodel to the degree-corrected stochastic blockmodel and obtain a general theorem for community detection consistency.

4.2. Dynamic stochastic blockmodels

4.2.1. Dynamic stochastic blockmodel with varying community memberships
—To analyze dynamic communities, Yang et al. [86] propose a model that captures the evolution of communities by explicitly modeling the transition of community memberships for individual nodes in the network.

Let Y_t be the snapshot of a network at a given time step t, with N the number of nodes in the network. Each element y_{ijt} in Y_t is the weight assigned to the edge between nodes i and j. Although this dynamic stochastic blockmodel can handle both the frequency of interactions (i.e., a natural number) and the binary number indicating presence or absence of interactions, the authors' main focus is on the binary edges. For a dynamic network, they use $\mathscr{Y}_T = \{Y_1, ..., Y_T\}$ to denote a collection of snapshots for the network over T discrete time steps. They also use $z_i \in \{1, ..., C\}$, where C is the total number of communities, to denote the community membership of node i. In addition, they introduce $z_{ic} = [z_i = c]$ to indicate if node i is in the cth community where $[x]$ outputs 1 if x is true and zero otherwise. Community matrix $Z_t = (z_{ict}: i \in \{1, ..., N\}, c \in \{1, ..., C\})$ then indicates the community assignments of all nodes in network at a given time step t. Finally, they set $\mathscr{Z}_T = \{Z_1, ..., Z_T\}$ to represent the collection of community assignments of all nodes over T time steps.

Assuming that the community matrix Z_{t-1} for time step $t-1$ is given, the model uses a transition matrix $A \in \mathbb{R}^{C \times C}$ to model the community matrix Z_t at time step t. Moreover, $\pi = \{\pi_1, ..., \pi_C\}$ is used as the initial probability where π_c is the probability for a node to be assigned to community c. Given the community memberships in Z_t, the edge between nodes

is determined stochastically by probabilities P. The joint probability of the data $\mathscr{Y}_T$ and the latent variable $\mathscr{Z}_T$ can be written in the following form:

$$P\left(\mathscr{Y}_T, \mathscr{Z}_T | \pi, P, A\right) = \prod_{t=1}^{T} P\left(Y_t | Z_t, P\right) \prod_{t=2}^{T} P\left(Z_t | Z_{t-1}, A\right) P\left(Z_1 | \pi\right)$$

where

$$P\left(Y_t | Z_t, P\right) = \prod_{i \sim j} P\left(y_{ijt} | z_{it}, z_{jt}, P\right)$$

$$= \prod_{i \sim j} \prod_{g, h \in \mathscr{C}} \left(P_{gh}^{y_{ijt}} \left(1 - P_{gh}\right)^{1 - y_{ijt}}\right)^{z_{igt} z_{jht}},$$

$$P\left(Z_t | Z_{t-1}, A\right) = \prod_{i=1}^{N} P\left(z_{it} | z_{i(t-1)}, A\right)$$

$$= \prod_{i=1}^{N} \prod_{g, h \in \mathscr{C}} A_{kl}^{z_{ig(t-1)} z_{iht}},$$

and

$$P\left(Z_1 | \pi\right) = \prod_{i=1}^{N} \prod_{c=1}^{C} \pi_c^{z_{ic1}},$$

respectively. Note that in their model self-loops are not considered and so in the above equations $i \sim j$ means over all i's and j's such that $i \neq j$.

For the inference of the model, the authors introduce a point estimation approach using a Variational EM algorithm. They also propose a method based on a combination of probabilistic simulated annealing algorithm and Gibbs sampling algorithm to infer the parameters.

This model is applied to three datasets. The southern women dataset records the attendance of 18 women to 14 social events over a 9-month period in 1930's in Natchez, Mississippi, as part of their work to study social class in both black and white societies. The blog dataset is collected by the NEC labs and includes 148,681 entry-to-entry links among 407 blogs during 15 months. The paper co-authorship (DBLP) dataset contains the co-authorship information of papers in 28 conferences in three areas (data mining (DM), database (DB), and artificial intelligence (AI)) over ten years (1997–2006). For the first two datasets, the model can detect the change of community membership over the study period. In the co-authorship dataset, the model discovers the trend that, over time, the community of DB gets smaller, the community of DM gets larger, and the community of AI remains relatively

stable. Moreover, they also find some highly productive researchers had many changes in community membership.

4.2.2. Dynamic stochastic blockmodel with varying community memberships and connectivity parameters

Both Xu and Hero [85] and Matias and Miele [53] propose methods which relax the constraint of fixed connectivity probabilities P of Yang et al. [86] and allow both community memberships and connectivity parameters vary over time. The main difference between the two papers is that the former entirely relaxes the constraint of fixed connectivity parameters while the latter keeps a weak constraint on the connectivity parameters to handle the label switching issues across different time steps. Here we first introduce Xu and Hero [85] followed by Matias and Miele [53].

Xu and Hero [85] propose a state space model through time on the probability of connection between groups and focus on directed networks with no self-edges. Let Y_t denote the adjacency matrix of the network observed at time step t. Let $\mathscr{Y}_t$ denote the set of all snapshots up to time t, i.e., $\{Y_1, \ldots Y_t\}$ The notation $i \in c$ indicates that node i belong class c and the classes of all nodes at time t is given by a vector z_t with $z_{it} = c$ if $i \in c$ at time t. The $\Theta_t = [\theta_{ght}]$ denotes the connectivity parameter between groups, where θ_{ght} denotes the probability of forming an edge between a node in class g and a node in class h, and C denotes the number of classes. The set Θ_t can be viewed as the states of a dynamic system that generates the noisy observation sequence. Since θ_{ght} is a probability that must be bounded between 0 and 1, they work with the $\psi_{ght} = \log(\theta_{ght}) - \log(1 - \theta_{ght})$, which is the logit of θ_{ght}.

The following linear dynamic system gives a model for the state evolution

$$\psi_t = F_t \psi_{t-1} + v_t$$

where F_t is the state transition model applied to the previous state, ψ_t is the vector representation of the $[\psi_{ght}] \in \mathbb{R}^{C \times C}$, and v^t is a random vector of zero-mean Gaussian entries, commonly referred to as process noise, with covariance matrix Γ_t. They assume Γ_t to be time-invariant and not necessarily diagonal because states could evolve in a correlated manner.

Since the class memberships z_t are not known and should be estimated along with ψ_t, the label-switching methods are used as in [41], [87], and maximize the posterior state density given the entire sequence of observations $\mathscr{Y}_t$ up to time t, in order to account for the prior information. The posterior state density is given by

$$P(\psi_t | \mathscr{Y}_t) \propto P(Y_t | \psi_t, \mathscr{Y}_{t-1}) P(\psi_t | \mathscr{Y}_{t-1}).$$

Their inference is based on on-line iterative estimation procedure alternating two steps. First, a label-switching method is used to explore the space of node group configurations. Then,

they use extended Kalman filter (EKF) that optimizes the likelihood when the group memberships are known.

This model is applied to two datasets. The MIT reality mining dataset records cell phone activity of 94 students and staff at MIT over a year. They use the participant affliations as ground-truth class memberships and compare the class estimation accuracy of their model to static stochastic blockmodel. The Enron email dataset consists of about 500,000 email messages of 184 Enron employees from 1988 to 2002. The model discovers a steady increasing trend in edge probabilities from Enron CEOs to presidents as Enron's financial situation worsened. On the other hand, edge probabilities between other employees remained at their baseline levels until Enron fell under federal investigation.

Matias and Miele [53] focus on detecting communities characterized by a stable within-group connectivity behavior by adding some constraints on the varying connectivity parameter. They consider weighted interactions between N nodes recorded over time in a set of data matrices $Y = (Y_t)_{1 \leq t \leq T}$. For each $t \in \{1, \ldots, T\}$, the adjacency matrix $Y_t = (Y_{ijt})_{1 \leq i \leq j \leq N}$ contains real-values measuring interactions between the nodes $i, j \in \{1, \ldots, N\}^2$. Their model can also consider undirected networks without self-loops. They assume that the N nodes are split into C latent communities that vary through time, as encoded by the random variables $Z = (Z_{it})_{1 \leq t \leq T, 1 \leq i \leq N}$. They use independent Markov chains to model the evolution of the nodes' memberships over time. For each node i, the process $(Z_{it})_{1 \leq t \leq T}$ is an irreducible, aperiodic stationary Markov chains with transition matrix $A = (A_{gh})_{1 \leq g,h \leq C}$ and initial stationary distribution $a = (a_1, \ldots, a_C)$. They consider Z_{it} as a random vector $Z_{it} = (Z_{i1t}, \ldots, Z_{iCt}) \in \{0, 1\}^C$ constrained to $\sum_{c=1}^{C} Z_{ict} = 1$.

To take account of possible sparse weighted networks, they assume that

$$Y_{ijt} \mid \{Z_{igt} Z_{jht} = 1\} \sim \left(1 - \beta_{ght}\right) \delta_0(\cdot) + \beta_{ght} F\left(\cdot, \gamma_{ght}\right),$$

where $g, g, h \in \mathscr{C}$ and $\{F(\cdot, \gamma), \gamma \in \Gamma\}$ is a parametric family of distribution with no mass at 0 with its density denoted by $f(\cdot, \gamma)$. Here $\beta_t = (\beta_{ght})_{1 \leq g,h \leq C}$ is the sparsity parameter which satisfies $\beta_{ght} \in [0, 1]$. Moreover, γ_{ght} is the connectivity parameter which depends on the choice of the parametric family. They also let $\phi(\cdot; \beta, \gamma)$ denote the density of the distribution.

The joint probability of the data Y and the latent variable Z can then be written in the following form:

$$P(Y, Z \mid A, \beta, \gamma) = \prod_{i=1}^{N} \prod_{c=1}^{C} \alpha_c^{Z_{ic1}} \prod_{t=2}^{T} \prod_{i=1}^{N} \prod_{1 \leq g,h \leq C} A_{gh}^{Z_{ig(t-1)} Z_{iht}}$$
$$\times \prod_{t=1}^{T} \prod_{1 \leq i < j \leq N} \prod_{1 \leq g,h \leq C} \phi\left(Y_{ijt}; \beta_{ght}, \gamma_{ght}\right)^{Z_{igt} Z_{jht}}.$$

To infer the model parameters and to cluster the nodes, the variational expectation-maximization (VEM) algorithm is employed. Furthermore, they propose an integrated classification likelihood (ICL) criterion for estimating the number of communities.

This model is applied to reveal social structure and organizations in a French high school and two animal interaction datasets. The French high school dataset consists of face-to-face encounters of 31 students in the class during four days in December 2011. The model finds four distinct groups showing different interaction patterns. The model also discovers the evidence of some gender homophily. The migratory birds (sparrows) dataset is composed of 3 time steps with 69 birds in total. The model successfully confirms the analysis in Shizuka et al. [71] that the sparrow community is stable. The model is also applied to the Indian equids (onagers) dataset by aggregating interactions of 23 onagers monthly from February 2003 to May 2003 and reveals a hierarchical social integration process. On the other hand, under the setting of constant community membership across different time steps and varying connectivity parameters, Bhattacharyya and Chatterjee [8] propose a spectral clustering algorithm for dynamic stochastic blockmodel which guarantees consistency of community detection.

4.2.3. Dynamic mixed membership stochastic blockmodel—Xing et al. [84] and Ho et al. [26] propose dynamic extensions of the static mixed membership stochastic blockmodel using a state space model for the time-varying parameters of the priors, both for the mixed membership vector of a node and the connectivity behavior. Their methods model the role of each node as a dynamic mixed membership vector that allows nodes to behave differently over time as well as carry out different roles when interacting with different nodes.

Xing et al. [84] consider a temporal series of networks $Y = \{ Y_1, \ldots, Y_T \}$ where $Y_t = \left\{ y_{ijt} \right\}_{i, j = 1}^{N}$ is the set of edges at time t between a fixed set of N nodes. In the static mixed membership stochastic blockmodels, Airoldi et al. [2] employ a simple Dirichlet prior because it is conjugate to the multinomial distribution over every latent membership label $\left\{ \overrightarrow{z}_i \rightarrow ., \overrightarrow{z}_i \leftarrow . \right\}$ defined by relevant $\overrightarrow{\pi}_i$. However, to model temporal dynamics of the node roles and capture the non-trivial correlations between different roles at the same time, the authors employ the logistic normal distribution. The logistic normal distribution, $LN(\mu_t, \Sigma_t)$, is used for the prior for the mixed membership vectors, $\overrightarrow{\pi}_{it}$ and another logistic normal distribution, $LN(\eta_t, S_t)$ is used for the prior for the entries of the connectivity matrix, β_{ght}.

Their basic model structure is based on the state space model, which defines a linear dynamic transformation of the mixed membership priors over adjacent time points:

$$\overrightarrow{\mu}_t = A\overrightarrow{\mu}_{t-1} + \overrightarrow{\omega}_t,$$

where $\overrightarrow{\pi}_t$ represents the mean parameter of the prior distribution of the transformed mixed membership vectors of all vertices at time t, and $\overrightarrow{\omega}_t \sim N(0, \Phi)$ represents normal transition

noise for the mixed membership prior, and the transition matrix A shapes the trajectory of temporal transformation of the prior. Also, the model makes a similar assumption for η_t such that

$$\eta_t = b\eta_{t-1} + \xi_t,$$

where b is scalar and $\eta_t \sim N(0, \psi)$.

The dynamic mixed membership stochastic blockmodel thus consists of three components: a state space model for the mixed membership vector, a state space model for entries of connectivity matrix, and a logistic normal mixed membership stochastic blockmodel for the networks. The first two components explain the temporal dynamics while the third component models the generative process of the network at each time point. Following is an outline of the generative process.

- State space model for mixed membership prior:

    - $\overrightarrow{\mu}_1 \sim N(\nu, \Phi)$, sample the means of the mixed membership prior at $t = 1$.

    - $\overrightarrow{\mu}_t \sim N(A\overrightarrow{\mu}_{t-1}, \Phi)$, sample the means of the mixed membership prior at time points $t = 2, \ldots, T$.

- State space model for connectivity matrix:

    - $\eta_{gh1} \sim N(\iota, \psi)$, sample the means of the entries of connectivity matrix prior between g and h at $t = 1$.

    - $\eta_{ght} \sim N(b\eta_{gh(t-1)}, \psi)$, sample the means of the entries of connectivity matrix prior between g and h at $t = 2, \ldots, T$.

    - $\beta_{ght} \sim LN(\eta_t, S_t)$, sample the entries of connectivity matrix between g and h at $t = 1, 2, \ldots, T$.

- Logistic normal mixed membership stochastic blockmodel:

    - $\overrightarrow{\pi}_{it} \sim LN(\overrightarrow{\mu}_t, \Sigma_t)$, sample a C-dimensional mixed membership vector for each node $i = 1, \ldots, N$, at $t = 1, 2, \ldots, T$.

    - $\overrightarrow{z}_{i \rightarrow j,t} \sim \text{Multinomial}(\overrightarrow{\pi}_{it})$, sample membership indicator for the initiator for each node $i = 1, \ldots, N$, at $t = 1, 2, \ldots, T$.

    - $\overrightarrow{z}_{i \leftarrow j,t} \sim \text{Multinomial}(\overrightarrow{\pi}_{jt})$, sample membership indicator for the receiver for each node $i = 1, \ldots, N$, at $t = 1, 2, \ldots, T$.

    - $e_{ijt} \sim \text{Bernoulli}(\overrightarrow{z}_{i \rightarrow j,t}' B_t \overrightarrow{z}_{i \leftarrow j,t})$, sample the edges between nodes i and j for $1 < i \quad j < N$, at $t = 1, 2, \ldots, T$.

The joint probability of the data Y and the latent variables $\left\{\overrightarrow{\pi}_{1:N,t}, Z_{\rightarrow,t}, Z_{\leftarrow,t}\right\}$ can be written in similar form as the static mixed membership stochastic blockmodel. For the posterior inference, a Laplace variational approximation scheme based on the generalized

means field (GMF) approximation is used to infer the latent variables and estimate the model parameters.

This model is applied to three datasets. The Sampson's monk dataset contains the liking relationship among 18 monks over three time points. The result of the model is consistent with previous works except for one controversial person, Mark. The model is also applied to a subset of Enron email dataset consisting of 151 persons from 2001 and discovers five distinct roles of actors. They visualize and track the trajectory of the mixed membership vector for an individual to understand how each evolves in his or her role. The third example studies the evolving gene network of fruit flies. The Drosophila melanogaster gene network dataset contains 22 networks at different time points across various developmental stages. The model discovers that many genes exhibit a sharp transition regarding their roles near the end of the embryonic stage. They further examine 45 ontological groups and find an overall pattern that each role consists of genes with a variety of functions, and the functional composition of each role varies across time.

Different from Xing et al. [84], which employs a time-evolving logistic normal distribution on all networks nodes, Ho et al. [26] generalize the prior on nodes to be a mixture of time-evolving logistic normal distributions. This mixture prior is multi-modal and captures correlations between roles, allowing it to fit complex data densities that the unimodal prior cannot. The state space model, which defines a linear dynamic transformation of the mixed membership priors, now contains N distinct trajectories over adjacent time points,

$$\overrightarrow{\mu}_{ct} = A\overrightarrow{\mu}_{c(t-1)} + \overrightarrow{\omega}_{ct}, \quad c = 1, 2, ..., C$$

where A is a transition matrix and $\omega_{ct} \sim N(0, \Phi)$ represents the normal transition noise. Now each mixed membership vector $\overrightarrow{\pi}_{it}$ is drawn from one of the C trajectories $\overrightarrow{\mu}_{ct}$. The choice of trajectory for $\overrightarrow{\pi}_{it}$ is given by the indicator z_{it}, which is drawn from some prior. For simplicity, they use a single multinomial prior with the parameter δ for all z_{it}. Notice that z_{it} can change over time, allowing nodes to switch clusters if that would fit the data better. Given z_{it}, the mixed membership vector $\overrightarrow{\pi}_{it}$ is sampled from $LN\left(\overrightarrow{\mu}_{z_{it}, t}, \Sigma_{z_{it}, t}\right)$. Once $\left\{\overrightarrow{\pi}_{it}\right\}_{i=1}^{N}$ have been drawn for some timepoint t, the remaining part follows the generative process of Xing et al. [84]. They also propose a similar variational EM algorithm for parameter learning and latent variable inference.

The authors analyze the United States 109th Congress voting records by generating time-varying networks containing 100 senators and eight time points corresponding to 3-month periods from Jan 1st, 2005 to Dec 31st, 2006. The analysis not only finds the two clusters capturing party affliations but also identifies outliers, for example, a senator who shifted from the Democrat cluster to the Republican cluster.

5. Conclusions and discussions

The literature on statistical network modeling may be divided into various lines of research, such as the development of network statistics for ERGMs, inference methods for SAOMs, and asymptotic network properties in physics-based models. Among various possibilities, in this paper, we mainly focus on the review of dynamic network models with latent variables —models that assume and estimate latent structures of networks—in detail. Latent space models assume the existence of unobserved network embedding into a low-dimensional space, and stochastic blockmodels assume that the nodes within same blocks are structurally equivalent. From these, we make a distinction between the latent space models (LSM) and stochastic blockmodels (SBM) and outline the evolution of each class of models from static to dynamic models, as well as their explicit connections. As a review of network models, to conclude, we provide two mini directed networks to summarize the relationships of the models discussed in previous sections, with latent space models in Figure 1 and stochastic blockmodels in Figure 2. In both figures, static models are in italic letters, with the names of nodes corresponding to the title of each section. A directed link indicates the sender model motivates the development of the receiver model, or the receiver is a dynamic extension of the sender.

Despite recent advances in statistical and computational methods of dynamic network modeling, there remain several challenges across the models described in previous sections. Some (e.g., computation, visualization) are challenges for network modeling in general, while others (e.g., lack of continuous-time models, identifiability) are only applicable to dynamic latent space and/or stochastic blockmodels. In this section, we briefly sketch out the three major issues in modeling dynamic networks with latent variables.

Computation, the most common type of problem for network modeling in general, is particularly an issue for latent variable models due to the estimation of the additional parameters. For example, the evaluation of likelihood equations for static and dynamic latent space models require the calculation of a distance matrix given the complete set of latent positions at a computational and storage cost of $O(N^2)$ [63] and $O(T \times N^2)$, respectively. This makes the estimation of latent space models impractical for networks larger than a few thousand nodes, and as a result, most of their applications have been limited to relatively small networks, typically less than 500 nodes. As the existing software packages for other well-known frameworks including temporal ERGMs [24, 45, 48, 1] or SAOMs [72, 75] can efficiently fit networks of size up to a few thousand nodes, there is a strong need for the development of efficient statistical software to fit these models. Recently, new estimation methods [51, 62, 63, 64] are considered for latent space models, so we expect a bright prospect to achieve a substantial reduction in computing time. In the case of stochastic blockmodels, the computation is especially non-trivial for large networks, since the optimization over all possible label assignments is NP-hard (i.e., non-deterministic polynomial-time hardness). With the help of variational methods [9], stochastic blockmodels can be fit for thousands of nodes. It is still a challenge to fit blockmodels to large networks with millions or tens of millions of nodes.

In general, latent variable models allow for better understanding of networks that are believed to contain hidden structure. As such, visualization is one of the biggest advantages of those models due to their model-based nature, such as Euclidean distances between nodes and clustering of the nodes. At the same time, there also remain limitations and open problems in visualizing the latent variables. For example, latent space models are usually visualized in 2D or 3D plots, but not in higher dimensions. Furthermore, when a dynamic network spans a large number of time points (e.g., T 50) or has a continuous-time scale, it is diffcult to create plots that effectively show the movement of latent positions.

Finally, motivated by the rise of large-scale online social networks, a variety of continuous-time network models have been introduced, mostly built upon survival and event history analysis [38, 60, 79]. In contrast, most of the dynamic latent variable models introduced in this paper focus on learning discrete-time networks or the discrete realization of a continuous process [15], where multiple cross-sectional "snapshots" of the network are recorded at discrete time points. One exception is the dynamic bilinear effects model in Section 3.3.3, however, the model's illustrated example is still the discrete (quarterly) network data. Considering the growing amount of network data recorded on a continuous-time scale, developing the latent variable models that take advantage of this temporal granularity might be a promising future research direction.

Acknowledgments

Appendix:: Dynamic network data sources

We provide a list of data sets mentioned in this paper and the current version of data sources. In case of multiple sources, we refer to multiple websites or a specific section of the paper in which the data sources are illustrated in detail.

References

[1]. Agarwal A, Lee K, and Xue L (2018). Temporal exponential-family random graph models with time-evolving latent block structure for dynamic networks. Technical report, Penn State Univeristy.

[2]. Airoldi EM, Blei DM, Fienberg SE, and Xing EP (2008). Mixed membership stochastic blockmodels. Journal of Machine Learning Research, 9(Sep):1981–2014. [PubMed: 21701698]

[3]. Amini AA, Chen A, Bickel PJ, and Levina E (2013). Pseudo-likelihood methods for community detection in large sparse networks. The Annals of Statistics, 31(4):2097–2122.

[4]. Amini AA and Levina E (2018). On semidefinite relaxations for the block model. The Annals of Statistics, 46(1):149–179.

[5]. Arbeitman MN, Furlong EE, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, and White KP (2002). Gene expression during the life cycle of drosophila melanogaster. Science, 297(5590):2270–2275. [PubMed: 12351791]

[6]. Asur S and Parthasarathy S (2007). An event-based framework for characterizing the evolutionary behavior of interaction graphs. Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 913–921.

[7]. Banfield J and Raftery A (1993). Model-based gaussian and non-gaussian clustering. Biometrics, 49:803–821.

[8]. Bhattacharyya S and Chatterjee S (2017). Spectral clustering for dynamic stochastic block model. Working Paper.

[9]. Blei DM, Kucukelbir A, and McAuliffe JD (2017). Variational inference: A review for statisticians. Journal of the American Statistical Association, to appear.

[10]. Chang J and Blei DM (2009). Relational topic models for document networks. In International Conference on Artificial Intelligence and Statistics, pages 81–88.

[11]. Chi Y, Song X, Zhou D, Hino K, and Tseng BL (2007). Evolutionary spectral clustering by incorporating temporal smoothness. In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 153–162.

[12]. Desmarais BA and Cranmer SJ (2017). Statistical inference in political networks research. arXiv preprint arXiv:1703.02870.

[13]. Durante D and Dunson DB (2014a). Bayesian dynamic financial networks with time-varying predictors. Statistics & Probability Letters, 93:19–26.

[14]. Durante D and Dunson DB (2014b). Bayesian logistic gaussian process models for dynamic networks. In Artificial Intelligence and Statistics, pages 194–201.

[15]. Durante D and Dunson DB (2014c). Nonparametric bayes dynamic modeling of relational data. Biometrika, 101(4):883–898.

[16]. Eagle N, Pentland AS, and Lazer D (2009). Inferring friendship network structure by using mobile phone data. Proceedings of the National Academy of Sciences, 106(36):15274–15278.

[17]. Fletcher RJ, Acevedo MA, Reichert BE, Pias KE, and Kitchens WM (2011). Social network models predict movement and connectivity in ecological landscapes. Proceedings of the National Academy of Sciences, 108(48):19282–19287.

[18]. Freeman LC (2003). Finding social groups: A meta-analysis of the southern women data. New York: National Academies Press.

[19]. Friel N, Rastelli R, Wyse J, and Raftery AE (2016). Interlocking directorates in irish companies using a latent space model for bipartite networks. Proceedings of the National Academy of Sciences, 113(24):6629–6634.

[20]. Girvan M and Newman ME (2002). Community structure in social and biological networks. Proceedings of the National Academy of Sciences, 99(12):7821–7826.

[21]. Goldenberg A, Zheng AX, Fienberg SE, Airoldi EM, et al. (2010). A survey of statistical network models. Foundations and Trends R in Machine Learning, 2(2):129–233.

[22]. Guédon O and Vershynin R (2016). Community detection in sparse networks via grothendieck's inequality. Probability Theory and Related Fields, 165(3–4):1025–1049.

[23]. Handcock MS, Raftery AE, and Tantrum JM (2007). Model-based clustering for social networks. Journal of the Royal Statistical Society: Series A (Statistics in Society), 170(2):301–354.

[24]. Hanneke S, Fu W, Xing EP, et al. (2010). Discrete temporal models of social networks. Electronic Journal of Statistics, 4:585–605.

[25]. Henry T, Banks D, Chai C, and Owens-Oas D (2016). Modeling community structure and topics in dynamic text networks. arXiv preprint arXiv:1610.05756.

[26]. Ho Q, Song L, and Xing E (2011). Evolving cluster mixed-membership blockmodel for time-varying networks. Journal of Machine Learning Research: Workshop and Conference Proceedings, 15:342–350.

[27]. Hoff P, Fosdick B, Volfovsky A, and Stovel K (2013). Likelihoods for fixed rank nomination networks. Network Science, 1(3):253–277. [PubMed: 25110586]

[28]. Hoff PD (2005). Bilinear mixed-effects models for dyadic data. Journal of the American Statistical Association, 100(469):286–295.

[29]. Hoff PD (2008). Modeling homophily and stochastic equivalence in symmetric relational data. In Advances in Neural Information Processing Systems, pages 657–664.

[30]. Hoff PD (2009). Multiplicative latent factor models for description and prediction of social networks. Computational and Mathematical Organization Theory, 15(4):261–272.

[31]. Hoff PD (2011). Hierarchical multilinear models for multiway data. Computational Statistics & Data Analysis, 55(1):530–543.

[32]. Hoff PD (2015a). Dyadic data analysis with amen. arXiv preprint arXiv:1506.08237.

[33]. Hoff PD (2015b). Multilinear tensor regression for longitudinal relational data. The Annals of Applied Statistics, 9(3):1169. [PubMed: 27458495]

[34]. Hoff PD et al. (2011). Separable covariance arrays via the tucker product, with applications to multivariate relational data. Bayesian Analysis, 6(2):179–196.

[35]. Hoff PD, Raftery AE, and Handcock MS (2002). Latent space approaches to social network analysis. Journal of the American Statistical Association, 97(460):1090–1098.

[36]. Hoff PD and Ward MD (2004). Modeling dependencies in international relations networks. Political Analysis, 12(2):160–175.

[37]. Holland PW, Laskey KB, and Leinhardt S (1983). Stochastic block-models: First steps. Social Networks, 5(2):109–137.

[38]. Hunter D, Smyth P, Vu DQ, and Asuncion AU (2011). Dynamic egocentric models for citation networks. In Proceedings of the 28th International Conference on Machine Learning (ICML-11), pages 857–864.

[39]. Ji P and Jin J (2016). Coauthorship and citation networks for statisticians. The Annals of Applied Statistics, 10(4):1779–1812.

[40]. Jin J, Ke ZT, and Luo S (2017). Estimating network memberships by simplex vertex hunting. arXiv preprint arXiv:1708.07852.

[41]. Karrer B and Newman MEJ (2011). Stochastic blockmodels and community structure in networks. Physical Review E, 83(1):016107.

[42]. Kolaczyk ED (2009). Statistical Analysis of Network Data: Methods and Models. Springer Science & Business Media.

[43]. Krafft P, Moore J, Desmarais B, and Wallach HM (2012). Topic-partitioned multinetwork embeddings. In Advances in Neural Information Processing Systems, pages 2807–2815.

[44]. Krivitsky PN and Handcock MS (2008). Fitting position latent cluster models for social networks with latentnet. Journal of Statistical Software, 24(5).

[45]. Krivitsky PN and Handcock MS (2014). A separable model for dynamic networks. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 76(1):29–46. [PubMed: 24443639]

[46]. Krivitsky PN and Handcock MS (2015). latentnet: Latent Position and Cluster Models for Statistical Networks. The Statnet Project (http://www.statnet.org). R package version 2.7.1.

[47]. Krivitsky PN, Handcock MS, Raftery AE, and Hoff PD (2009). Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. Social Networks, 31(3):204–213. [PubMed: 20191087]

[48]. Lee K, Xue L, and Hunter DR (2017). Model-based clustering of time-evolving networks through temporal exponential-family random graph models. Technical report, Penn State Univeristy.

[49]. Leibenstein H (1966). Shaping the World Economy: Suggestions for an International Economic Policy.

[50]. Lin Y, Chi Y, Zhu S, Sundaram H, and Tseng BL (2008). Facetnet: a framework for analyzing communities and their evolutions in dynamic networks. In Proceedings of the 17th International Conference on World Wide Web, pages 685–694.

[51]. Ma Z and Ma Z (2017). Exploration of large networks via fast and universal latent space model fitting. arXiv preprint arXiv:1705.02372.

[52]. Mao X, Sarkar P, and Chakrabarti D (2017). Estimating mixed memberships with sharp eigenvector deviations. arXiv preprint arXiv:1709.00407.

[53]. Matias C and Miele V (2017). Statistical clustering of temporal networks through a dynamic stochastic block model. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79(4):1119–1141.

[54]. Matias C and Robin S (2014). Modeling heterogeneity in random graphs through latent space models: a selective review. ESAIM: Proceedings and Surveys, 47:55–74.

[55]. Minhas S, Hoff PD, and Ward MD (2016a). Inferential approaches for network analyses: Amen for latent factor models. arXiv preprint arXiv:1611.00460.

[56]. Minhas S, Hoff PD, and Ward MD (2016b). A new approach to analyzing coevolving longitudinal networks in international relations. Journal of Peace Research, 53(3):491–505.

[57]. Newman M, Barabasi A-L, and Watts DJ (2011). The Structure and Dynamics of Networks. Princeton University Press.

[58]. Newman ME (2004). Coauthorship networks and patterns of scientific collaboration. Proceedings of the National Academy of Sciences, 101 (suppl1):5200–5205.

[59]. Nowicki K and Snijders TA (2001). Estimation and prediction for stochastic blockstructures. Journal of the American Statistical Association, 96(455):1077–1087.

[60]. Perry PO and Wolfe PJ (2013). Point process modelling for directed interaction networks. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 75(5):821–849.

[61]. Qin T and Rohe K (2013). Regularized spectral clustering under the degree-corrected stochastic blockmodel. Advances in Neural Information Processing Systems, pages 3120–3128.

[62]. Raftery AE, Niu X, Hoff PD, and Yeung KY (2012). Fast inference for the latent space network model using a case-control approximate likelihood. Journal of Computational and Graphical Statistics, 21(4):901–919. [PubMed: 27570438]

[63]. Rastelli R, Friel N, and Raftery AE (2016). Properties of latent variable network models. Network Science, 4(4):407–432.

[64]. Rastelli R, Maire F, and Friel N (2018). Computationally efficient inference for latent position network models. arXiv preprint arXiv:1804.02274.

[65]. Rohe K, Chatterjee S, and Yu B (2011). Spectral clustering and the high-dimensional stochastic blockmodel. The Annals of Statistics, pages 1878–1915.

[66]. Rubenstein DI, Sundaresan SR, Fischhoff IR, Tantipathananandh C, and Berger-Wolf TY (2015). Similar but different: Dynamic social network analysis highlights fundamental differences between the fission-fusion societies of two equid species, the onager and grevy's zebra. PloS ONE, 10(10).

[67]. Sarkar P and Moore AW (2005). Dynamic social network analysis using latent space models. ACM SIGKDD Explorations Newsletter, 7(2):31–40.

[68]. Sarkar P, Siddiqi SM, and Gordon GJ (2007). A latent space approach to dynamic embedding of co-occurrence data. In AISTATS, pages 420–427.

[69]. Sewell DK and Chen Y (2015). Latent space models for dynamic networks. Journal of the American Statistical Association, 110(512):1646–1657.

[70]. Sewell DK and Chen Y (2016). Latent space models for dynamic networks with weighted edges. Social Networks, 44:105–116.

[71]. Shizuka D, Chaine AS, Anderson J, Johnson O, Laursen IM, and Lyon BE (2014). Across-year social stability shapes network structure in wintering migrant sparrows. Ecology Letters, 17(8): 998–1007. [PubMed: 24894316]

[72]. Snijders TA (1996). Stochastic actor-oriented models for network change. Journal of Mathematical Sociology, 21(1–2):149–172.

[73]. Snijders TA (2011). Statistical models for social networks. Annual Review of Sociology, 37.

[74]. Snijders TA and Nowicki K (1997). Estimation and prediction for stochastic blockmodels for graphs with latent block structure. Journal of Classification, 14:75–100.

[75]. Snijders TA, Van de Bunt GG, and Steglich CE (2010). Introduction to stochastic actor-based models for network dynamics. Social Networks, 32(1):44–60.

[76]. Sweet TM and Junker B (2011). Modeling intervention effects on social networks in education research. Educational Evaluation and Policy Analysis, 30:203–235.

[77]. Sweet TM, Thomas AC, and Junker BW (2013). Hierarchical network models for education research: Hierarchical latent space models. Journal of Educational and Behavioral Statistics, 38(3):295–318.

[78]. Vivar JC and Banks D (2012). Models for networks: a cross-disciplinary science. Wiley Interdisciplinary Reviews: Computational Statistics, 4(1):13–27.

[79]. Vu DQ, Hunter D, Smyth P, and Asuncion AU (2011). Continuous-time regression models for longitudinal networks. In Advances in Neural Information Processing Systems, pages 2492–2500.

[80]. Ward MD, Ahlquist JS, and Rozenas A (2013). Gravity's rainbow: A dynamic latent space model for the world trade network. Network Science, 1(1):95–118.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

[81]. Ward MD and Hoff PD (2007). Persistent patterns of international commerce. Journal of Peace Research, 44(2):157–175.

[82]. Ward MD, Siverson RM, and Cao X (2007). Disputes, democracies, and dependencies: A reexamination of the kantian peace. American Journal of Political Science, 51(3):583–601.

[83]. Ward MD, Stovel K, and Sacks A (2011). Network analysis and political science. Annual Review of Political Science, 14:245–264.

[84]. Xing EP, Fu W, and Song L (2010). A state-space mixed membership blockmodel for dynamic network tomography. The Annals of Applied Statistics, 4(2).

[85]. Xu KS and Hero AO (2014). Dynamic stochastic blockmodels for time-evolving social networks. IEEE Journal of Selected Topics in Signal Processing, 8(4):552–562.

[86]. Yang T, Chi Y, Zhu S, Gong Y, and Jin R (2011). Detecting communities and their evolutions in dynamic social networks-a bayesian approach. Machine learning, 82(2):157–189.

[87]. Zhao Y, Levina E, and Zhu J (2012). Consistency of community detection in networks under degree-corrected stochastic block models. The Annals of Statistics, 40(4):2266–2292.
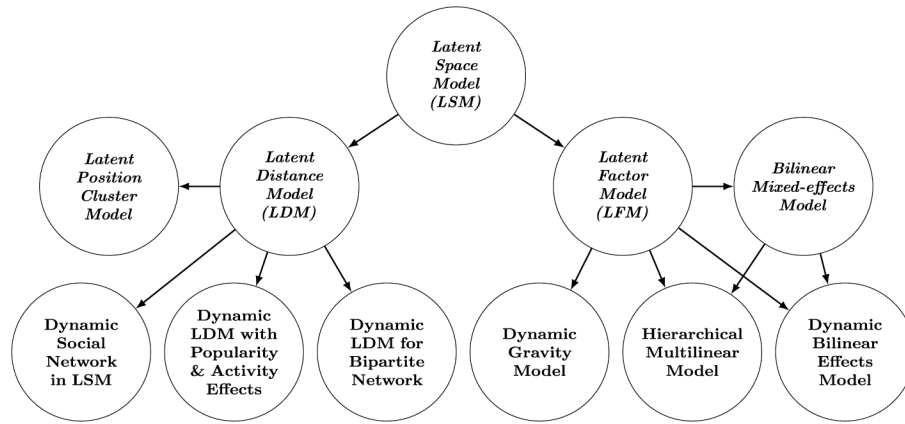
Fig 1.
Network summarizing the relations between models discussed in our review of latent space models (LSM): static latent space models (upper) and dynamic latent space models (lower). Arrows indicate inspiration or influence of the model at the source on the model at the target.
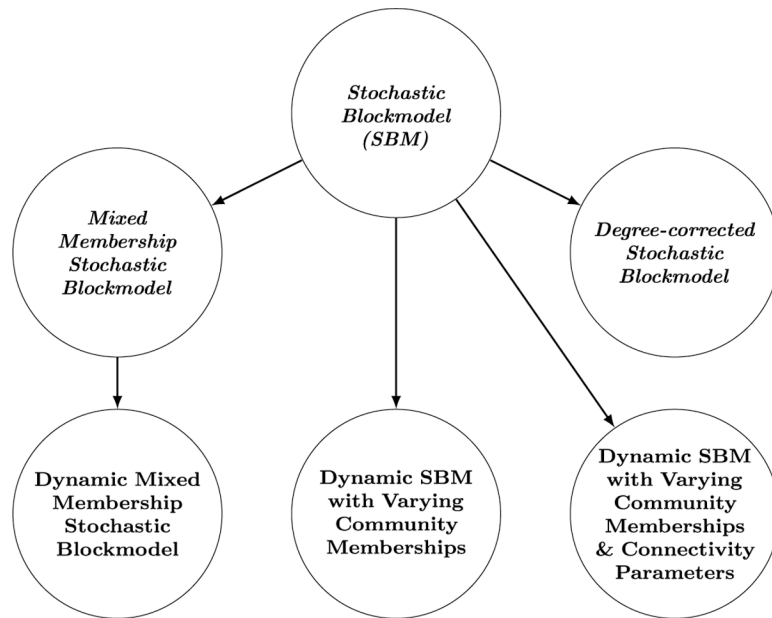
Fig 2.
Network summarizing the relations between models discussed in our review of stochastic blockmodels (SBM): static stochastic blockmodels (upper) and dynamic stochastic models (lower). Arrows indicate inspiration or influence of the model at the source on the model at the target.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Dynamic network data sources.

Data	Article	Section	Source
NIPS data	[67, 68]	2.2.1	http://ai.Stanford.edu/~gal/data.html
Dutch classroom data	[69]	2.2.2	http://visone.info/wiki/index.php/Knecht_Classroom_(data)
Cosponsorship data	[69]	2.2.2	http://jhfowler.ucsd.edu/cosponsorship.htm
Friends and Family data	[70]	2.2.2	http://realitycommons.media.mit.edu/friendsdataset.html
World trade data	[70]	2.2.2	http://www.economicswebinstitute.org/worldtrade.htm
Irish companies data	[19]	2.2.3	Irish Stock Exchange and Irish Companies Registration Office websites
International commerce data	[81]	2.3.1	various sources illustrated in Appendix A [81]
Militarized interstate disputes data	[82]	2.3.1	various sources illustrated in Appendix A [82]
World trade network	[80]	2.3.1	various sources illustrated in Section 3.4 [80]
National Stock Market Indices	[15]	2.3.2	http://finance.yahoo.com/
Financial Network	[13]	2.3.2	http://finance.yahoo.com/ and http://www.gdeltproject.org/data.html
Italian soccer Championship data	[14]	2.3.2	not available
International relations data	[33, 56]	2.3.3	https://dataverse.harvard.edu/dataverse/icews
Southern women data	[86]	3.2.1	Freeman [18]
Blog (NEC Labs) data	[86]	3.2.1	Chi et al. [11]; Lin et al. [50]
Paper co-authorship (DBLP) data	[86]	3.2.1	Asur et al. [6]; Lin et al. [50]
MIT reality mining data	[85]	3.2.2	Eagle et al. [16]
French high school data	[53]	3.2.2	http://www.sociopatterns.org
Migratory birds (sparrows) data	[53]	3.2.2	Shizuka et al. [71] http://datadryad.Org/resource/doi:10.5061/dryad.d3m85
Indian equids (onagers) data	[53]	3.2.2	Rubenstein et al. [66] http://datadryad.Org/resource/doi:10.5061/dryad.q660q
Sampson's monk data	[84]	3.2.3	https://networkdata.ics.uci.edu/netdata/html/sampson.html
Enron email data	[85, 84, 26]	3.2.2 / 3.2.3	https://snap.stanford.edu/data/email-Enron.html
Drosophila melanogaster gene network	[84]	3.2.3	Arbeitman et al. [5]
US congress voting data	[26]	3.2.3	https://www.senate.gov