# A Nonparametric Latent Space Model for Dynamic Relational Networks

## *Un modello non parametrico a spazio latente per reti sociali dinamiche*

Isadora Antoniano-Villalobos, Maxim Nazarov, Sonia Petrone

**Abstract** Recent years have seen a growing interest in the study of social networks and relational data and, in particular, of their evolution over time. In the context of static networks, a commonly used statistical model defines a latent social space and assumes the relationship between two actors to be determined by the distance between them in such latent space. In this manner, it is possible to introduce additional information about each actor and to quantify the residual dependence through a row-column exchangeability assumption on the adjacency matrix associated to the error terms of the model. The present paper analyzes the behavior of the stochastic model given by changes in a "global sociability" parameter which describes the dispersion of the residuals of the positions of the actors in the latent space. This justifies the definition of a Bayesian model for dynamic networks which extends the latent space representation through an infinite hidden Markov model on such positions.

**Abstract** *Negli ultimi anni è cresciuto l'interesse verso lo studio statistico delle reti sociali e dei dati relazionali, e della loro evoluzione temporale. Un modello statico comunemente usato introduce uno "spazio sociale" latente, e ipotizza che la probabilità di una relazione fra due attori sia dipendente dalla posizione da essi ricoperta in tale spazio sociale. In questo modo è possibile introdurre informazioni relative a ciascun attore, e tener conto della dipendenza residua attraverso la scambiabilità tra righe e colonne della matrice di connessione legata al termine di errore. In questo lavoro si analizza il comportamento stocastico del modello in relazione ad un parametro di "socialità globale", espresso dalla dispersione delle posizioni degli attori nello spazio sociale. Si giustifica quindi un modello Bayesiano per reti sociali dinamiche che estende la rappresentazione latente attraverso un "infinite hidden Markov model", in grado di cogliere cambiamenti nella coesione della rete.*

Isadora Antoniano-Villalobos
Universita Bocconi, Milano, Italia, e-mail: isadora.antoniano@unibocconi.it

Sonia Petrone,
Universita Bocconi, Milano, Italia e-mail: sonia.petrone@unibocconi.it

**Key words:** Infinite hidden Markov model, Row-column exchangeability, latent distance model

## 1 Introduction

Network data consist of measured relations between pairs of actors or particles, and it arises naturally in many fields, such as computer science, statistical physics, biology and sociology, to name a few. The first formal treatment of this type of data dates back to the first half of the 20th century, and throughout its second half, an active study of mathematical network theory developed. In recent years, however, we have witnessed a revolution brought about by the intensive data collection from internet, mobile telecommunications and social network traffic, among others. In order to analyze such data, scientific focus has turned back to the study of networks and, in particular, new statistical models have been developed at a pace.

In the present work, we focus on binary network data. Formally, for a fixed set of $N$ actors, each observation is a binary square matrix $Y = (Y_{i,j})_{i,j=1\ldots N}$, for which each entry $Y_{i,j}$ takes the value 1 if there a relationship is present between actors $i$ and $j$, and the value 0 otherwise.

Since the earliest models for this type of data, which constitute what is now considered the *classical* approach, a distinction can be made between two main modeling ideas. On one side, models can be constructed by defining a joint distribution directly on the complete adjacency matrix $Y$. The simplest version of this type of model, proposed by Erdős and Rényi (1959) assumes a uniform distribution over all possible graphs with a given number, $E$, of edges between the $N$ nodes; different statistical models can be obtained depending on the distribution assigned to $E$. A relevant extension, due to Frank and Strauss (1986) has been widely studied, giving rise to the *exponential family random graph models*, based on basic network structure assumption represented by some form of sufficient statistics.

On the other side, Gilbert (1959), proposed a simple model which constructs the joint distribution over $Y$ by considering the individual entries $Y_{i,j}$ as i.i.d. Bernoulli random variables. The model has been extended by replacind the independence of the links with a milder condition of row-column exchangeability or partial exchangeability, in the sense of Aldous-Hoover, of the adjacency matrix $(Y_{i,j})$. One example of such type of extension is the stochastic block model, first introduced by (Nowicki and Snijders, 1959), and extended in recent years, in particular, for the inroduction of a dynamic evolution of the network. Another example is the *latent space model* of (Hoff et al., 2002), in which the probability of link between pairs of actors is defined through latent variables which can be interpreted as the positions of such actors in some latent social space; as an additional advantage, this model may incorporate covariate information on the actors.

The problem of flexibly modeling network structures becomes more crucial when we consider that real-life networks often vary with time. Dynamic modeling is rapidly evolving, but effective inference remains an interesting challenge. We pro-

pose a dynamic extension of the latent space model which aims at preserving the row-column exchangeability assumption at each fixed time, while incorporating a dynamic structure by means of an infinite hidden Markov model driving the temporal evolution of the latent positions of the actors, which allows us to capture changes in the overall cohesion of the network over time. In the present work, we study the sensibility of Hoff et al.'s 2002 model to changes in one of the parameters, which can be interpreted as a measure of the network cohesion. The variety of behaviors that a network can exhibit, depending on this cohesion parameter, justifies our proposal of introducing the time evolution through it.

A positive aspect of this idea is the simplicity of the interpretability of the mechanism driving the dynamics in terms of a time-evolving "global sociability" parameter which, in the presence of covariates, still reflects the unexplained dependence. The nonparametric nature of the infinite hidden Markov model provide enough flexibility to naturally incorporate the evolution of some relevant quantities which are representative of the network structure. In this sense, our approach has some relation with exponential random graph models, which assume summary quantities as sufficient statistics, but suffer from drawbacks (see Chatterjee and Diaconis, 2013). We avoid the assumption of sufficiency in our dynamic model, but still carefully monitor the evolution of statistics of interest.

## 2 Towards a Dynamic Latent Distance Model: the role of dispersion parameters in the latent space model.

Recall that we are considering network data represented by an adjacency matrix $Y \in \{0,1\}^{N\times N}$ for a fixed set of $N$ actors, for which each entry $Y_{i,j}$ is, marginally modeled as, a Bernoulli random variable indicating the presence or absence of a link between actors or nodes $i$ and $j$. The latent space model introduced by Hoff et al. (2002) may incorporate additional information on the actors, in the form of covariates $\mathbf{X} = (X_{i,j})_{i,j\geq 1}$. The probability of a link between two edges is defined as

$$P(Y_{i,j} = 1|\beta, \mathbf{X}, \Gamma) = f(\beta^T \mathrm{X}_{i,j} + \gamma_{i,j}),$$

where $f$ is some adequately chosen link function, and the $Y_{i,j}$'s are considered conditionally independent given the $\beta, \mathbf{X}$ and $\Gamma$.

Row-column exchangeability of the random effects $\Gamma = (\gamma_{i,j})$ can be imposed (see Aldous, 1981) by making

$$\gamma_{i,j} = d(Z_i, Z_j),$$

where $d$ is a distance and $\{Z_1, \ldots, Z_N\}$ are i.i.d. latent variables on $\mathbb{R}^p$, interpreted as the positions of the actors in some latent social space.

For simplicity, we restrict here to the case where no covariate information is present. This implies that the resulting adjacency matrix $Y$ will be, itself, row-

column exchangeable. Furthermore, we take $Z_i \in \mathbb{R}^2$, which provides a useful graphical representation of the network, and define $d$ to be the Euclidean distance in $\mathbb{R}^2$.

The model is completed by the choice of the link function $f$ and the introduction of a prior distribution for the vector $Z = \{Z_1, \ldots, Z_N\}$ of latent positions. Following, Hoff et al. (2002), we consider the $Z_i$ to be i.i.d. random variables with a bivariate normal distribution $\mathrm{N}_2(0, \sigma^2 \mathrm{I}_2)$ and we use a logit link. In addition, in order to simplify notation, we consider here only simple undirected networks, in other words $Y_{i,i} = 0$ and $Y_{i,j} = Y_{j,i}$ for $i, j = 1, \ldots, N$. Therefore, the full model can be expressed as:

$$\mathbb{P}(Y|Z, \beta : 0) = \prod_{i<j} \mathbb{P}(Y_{ij}|Z_i, Z_j, \beta_0) = \prod_{i<j} \frac{\exp\{Y_{i,j}(\beta_0 - ||Z_i - Z_j||_2)\}}{1 + \exp\{\beta_o - ||Z_i - Z_j||_2\}}, \qquad (1)$$

$$Z_i|\sigma^2 \overset{iid}{\sim} \mathrm{N}_2(0, \sigma^2 \mathrm{I}_2); \quad i = 1, \ldots, N.$$

The interpretation of the latent $Z$ in terms of locations of the actors in an imaginary social space makes it natural to introduce a time factor into the model precisely by modeling the change, across time, of such locations. Sewell and Chen (to appear), for example, propose a dynamic model where the latent positions evolve as independent Markov processes. We wish to relax the markovian lack of memory assumption, together with the independence of the evolution of the different actors, while maintaining the row-column exchangeability at each fixed time. We do so via a nonparametric hidden Markov model which places the time dependence on the parameter $\sigma^2$, as will be explained in the following section.

Indeed, the parameter $\sigma^2$ plays a relevant role in the model, capturing the dispersion of the actors' positions, thus the "cohesion" or "global sociability" of the network, so that its temporal evolution may reflect relevant aspects of the network's dynamics. To illustrate this, we provide a simulation study of how the stochastic structure of the network varies when the model parameters influencing the link probability, $(\beta_0, \sigma^2)$, change. Figure 1 shows four networks simulated from the model, with different values of the parameters, namely $\sigma^2 = 4; 10$ and $\beta_0 = 1; 3$. The plots suggest that an increase in $\sigma^2$ corresponds to a decrease in the number of links in the network, while an increase in $\beta_0$ has the opposite effect. Notice that, as the distance between two latent locations $Z_i$ and $Z_j$ approaches 0, the probability of not having a link between the corresponding actors converges to $1/(1 + \exp\{\beta_0\})$ and therefore $\beta_0$ will control the degree to which the social space is believed to represent the relationships between the actors. When $\beta_0 = 0$, two actors in the same location will still have a 50% probability of not being linked, and this probability decreases as $\beta_0$ increases. In what follows we consider the value $\beta_0 = 3$ for which we consider the probability of link between individuals in the same location to be sufficiently close to zero to say that the social space does represent the relationships between the actors.

As mentioned before, in the context of network analysis, one is often interested in some typical summary quantities. We focus on the two of them, which in our case give a clearer clearer indication of the role of $\sigma^2$. The first one, commonly known

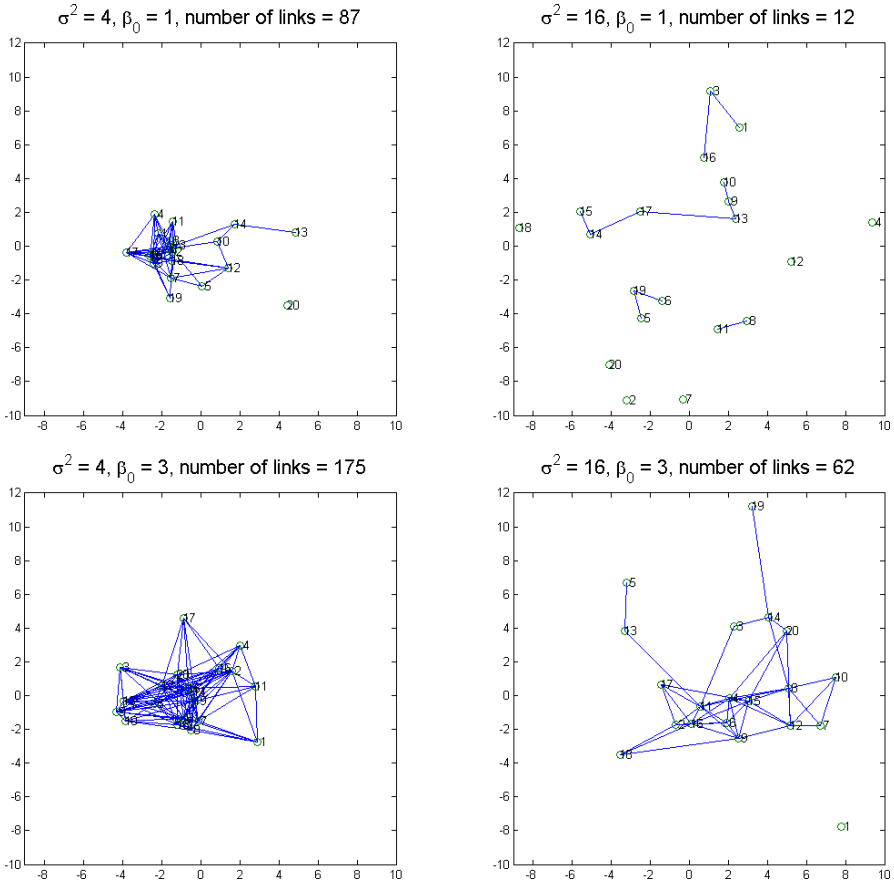
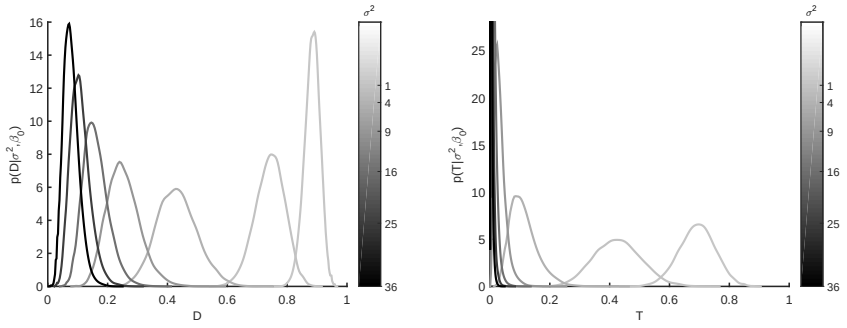



Fig. 1: Four networks with $N = 20$, simulated from the static latent distance model with different values of parameter $\sigma^2$ in the prior distribution of the latent positions $Z$'s and different $\beta_0$. By columns $\sigma^2 = 4$ and $\sigma^2 = 16$, by rows $\beta_0 = 1$ and $\beta_0 = 3$.

as *edge density* $D$, defined as the total number of edges present in the graph, divided by the total possible number of edges given by $N(N-1)$. The second one, which we refer to as the *triangle density*, $T$, is given by the total number of triangles (also known as 3-cliques) present in the graph, divided by the total possible number of triangles, $N(N-1)(N-2)/6$.

Figure 2 shows smoothed histograms of the edge and triangle densities obtained from simulated graphs from the static model with varying values of $\sigma^2$. The static model is clearly sensitive to changes in the parameter, which thus can be said to affect the stochastic behavior of the network structure. $D$ and $T$ were not selected randomly, they are quantities of interest in the context of network analysis and they even constitute commonly used sufficient statistics for the definition of exponential random graphs. One could monitor other quantities, such as the eigenvalues of the




(a) Edge density (b) Triangle density

Fig. 2: Smoothed histograms of the edge density, $D$ and the triangle density, $T$, for a sample of $n = 10000$ graphs with $N = 20$ nodes, simulated from the static latent distance model with parameters $\beta_0 = 3$ and $\sigma^2 = 0.25, 1, 4, 9, 16, 36$.

spectral representation for the graph or given types of sub-graph relevant to a particular application. We limit the present discussion to the two quantities mentioned above only as an example of the different types of behavior that can be captured by changes in $\sigma^2$.

## 3 Dynamic Latent Distance Model

The above study motivates our proposal of a fully Bayesian nonparametric model latent distance model for dynamic network data of the form $Y \in \{0,1\}^{N\times N\times T}$. For this, we assume that the links are conditionally independent given the latent positions and parameters, i.e.

$$\mathbb{P}(Y|Z,\beta_0) = \prod_{t>=0;i<j} \mathbb{P}\left(Y_{ijt}|Z_{it},Z_{jt},\beta_0\right).$$

with marginal distribution given by the logistic regression link function, as in the static model 1. In other words,

$$\mathbb{P}\left(Y_{ijt}=1|Z_{it},Z_{jt},\beta_0\right) = \text{logit}^{-1}\left(\beta_0 - \|Z_{it}-Z_{jt}\|_d\right)$$

The temporal evolution of the process is driven by the change, over time, of the latent positions, $Z_{i,t}$ which we model through an infinite hidden Markov model, i.e.

$$Z_{it}|S_t,(\sigma_k^2) \overset{ind}{\sim} \text{N}_2(0,\sigma_{S_t}^2\text{I}_2)$$
$$P(S_t = j|S_{t-1} = i,\pi) = \pi_{i,j}$$

$$(\pi_i) \sim \text{ hierarchical DP},$$

where $\pi_i$ denotes the $i$-th row of the transition matrix $\pi = (\pi_{i,j})$. In this way, we obtain a dynamic network model in which, for a fixed time $t$, the corresponding adjacency matrix $Y_t$ is row-column exchangeable. The infinite hidden Markov model controls the temporal evolution of the position of the actors in the latent space, through the global sociability parameter $\sigma^2$, which takes different values over time depending on the realizations of the state process $S_t$. The hierarchical Dirichlet process Teh et al. (2006), as a prior distribution on the transition matrix $\pi$, avoids the need for the number of possible states to be fixed a priori. The non Markovian nonparametric structure of the time dependence so defined over $Y$ provides a great flexibility to match real life applications.

Bayesian inference for the propose model is challenging due to the complexity of the data as well as of the modelling structure, and it falls outside of the scope of the present work. We only mention here that it is possible to implement an algorithm for MCMC posterior inference based on the beam sampling algorithm of (Van Gael et al., 2008) and the slice sampling methods of (Kalli et al., 2011), as we will show elsewhere.

## References

D. J. Aldous. Representations for Partially Exchangeable Arrays of Random Variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.

Sourav Chatterjee and Persi Diaconis. Estimating and understanding exponential random graph models. *Ann. Statist.*, 41(5):2428–2461, 10 2013. doi: 10.1214/13-AOS1155. URL `http://dx.doi.org/10.1214/13-AOS1155`.

P. Erdős and A. Rényi. On Random Graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

O. Frank and D. Strauss. Markov Ggraphs. *Journal of the American Statistical Association*, 81(1395):832–842, 1986.

E. N. Gilbert. Random Graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

P. D. Hoff, A. E. Raftery, and M. S. Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098, 2002.

M. Kalli, J. E. Griffin, and S. G. Walker. Slice sampling mixture models. *Statistics and Computing*, 21:93–105, 2011.

K Nowicki and T. A. B. Snijders. Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association*, 455(96):1077–1087, 1959.

D.K. Sewell and Y. Chen. Latent space models for dynamic networks. *Journal of the American Statistical Association*, to appear. doi: 10.1080/01621459.

2014.988214. URL http://dx.doi.org/10.1080/01621459.2014.988214.

Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566–1581, 2006.

J. Van Gael, Y. Saatci, Y. Teh, and Z. Gahramani. Beam Sampling for the Infinite Hidden Markov Model. In *Proceedings of the 25th International Conference on Machine Learning*, pages 1088–1095. ACM, 2008.