Hierarchical Models for Time-Dependent Networks

Andrew C. Thomas

This draft: October 13, 2009

Abstract

1 Time Evolution Models of Networks in the GLM Framework

The modelling approach of Thomas [2009] is integrative at its heart. The adaptation of time evolution into this modelling approach is almost automatic: the addition of previous observations into the observation for the present forms the basis for autoregressive time series modelling.

This is not the only approach used for the time evolution of serial data. In particular, the Markov chain approach has been used for a considerably longer period to capture discrete and continuous time evolution of both discrete and continuous states. The simplest autoregressive time series model, AR-1, works in exactly this fashion: for a time series $\mathbf{Z}$, the set of equations follows the form

$$Z_t = \alpha Z_{t-1} + \varepsilon_t \tag{1}$$

so that given on the autoregressive term α, and the starting value Z_0, each of the equations is conditionally independent of the others; this is easiest to demonstrate if the errors ε_t are normally distributed and jointly form a multivariate normal distribution, such that their lack of correlation implies independence.

However, there is an additional element of social observation that must be respected: that the observation of a friendship at a particular time is a noisy measure of the true, underlying friendship state. For example, if the inference of friendship is made based on an observation, such as two people having lunch together, there are daily processes that affect this outcome – such as bad weather – that would not necessarily affect the underlying friendship, or the observation on the next day. It is essential to explore models that incorporate this mechanism when predicting future outcomes based on past and present observations.

As the treatment of network models is being addressed from this type of perspective, it is natural to ask how existing implementations of time dependence can be extended to this class of data. After reviewing the history of the discrete-time, discrete-state Markov approach, the use of autoregressive processes in the literature for social network data is reviewed; this is followed by the introduction of a filtration approach to social network data that disentangles short-term observational noise from long-term trends, for binary and valued data.

1.1 Review: Discrete-Space Markov-based Approaches

In the case of binary data, there is an intuitive appeal to using Markov processes when constructing time evolution: for a single dyad, it is easy to estimate the probability of making or breaking a new friendship, or maintaining the existing state (see Figure 1, first diagram). Because the transition matrix is row-stochastic, the change at each time step can be treated as a single multinomial outcome, governed by the probabilities in each row. The approaches for solving this class of systems are shared with the estimation methods for models of the p_1 type. Initial work on this area was conducted by Wasserman [1977], followed by Wasserman [1980b,a], with a slightly different approach in Wasserman and Iacobucci [1988]. More advancements on the Markov process have subsequently been made by Huisman and Snijders [2003]; Snijders [2004], and implemented in the software SIENA [Snijders and Huisman, 2003].

Because there are $2^{\binom{n}{2}}$ possible binary states for an n-node system, various simplifying assumptions can be made to dramatically reduce the number of parameters. One of the more popular choices is the Markov Random Field, covered in more detail in Frank and Strauss [1986], such that the evolution of a tie depends only on itself and on any ties that share a common node.

A natural extension to the standard Markov model is a Hidden Markov Model (See Figure 1, second diagram), which supposes that there is an underlying state that evolves over time, governing the likelihood of the observations, which are now independent conditional on this underlying state evolution. As a simple example, the underlying Markov chain can describe the strength of the underlying friendship (strong or weak), and the observation of a friendship (say, a phone call) would be conditional on that underlying strength. Note that this supposes the independence between the visible outcome at time t and the latent state at time $t + 1$, conditional on the latent state at time t, but nonetheless represents an extension of standard Markov models. Inferences on the parameters can be found using the Baum-Welch algorithm (an Expectation-Maximization algorithm) or by straight MCMC methods.

1.2 Autoregressive Models of Network Data

Given the representation of binary ties in terms of latent normal random variables, it is a natural extension to bring the tools from the time series literature onto this area. The AR-1 model in Equation 1 is the basis for this extension; as seen in the third panel of Figure 1, the underlying

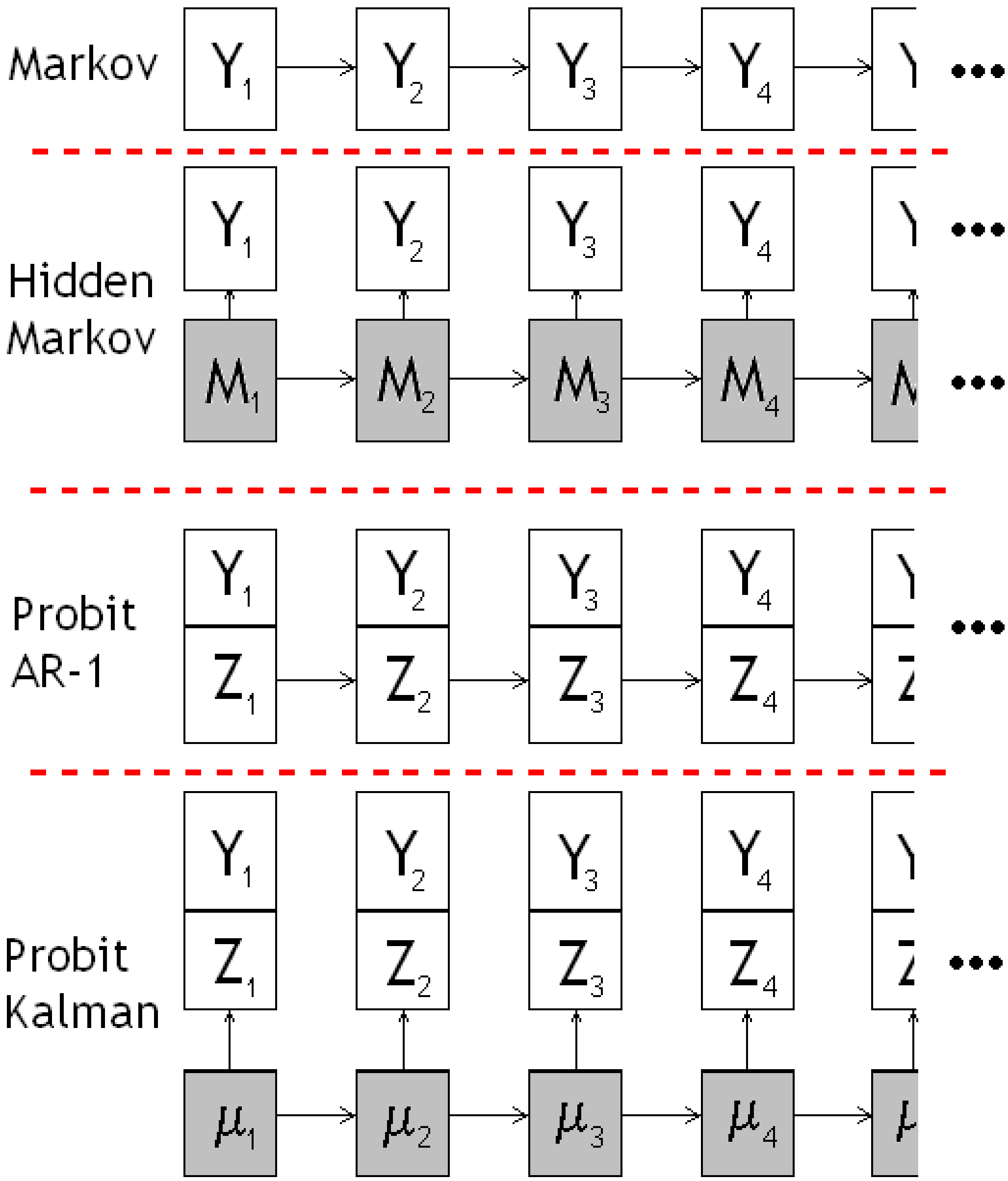

Figure 1: Four potential frameworks for time evolution of binary systems. Top, the standard Markov Chain model. Second, a broad implementation of a Hidden Markov Model (mentioned for completeness), in which the state observations through time are independent conditional on their latent states, which themselves evolve through time. Third, a probit-type autoregressive model with one time lag term; the binary observation is an indicator for the state of a latent random normal which evolves through time. Fourth, a probit Kalman filter approach: the underlying mean of the latent normal evolves through time, while the binary observation is an indicator for the state of the normal. Directly connected boxes imply a functional relationship; arrows imply a stochastic relationship.

normal random variables evolve over time, and their position determines the observed binary outcome.

Among other approaches, Westveld [2007] applies this idea to two different data sets: first, a measure of trade patterns between nations first presented in Ward and Hoff [2005], where the outcome is modelled as Gaussian (so that no latent pattern is needed); the second, a measure of international conflict, which uses said Gaussian implementation, but with the data augmentation step to produce a probit outcome.

The specification of each term's time evolution, while being the main complication of the process, is also its strength. Consider a first order autoregressive time-evolving p_1 model (AR(1)-p_1), in which only sender and receiver effects, along with reciprocity, are allowed. The latent Gaussian in each dyad takes the form

$$\begin{bmatrix} Z_{ijt} \\ Z_{jit} \end{bmatrix} = \tau \begin{bmatrix} Z_{ij(t-1)} \\ Z_{ji(t-1)} \end{bmatrix} + \begin{bmatrix} \alpha_{it} + \beta_{jt} \\ \alpha_{jt} + \beta_{it} \end{bmatrix} + N_2 \left(0, \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \right),$$

so that in the simplest case, the time-varying sender, receiver and reciprocity effects are time-invariant (e.g. $\rho_t = \rho$. However, their own evolution through time may also be considered, such as a random walk (itself a degenerate AR(1)), of the form

$$\alpha_{it} = \alpha_{i(t-1)} + \epsilon_\alpha.$$

1.3 Filtration Models of Network Data

In many circumstances, it may be more useful to consider the evolution of the system as being completely latent, in particular if the event itself is merely a noisy reflection of the true quantity of interest. One mechanism that can accomplish this is the Kalman filter, which has had great success as as a mechanism for separating the noise from the time evolution of a system. The probit form of the Kalman filter has been proposed in econometrics [Dueker, 2006] though not implemented for this class of data.

One particular implementation of this format is

$$\begin{aligned} Y_{it} &= \mathbb{Z}_{it} > 0) \\ Z_{it} &= \mu_{it} + X_{it}\beta_t + \varepsilon_{it} \\ \mu_{it} &= \tau\mu_{i(t-1)} + Z_{it}\nu_t + \gamma_{it}; \end{aligned}$$

In this case, there are covariates both for time evolution (some change in the relationship) and for the noise of observation (say, weather).

An alternative filtration is on the underlying latent space; Sarkar and Moore [2005] considers the time evolution of the positions of actors in a latent space model, under the assumption that they undergo a random walk in this space. This can also be implemented with the existing framework.

It is not necessary to process all time points simultaneously when computing model parameters. Dynamic programming methods allow for a two-step process, the forward-backward algorithm, to solve for the state of a system (or, only the forward portion if the goal is prediction of the next state).

References

Dueker, M. (2006). Kalman filtering with truncated normal state variables for bayesian estimation of macroeconomic models. *Economics Letters*, **93** 58–62.

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

Huisman, M. and Snijders, T. (2003). Statistical analysis of longitudinal network data with changing composition. *Sociological Methods and Research*, **32** 253 287.

Sarkar, P. and Moore, A. W. (2005). Dynamic social network analysis using latent space models. *SIGKDD Explorations*, **7**.

Snijders, T. and Huisman, J. (2003). *SIENA*. URL http://stat.gamma.rug.nl/stocnet/.

Snijders, T. A. (2004). Models for longitudinal network data. In *Models and Methods for Social Network Analysis*. Cambridge University Press, 215–247.

Thomas, A. C. (2009). *Hierarchical Models for Relational Data*. Ph.D. thesis, Harvard University.

Ward, M. D. and Hoff, P. D. (2005). Persistent patterns of international commerce. Tech. rep., Center for Statistics and the Social Sciences, University of Washington.

Wasserman, S. (1977). *Stochastic models for directed graphs*. Ph.D. thesis, Harvard University.

Wasserman, S. (1980a). Analyzing social networks as stochastic processes. *Journal of the American Statistical Association*, **75** 280 294.

Wasserman, S. (1980b). A stochastic model for directed graphs with transition rates determined by reciprocity. *Sociological Methodology*, **11** 392–412.

Wasserman, S. and Iacobucci, D. (1988). Sequential social network data. *Psychometrika*, **53** 261–282.

Westveld, A. H. (2007). *Statistical Methodology for Longitudinal Social Network Data*. Ph.D. thesis, University of Washington.