18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Simulating exponential random graph models for social networks: From local to global structure

G.L. Robins

School of Behavioural Science, University of Melbourne, Australia (email: garrylr@unimelb.edu.au)

Abstract: This paper reviews current progress in the development of exponential random graph models, a family of statistical models for modeling social networks. Full model specification requires formulation of a dependence hypothesis which proposes how possible network ties are conditionally dependent on each other. The dependence hypothesis in effect sets out the processes whereby network ties self organize into small micro-structures (network configurations). These can be construed as the building blocks of the global network structure. The paper presents three dependence assumptions that result in different types of graph distribution: (1) Bernoulli graph distributions whereby ties are assumed to be independent of each other; (2) Markov random graph distributions whereby ties are assumed to be conditionally dependent if they share a node; and (3) social circuit dependence where dependence emerges from observed ties, such that two possible ties become dependent if their observation would create a four cycle.

Methods of simulation and parameter estimation from empirical data are described. Simulation studies show that a variety of global network outcomes can be produced from these models. Markov models can produce graphs with small world type properties and social circuit models can result in core-periphery structures as well as more segmented graphs with connected islands of denser network regions. Nevertheless, there are regions of parameter space for some models that result in frozen, determinist structures that do not necessarily represent real network data well.

Keywords: *Social networks, exponential random graph models, statistical models, estimation.*

1. INTRODUCTION

Exponential random graph models (ergm) provide families of statistical models suitable for modeling social networks. They were first introduced by Frank and Strauss (1986), applying ideas from the spatial statistics literature to social network data structures. Wasserman and Pattison (1996) further extended this class of models, describing them as p^* models. Over the last decade there has been considerable methodological work further developing the class of models, including more recent model specifications, and advances in simulation and estimation, as described briefly below.

Although the class of models can be applied to various complex forms of relational data, in this paper in this paper we concentrate on the simplest variant: the modeling of a single binary social network. We use the following notation. A social network can be represented by a graph **G** with node set **N** of size *n*. To model **G** we apply a probability distribution across all possible graphs with node set **N**. For each *i* and *j* who are distinct members of **N**, let X_{ij} be a random variable where $X_{ij} = 1$ if there is an edge (or tie) between *i* and *j*, and where $X_{ij} = 0$ otherwise. Denote x_{ij} as a realization of the variable X_{ij} and we let **X** be the set of all variables X_{ij} with **x** a set of realized ties. Then, as shown by Frank and Strauss (1986), the Hammersley-Clifford theorem (Besag, 1974) implies a general parametric form for the probability distribution of graphs, $\Pr(\mathbf{X} = \mathbf{x})$:

$$Pr(\mathbf{X} = \mathbf{x}) = \exp(\Sigma_A \lambda_A z_A(\mathbf{x}))/\kappa \qquad (1)$$

where *A* is a neighbourhood of ties, as explained below; λ_A is a parameter associated with the neighbourhood *A*; the quantity $z_A(\mathbf{x})$ is a *network statistic* indicating whether all ties in *A* are present in **x**; and κ is a normalizing constant.

A neighbourhood is a subset of possible ties, all of which are conditionally dependent on each other. By conditionally dependent, we mean that two variables X_{ij} and X_{rs} are dependent on one another irrespective of the observations on all other variables.

Obviously the model is not specified unless the neighbourhoods are well defined. Typically a dependence assumption is made, equivalent to a basic assumption about the ways in which ties self organize in a social network. The neighbourhoods are collections of possible ties, and if those ties were realized, they would result in certain subgraph patterns. These patterns – termed *configurations* – can then be construed as the basic structural building blocks from which the network emerges. So a social neighbourhood corresponds to a network configuration and the probability of any graph **x** is determined by whether these configurations are observed in **x** or not (the graph statistic $z_A(\mathbf{x})$) weighted by the parameter values λ. Through the possibly overlapping neighbourhoods or configurations *A*, the model expresses the probability of network **x** as a function of self-organising processes in local regions of the network. These processes result in the presence (or absence) of configuration *A* in **x**, with a positive parameter value λ_A resulting in a tendency for configuration *A* to be observed in **x** than would otherwise be expected by chance. In *homogeneous* models, parameters for isomorphic configurations are equated. In that case a positive value of λ_A indicates that networks with more configurations of type *A* are more probable.

Equation (1) is the general form of an *exponential random graph model* for social networks, also known as the p^* model (Frank & Strauss, 1986; Pattison & Wasserman, 1999; Robins, Pattison & Wasserman, 1999; Wasserman & Pattison, 1996 – for a recent review, see Wasserman & Robins, 2005). The critical step is to specify the model is to apply a dependence assumption that then constrains the possible configurations in the model.

2. DEPENDENCE ASSUMPTIONS

2.1. Bernoulli graph models

The simplest dependence assumption is that all possible ties are independent of one another. A homogeneous model then implies that ties occur with constant probability, as in the *Bernoulli graph* model (Frank & Nowicki, 1993). This is equivalent to a simple random graph or Erdös-Renyi model (Erdös & Renyi, 1959). Because possible ties are independent, there is a separate neighbourhood for each X_{ij}, and no other neighbourhoods exist under this assumption. The only configuration in a homogeneous model is a single edge:

$$\Pr(\boldsymbol{X} = \boldsymbol{x}) = (1/\kappa) \exp(\theta L(\mathbf{x}))$$

where L is the number of edges in the graph $\mathbf{x}$, and θ is an *edge parameter* reflecting the probability with which the ties occur.

2.2. Markov random graph models

Frank and Strauss (1986) proposed a Markov dependence assumption whereby two network variables X_{ij} and X_{rs} are assumed conditionally independent unless they share a node ($\{i,j\} \cap \{l,m\} \neq \varnothing$). Frank and Strauss showed that this neighbourhood assumption resulted in configurations of single edges, all k-stars for k=2, ..., n – 1, (where a k-star is a configuration centred on a node i with k edges expressed from it) , and triangles:

$$\Pr(\boldsymbol{X} = \boldsymbol{x}) = (1/c) \exp\{\theta L(\boldsymbol{x}) + \sigma_2 S_2(\boldsymbol{x}) + \sigma_3 S_3(\boldsymbol{x}) + \ldots + \sigma_{n-1} S_{n-1}(\boldsymbol{x}) + \tau T(\boldsymbol{x})\}$$

where σ_k and $S_k(\boldsymbol{x})$ denote the k-star parameter and the number of k- stars in $\boldsymbol{x}$, respectively; and τ and $T(\boldsymbol{x})$ denote the triangle parameterand the number of triangles, respectively. It is possible to restrict the number of k-star parameters used in the model. For instance, Robins, Pattison and Woolcock (2005) used edge, 2-star, 3-star and triangles in a four parameter Markov model.

2.3. Social circuit dependence

Pattison and Robins (2002) suggested that Markov models may be inadequate and proposed that dependence may recursively emerge from observed network ties: that is, certain neighbourhoods could be present based on the observed network ties in $\mathbf{x}$. Snijders, Pattison, Robins and Handcock (2006) used this approach to propose what has been termed *social circuit dependence*. Two possible ties X_{ij} and X_{rs} are conditionally dependent if $X_{ir} = X_{js} = 1$. That is, two possible ties are dependent if, when observed, they complete a four cycle.

Snijders et al (2006) showed that when this assumption is used alongside Markov dependence a number of additional configurations arise in the model. The most important new configuration was a k-triangle which is k triangles all sharing the one edge (the *base* of the k-triangle): that is, if the base is an edge between nodes i and j, then there are k single triangles among nodes i, j and r_p where p goes from 1 to k. In other words i and j are both connected to k *shared network partners*.

Snijders et al (2006) also introduced novel methods to combine parameters of the same class (e.g. k – star parameters and k – triangle parameters). They introduced a geometric weighting (with alternating signs) between parameters of different orders. For instance, the *alternating k-star hypothesis* assumes that $\sigma_{(k+1)} = -\sigma_k/\lambda$ for the σ_k star parameters for some λ greater than 1 (typically λ=2).

3. SIMULATION OF EXPONENTIAL RANDOM GRAPH DISTRIBUTIONS

Strauss (1986) was the first author to propose simulation of exponential random graph models using well-established statistical algorithms. Other later authors describe similar approaches (e.g., Robins et al, 2005; Snijders, 2002). For instance, the Metropolis algorithm (e.g., see Gilks, Richardson, & Spiegelhalter, 1996) can be used to simulate well specified graph distributions of the form of (1) with parameter values λ_A fixed. Start with a randomly chosen graph on node set $\mathbf{N}$. At each iteration, a candidate graph $\boldsymbol{x}'$ is proposed as the next step in a Markov chain. The candidate graph $\boldsymbol{x}'$ is obtained from the current graph $\boldsymbol{x}$ by selecting at random a possible edge from i to j and changing x_{ij} to $1 - x_{ij}$. The candidate graph is accepted if it has an increased probability of being observed according to (1). When the probability is not increased, the candidate graph is accepted with probability

$$r = \exp\text{-}\{\Sigma_A \lambda_A (z_A(\mathbf{x}) - z_A(\mathbf{x}'))\}$$

This process establishes a Markov chain on the state space of all graphs with nodes set $\mathbf{N}$ and, after a suitable burnin, converges to the probability distribution (1). Once burn in has been achieved, samples of graphs can be obtained to study the distribution of graph properties typical of the distribution (e.g. Robins et al, 2005).

3.1. Small world properties for Markov models

Robins et al (2005) showed that Markov random graph models, when simulated with a judicious choice of parameters, can produce the properties of a *small world.* Watts (1999) specified the properties of small world networks in graph-theoretic terms: a small world graph has low *density* and is highly *clustered* but with short *average geodesics*. Using the four parameter Markov model described above, Robins et al (2005) found that small world graphs tended to be produced with the following pattern of parameter values: negative edge parameter, positive 2-star and negative 3-star parameters and positive triangle parameters. Robins et al interpreted the parameters as implying that small world graphs could arise in social networks when:

1. individuals sought more than one network partner,
2. but the costs of maintaining many partners were high (with a tendency against too many partners);
3. a tendency for network partners to agree about other possible partners (i.e. for clustering);
4. but this tendency is neither too strong (else the network becomes too clique-like with insufficient links between cliques for smaller geodesics), nor too weak (else there is little clustering in the network).

These results indicate that the global patterns produced from cumulation of configurations can be interpreted in meaningful ways. The fourth condition above, however reveals an important property of these models, that there are regions of parameter space whereby the resulting graphs are"frozen" in one part of the state space, producing either empty or full graphs, or "crystalline" structures involving, for instance, disconnected cliques. There has been an interesting recent discussion in the literature of these "degenerate" or "near-degenerate" regions (e.g. Burda, Jurkiewicz, & Kryzwicki, 2004; Häggström & Jonasson, 1999; Handcock, 2002, 2003; Jonasson, 1999; Park & Newman, 2004; Robins et al, 2005; Robins et al, 2007; Snijders, 2002; Snijders et al 2006.) Suffice to say that these parameter regions are not good at representing the social network data that is typically collected. Moreover, these models are typical of complex systems, with phase transitions (e.g. Robins et al, 2005) and chaotic behaviour, whereby the same parameter values may yield graphs that may be vastly different (e.g. Snijders, 2002). Markov graph models are especially prone to these phenomena, a principal motivation for Snijders et al (2006) to propose social circuit models. Social circuit models generally display more stable behaviour and can more readily represent real network data (Robins et al, 2007).

3.2. Simulation results for social circuit models

A model with only an edge and positive *k*-triangle parameter tends to produce a core-periphery global graph structure. However, if a negative *k*-star parameter is included, the core "breaks down", leading to a more segmented (but connected) network structure with a larger number of smaller dense regions. These results are demonstrated by Snijders et al (2006) and Robins et al (2007).

4. ESTIMATION OF PARAMETERS

It is possible to estimate parameter values for these models using Monte Carlo Markov Chain maximum likelihood estimation (MCMCMLE) as proposed and implemented by Snijders (2002) and Hunter and Handcock (2006). The basic idea is to begin estimation with an initial set of parameter values, simulate from that parameter set and compare the resulting graph distribution with the observed data, adjusting the parameter values until the estimation converges. If the model is degenerate, the estimation will not successfully converge, which is why the social circuit specifications are typically to be preferred over Markov random graph models when fitting models to data.

5. CONCLUSIONS

Work continues on model specification for exponential random graph models. The social circuit models have been shown to be useful for smaller scale networks but additional parameters may be required to model large scale systems successfully. One of the strengths of the models is that a variety of effects (including node-level properties, multiple networks, and spatial covariates) can be included in the one model and thereby tested against each other in regard to real data.

6. REFERENCES

Burda, Z., Jurkiewicz, J., and Kryzwicki, A. (2004) Network transitivity and matrix models. *Physical Review E* 69: 026106.

Erdös P, and Renyi A (1959) On random graphs. I. *Publicationes Mathematicae* (Debrecen) 6: 290-297.

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

Häggström O., and Jonasson, J. (1999). Phase transition in the random triangle model. *Journal of Applied Probability* 36: 1101-1115.

Handcock, M. (2002). Statistical models for social networks: Degeneracy and inference. In: R. Breiger, K. Carley, and P. Pattison. (eds) *Dynamic social network modeling and analysis*. Washington DC: National Academies Press, pp. 229-240

Handcock, M. (2003). *Assessing degeneracy in statistical models of social networks*. Center for Statistics and the Social Sciences, Working Paper no 39, University of Washington

Hunter, D. and Handcock, M. (2006). Inference in curved exponential families for networks. *Journal of Computational and Graphical Statistics* 15: 565-583.

Jonasson, J. (1999) The random triangle model. *Journal of Applied Probability* 36: 852-867.

Park, J, and Newman, M, (2004) Solution of the 2-star model of a network. *Physical Review* E70: 066146.

Pattison, P., and Robins, G. (2002). Neighbourhood based models for social networks. *Sociological Methodology* 32: 301-337.

Robins, G., Pattison, P., Woolcock, J. (2005). Social networks and small worlds. *American Journal of Sociology* 110: 894-936.

Robins, G.L., Snijders, T.A.B., Wang, P., Handcock, M., and Pattison, P. (2007). Recent developments in exponential random graph (*p**) models for social networks. *Social Networks* 29: 192-215.

Strauss, D. (1986). On a general class of models for interaction. *SIAM Review* 28: 513-527.

Snijders, T.A.B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure* 3: 2.

Snijders, T.A.B., Pattison, P., Robins, G., and Handcock, M. (2006). New specifications for exponential random graph models. *Sociological Methodology* 36: 99-153.

Wasserman, S., and Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and *p**. *Psychometrika* 61: 401-425.