

Statistical Models for Networks: A Brief Review of Some Recent Research

Stanley Wasserman^a, Garry Robins^b, and Douglas Steinley^c

December 13, 2006

Technical Report 06-01
Department of Statistics
Indiana University
Bloomington, IN 47405

^a Department of Statistics, Department of Sociology, and Department of Psychological and Brain Sciences,

Indiana University. stanwass@indiana.edu

^b School of Behavioural Science,

University of Melbourne. g.robins@psych.unimelb.edu.au

^c Department of Psychology and Department of Statistics,

University of Missouri. steinleyd@missouri.edu

We begin with a graph (or a directed graph), a single set of nodes \mathcal{N} , and a set of lines or arcs \mathcal{L} . It is common to use this mathematical concept to represent a *network*. We use the notation of Wasserman and Faust (1994), especially Chapters 13 and 15. There are extensions of these ideas to a wide range of networks, including multiple relations, affiliation relations, valued relations, and social influence and selection situations (in which information on attributes of the nodes is available), all of which can be found in the chapters of Carrington, Scott, and Wasserman (2005).

The purpose of this short exposition is to discuss the developments in statistical models for networks that have occurred over the past five years, since the publication of the statistical chapters (8, 9, 10, and 11) of Carrington, Scott, and Wasserman (which were written in 2002). The statistical modeling of social networks is advancing quite quickly. The many exciting new developments include, for instance, longitudinal models for the co-evolution of networks and behavior (Snijders, Steglich & Schweinberger, in press) and latent space models for social networks (Hoff, Raftery & Handcock, 2002). In this chapter, we do not intend to review *all* the recent advances but rather limit our scope to a few developments that we have worked on.

1 Background

Early work on distributions for graphs was quite limiting, forcing researchers to adopt independence assumptions that were not terribly realistic (see Chapters 13-16 of Wasserman and Faust, 1994). It is hard to accept the standard assumption common in much of the literature, especially in physics, of complete independence and then to adopt the mis-named and overly simplistic “random graph” distribution (there are, of course, an infinite number of random graph distributions). The random graph distribution to the physicists, that is usually referred to as a Bernoulli graph, assumes no dependencies at all among the random components of a graph. Equally hard to believe as a true representation of social behavior are the many conditional uniform distributions and p_1 , which assumes independent dyads.

The breakthrough in statistical modeling of networks was first expounded by Frank and Strauss (1986), who termed their model a *Markov random graph*. Further developments, especially commentary on estimation of distribution parameters, were given by Strauss and Ikeda (1990). Wasserman and Pattison (1996) elaborated upon the model, describing a more general family of distributions. Pattison and Wasserman (1999), Robins, Pattison, and Wasserman (1999), and Anderson, Wasserman, and Crouch (1999) further developed this family of models, showing how a Markov parametric assumption gives just one, of many, possible sets of parameters. This family, with its variety and extensions, was named p^* , a label by which it has come to be known. The parameters (which are determined by the hypothesized dependence structure) reflect structural concerns, which are assumed to be governing the probabilistic nature of the underlying social and/or behavioral process.

This pre-2000 early work by the first researchers extended p^* in a variety of ways, and laid the foundation for work in this decade on the estimation problems inherent in the early formulations. This research also was an important forerunner of the new parametric specifications that promise wider usage of the family. A more thorough history of this family of distributions, including a discussion of its roots in spatial modeling and statistical physics, can be found in Borner, Sanyal, and Vespignani (in press). Wasserman and Robins (2005) offer a review of p^* circa-2003, while Robins, Pattison, Kalish, and Lusher (2006) review the 2003-2006 period.

The work of Frank and Strauss (1986) did indeed begin a new era for statistical modeling of networks, although it took ten years for Markov random graphs to be discussed at more length by network methodologists. We briefly describe the highlights of the past decade here.

2 Some Notation and a Bit of Background

A *network* is a set of n actors and a collection of r relations that specify how these actors are related to one another. As defined by Wasserman and Faust (1994; Chapter 3), a network can also contain a collection of attribute characteristics, measured on the actors.

We let $\mathcal{N} = \{1, 2, \dots, g\}$ denote the set of actors, and \mathcal{X} denote a particular relation defined on the actors (here, we let $r = 1$). Specifically, \mathcal{X} is a set of ordered pairs recording the presence or absence of relational ties between pairs of actors. This binary relation can be represented by a $g \times g$ matrix \mathbf{X} , with elements

$$X_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \mathcal{X}, \\ 0 & \text{otherwise.} \end{cases}$$

We will use a variety of graph characteristics and statistics throughout; such quantities are defined in the early chapters of Wasserman and Faust (1994). We assume throughout that \mathbf{X} and its elements are random variables. Typically, these variables are assumed to be interdependent, given the interactive nature of the social processes that generate and sustain a network. Much of the work over the past decade has been on the explicit hypotheses underlying different types of interdependencies.

In fact, one of the new ideas for network analysis, utilized by the p^* family of models is a *dependence graph*, a device which allows one to consider which elements of \mathbf{X} are independent. Wasserman and Robins (2005) discuss such graphs at length. A dependence graph, which we illustrate in the next section, is also the starting point for the Hammersley–Clifford Theorem, which posits a very general probability distribution for network random variables using the postulated dependence graph. The exact form of the dependence graph depends on the nature of the substantive hypotheses about the network under study.

As outlined by Robins, Pattison, et al., (in press), a statistical model for a network can be constructed using this approach through a series of five steps:

- Regard each relational tie as a random variable

- Specify a dependence hypothesis, leading to a dependence graph
- Generate a specific model from the p^* family from the specified dependence graph
- Simplify the parameter space through homogeneity or other constraints
- Estimate, assess, and interpret model parameters

We have mentioned the first two of these steps. To discuss the latter three, we need to introduce p^* via a dependence graph.

3 Statistical Theory

Any observed single relational network may be regarded as a realization $\mathbf{x} = [x_{ij}]$ of a random two-way binary array $\mathbf{X} = [X_{ij}]$. The dependence structure for these random variables is determined by the *dependence graph* \mathcal{D} of the random array \mathbf{X} . \mathcal{D} is itself a graph whose nodes are elements of the index set $\{(i, j); i, j \in \mathcal{N}, i \neq j\}$ for the random variables in \mathbf{X} , and whose edges signify pairs of the random variables that are assumed to be conditionally dependent (given the values of all other random variables).

More formally, a dependence graph for a univariate network has node set

$$\mathcal{N}_D = \{(i, j); i, j \in \mathcal{N}, i \neq j\}.$$

The edges of \mathcal{D} are given by

$$\mathcal{E}_D = \{((i, j), (k, l)), \text{ where } X_{ij} \text{ and } X_{kl} \text{ are not conditionally independent}\}.$$

Consider now a general dependence graph, with an arbitrary edge set. Such a dependence graph yields a very general probability distribution for a (di)graph, which we term p^* and focus on below.

For an observed network, which we consider to be a realization \mathbf{x} of a random array \mathbf{X} , we assume the existence of a dependence graph \mathcal{D} for the random array \mathbf{X} . The edges of \mathcal{D} are crucial here; consider the set of edges, and determine if there are any *complete subgraphs*, or cliques found in the dependence graph. (For a general dependence graph, a subset A of the set of relational ties \mathcal{N}_D is *complete* if every pair of nodes in A (that is, every pair of relational ties) is linked by an edge of \mathcal{D} . A subset comprising a single node is also regarded as complete.). These cliques specify which subsets of relational ties are all pairwise, conditionally dependent on each other.

The Hammersley–Clifford theorem (see Wasserman and Robins, 2005, for a summary) establishes that a probability model for \mathbf{X} depends only on the cliques of the dependence graph \mathcal{D} . In particular, application of the Hammersley–Clifford theorem yields a characterization of $Pr(\mathbf{X} = \mathbf{x})$ in the form of an exponential family of distributions:

$$Pr(\mathbf{X} = \mathbf{x}) = \left(\frac{1}{\kappa}\right) \exp\left(\sum_{A \subseteq \mathcal{N}_D} \lambda_A \prod_{(i,j) \in A} x_{ij}\right) \quad (1)$$

where:

- $\kappa = \sum_{\mathbf{x}} \exp\{\sum_{A \subseteq \mathcal{D}} \lambda_A \prod_{(i,j) \in A} x_{ij}\}$ is a normalizing quantity;
- \mathcal{D} is the dependence graph for \mathbf{X} ; the summation is over all subsets A of nodes of \mathcal{D} ;
- $\prod_{(i,j) \in A} x_{ij}$ is the sufficient statistic corresponding to the parameter λ_A ; and
- $\lambda_A = 0$ whenever the subgraph induced by the nodes in A is not a clique of \mathcal{D} .

The set of non-zero parameters in this probability distribution for $Pr(\mathbf{X} = \mathbf{x})$ depends on the *maximal* cliques of the dependence graph (A maximal clique is a complete subgraph that is not contained in any other complete subgraph). Any subgraph of a complete subgraph is also complete (but not maximal), so that if A is a maximal clique of \mathcal{D} , then the probability distribution for the (di)graph will contain non-zero parameters for A and all of its subgraphs. Each clique, and hence each nonzero parameter in the model, corresponds to a *configuration*, a small subgraph of possible network ties. Different dependence assumptions result in different types of configurations. For instance, Frank and Strauss (1986) showed that configurations for Markov dependence (described below) were edges, stars of various types (a single node with arcs going in and/or out), and triangles for nondirected graphs. The model in effect supposes that the observed network is built up from combinations of these various configurations, and the parameters express the presence (or absence) of the configurations in the observed network. For instance, a strongly positive triangle parameter is evidence for more triangulation in the network implying that networks with large numbers of triads have larger probabilities of arising.

All models from this family, which we refer to as p^* , have this form. Some recent literature refers to these models as ERGMs — exponential random graph models. It is our course uninformative to refer to these distributions as “exponential random graphs” — almost any probability distribution for a graph can be made “exponential”. Further, strictly speaking, the model is not exponential, but in the statistical sense, an exponential *family*, which conveys a special meaning in statistical theory (and has important implications for some of the estimation procedures described below – see Hunter, in press). Hence, we much prefer the more informative moniker p^* , and the label, *an exponential family of distributions* for random graphs. The p^* label (first used by Wasserman & Pattison, 1996) derives from the research on statistical modeling commenced by Holland and Leinhardt with their dyadic-independence p_1 model.

As for the details: the probability of a particular realization of a random graph depends on the cliques of the dependence graph, and from that, the sufficient statistics (arising from the *configurations*) specified by the hypothesized dependencies. The sufficient statistics are the counts of these configurations arising in the realization being modeled.

There are a variety of dependence graphs well-known in the literature. One very general and simple member of the p^* family is the Bernoulli graph; another is Holland and Leinhardt’s p_1 . The dependence graph first proposed for networks, for which this distribution was first developed, assumes conditional independence of X_{ij} and X_{kl} if and only if $\{i, j\} \cap \{k, l\} = \emptyset$. This dependence graph links

any two relational ties involving the same actor(s); thus, any two relational ties are associated if they involve the same actor(s). Because of the similarity to the dependence inherent in a Markov spatial process, such a random graph was labeled “Markov” by Frank and Strauss (1986). Further discussion can be found in Section 4 of Robins, Pattison, et al. (in press).

One can also formulate dependence graphs when data on attribute variables measured on the nodes is available. If the attribute variables are taken as fixed, with network ties varying depending on the attributes, then models for social selection arise (Robins, Elliott & Pattison, 2001). If on the other hand the network is assumed fixed, with the distribution of attributes dependent on the pattern of network ties, the outcomes are models for social influence (Robins, Pattison & Elliott, 2001).

4 Parameters – New Specifications

Limiting the number of parameters is wise – one can either postulate a simple dependence graph, or by making assumptions about the parameters. The usual assumption is *homogeneity*, in which parameters for isomorphic *configurations* of nodes are equated.

Even with homogeneity imposed, models may not be identifiable. Typically, parameters for higher order configurations (for example, higher order stars or triads) are set to zero (equivalent to setting higher order interactions to zero in general linear models).

As mentioned, Markov random graph models were indeed a breakthrough in moving towards more realistic dependence assumptions. But recently it has been shown that Markov dependence is often inadequate in handling typical social network data. Frequently, parameters arising from Markov dependence assumptions are consistent with either complete or very sparse networks, which are of course unhelpful in modeling realistic data. Several authors have provided technical demonstrations of this problem (Handcock, 2002; Park & Newman, 2004; Robins, Pattison & Woolcock, 2005; Robins, Snijders, Wang, Handcock & Pattison, in press; Snijders, 2002; Snijders, Pattison, Robins & Handcock, 2006). An intuitive explanation of the difficulty follows.

Markov random graphs assume that stars and triangles are rather evenly spread throughout the network. In many observed networks, of course, there are dense regions (that is, concentrations of many triangles) and some high degree nodes (that is, concentrations of many stars). As a result, for such data, parameter estimation for Markov random graphs is problematic: there is difficulty in finding “average” parameter values that can adequately capture such structural heterogeneity (see Robins, Snijders, et al., in press, and Snijders, et al., 2006, for further discussion.) When sensible parameter estimates cannot be obtained, the model is said to be *degenerate*, or *nonconvergent*.

Snijders, et al. (2006) proposed a method of combining counts of all the Markov star parameters into the one statistic, with geometrically decreasing weights on the higher order star counts so that they did not come to dominate the

calculation. The resulting parameter is termed a *geometrically weighted degree parameter*, or an *alternating k -star parameter* (the term alternating comes from alternating signs in the calculation of the statistic.) Various versions of this new degree-based parameter have been proposed (see Hunter, in press, who shows the linkages between them), but whatever the precise form of the parameter, it permits greater heterogeneity in the degree distribution, so that it is more capable of modeling high degree nodes than a small number of low order Markov star parameters. Such parameters appear to greatly increase the “fittability” of models.

Perhaps the most important innovation of Snijders et al (2006) was the introduction of *k -triangles*, configurations with k separate triangles sharing one edge, the base of the k -triangle. These configurations also introduce a new distribution of graph features (alongside the degree distribution and the geodesic distribution): the *edge-wise shared partner distribution* (see Hunter, in press; Hunter & Handcock, 2006.) Counts of the k -triangle configurations are combined into one statistic just as for the case of the geometrically weighted degree parameter, producing a new statistic and associated parameter for alternating k -triangles. This parameter models triangulation in the network but permits more heterogeneity. Alternating k -triangles are much better than the Markov single triangle parameter in dealing with clumps of triangles that form the denser regions of the network. The parameter has a simple general interpretation: a large positive parameter value indicates that there is substantial triangulation in the network, and that this is likely to be expressed in the formation of denser regions.

Snijders et al. (2006) also proposed *k -paths*, configurations identical to k -triangles except that the edge at the base of the k -triangle is not necessarily present. This configuration quantifies multiple independent paths between pairs of nodes. Again, Snijders et al. combined these configurations into the one parameter, alternating k -paths. There is an associated distribution across the graph, the *dyad-wise shared partner distribution* (that is, shared partners based on dyads, not just on edges; see Hunter, in press; Hunter & Handcock, 2006.)

As the new degree parameter is based on a combination of Markov star configurations, it can be derived from Markov dependence. Markov dependence alone, however, is not sufficient to produce the alternating k -triangle and k -path parameters, which require higher order dependence structures (Pattison & Robins, 2002). The additional dependence assumption is referred to by Robins, Snijders, et al. (in press) as *social circuit dependence*, where the presence of two edges in the observed graph creates dependence among two other possible edges, assuming the four edges constitute a 4-cycle. Social circuit dependence appears complicated but it reflects a simple feature of social interaction. For example, if John and Mary work together, and if Joanne and Mark work together, then a working relationship between John and Joanne may increase the chances of a working relationship between Mary and Mark. But the argument simply does not work without the existing John/Mary and Joanne/Mark working relationships. In other words, the ties between John and Mary, and Joanne and Mark, create the dependence between possible John/Joanne and Mary/Mark ties.

This is a special, and rather different, dependence assumption. First, it explicitly permits the emergence of dependence through existing observations, in the sense that the presence of certain ties creates dependencies that otherwise could not exist. Secondly, this emergent dependence permits the appearance of higher-order structures (for example, “clumps” of triangles). Thirdly, the necessity of this assumption gives evidence to the importance of emergent processes in networks. It is a *self-organizing* quality, apparent in real social networks. Self-organizing systems imply that the presence or absence of certain ties may lead to the creation, maintenance, or disappearance of other ties. While the simpler Markov models can also be interpreted in terms of such self-organizing qualities, the social circuit dependence enables the appearance of higher order structures (e.g. dense regions of triangles) that are expressly implied by the model, rather than a simple chance accumulation of basic Markov configurations. Social networks are complex systems that typically cannot be adequately represented with simplified assumptions.

We note that these new specifications do not resolve all the problems of degeneracy and non-convergence. There are other forms of higher order dependence assumptions that might also be necessary for a particular data set. However, the new specifications have proven very adequate. Robins, Snijders et al (2006) show that the models containing these new dependence parameters perform dramatically better than Markov models in terms of convergence, when applied to a number of classic small-scale network data sets. Goodreau (in press) fits the new specifications to a network of over 1000 nodes and shows how to assess model fit across many graph features.

5 Simulation, Estimation, and Assessment

It is relatively straightforward to simulate p^* models, and estimate parameters (as mentioned below), using long-established statistical approaches such the Metropolis algorithm (Snijders, 2002) implementation of a Markov chain Monte Carlo. As first noted by Anderson, Crouch, and Wasserman (1999), if the model is not degenerate, the algorithm will “burn-in” to a stationary distribution of graphs reflecting the parameter values in the model. The length of the burn-in depends on the starting graph for the simulation, the complexity of the model, and the size of the network. For small networks of 30 nodes, for instance, non-degenerate models can burnin within a few tens of thousands of iterations, which can be achieved within seconds on a fast enough computer. It is then possible to sample a number of graphs from this distribution and look at typical features of them, for instance, the density, the geodesic distribution, the frequencies of various triads, and so on (Robins, Pattison, et al., in press). In other words, although the model is based on certain configurations, the graphs from the distribution typically will exhibit certain other features of interest that can be investigated.

These models are especially appealing not only because they are readily simulated but also because the parameters can be estimated from available data. In the past, p^* models were fitted using pseudo-likelihood estimation based on

logistic regression procedures (Strauss & Ikeda, 1990; see Anderson, et al., 1999, for a review). Although pseudo-likelihood can provide information about the data, especially in terms of identifying major effects (Robins, Snijders, et al., in press), when models are close to degeneracy or when dependency is strong, the precise pseudo-likelihood parameter estimates are likely to be misleading.

A more reliable way to fit the models is through Markov chain Monte Carlo Maximum Likelihood Estimation (MCMCMLE). There are various algorithms possible to do this (see Hunter & Handcock, 2006; Snijders, 2002). While the technical details are complicated, the underlying conceptual basis is straightforward. MCMCMLE is based on simulation (hence, the MCMC part of the acronym). A distribution of graphs is simulated from an initial guess at parameter estimates. A sample from the resulting graph distribution is compared to the observed graph to see how well the observed graph is reproduced by the modeled configurations. If it is not well-reproduced the parameter estimates are appropriately adjusted. If the model is well-behaved, this procedure usually results in increasingly refined parameter estimates, until finally the procedure stops under some criterion. We do note one large difference between Markov models and models containing parameters from the new specifications: the new specifications are more likely to be well-behaved and result in convergent parameter estimates.

Once estimates have been obtained, the model can be simulated and assessed. The assessment is accomplished by comparing a statistic calculated from the observed graph to the distribution of the statistic generated by the model. This can be seen as a (rather demanding) goodness of fit diagnostic for the model. Goodreau (in press) shows how this approach can be used to improve models by the addition of extra effects. It is also an approach that permits judgments about how well competing models might represent the network.

We note that currently, there are three programs publicly available for the simulation, estimation, and goodness of fit of p^* models:

- the *StOCNET* suite of programs from the University of Groningen
<http://stat.gamma.rug.nl/stocnet/> (especially *SIENA*)
- the *statnet* program from the University of Washington
<http://csde.washington.edu/statnet>
- the *pnet* program from the University of Melbourne
<http://www.sna.unimelb.edu.au>

6 New Ideas — Network Imputation

One of the most interesting developments in the statistical modeling of networks centers on an approach to estimate missing nodes and missing links. We refer to it as *network imputation*.

To outline these ideas, first suppose a given network \mathcal{X} is fit with a particular statistical model containing a set of network parameters collected in the vector θ . The likelihood given the data can then be denoted as

$$L(\theta|\mathbf{X} = \mathbf{x}),$$

where the dependence of the likelihood on the particular model is suppressed. The parameters in $\boldsymbol{\theta}$ are estimated as discussed earlier, with the associated estimates denoted by $\hat{\boldsymbol{\theta}}$. The likelihood evaluated at the estimated parameters is

$$L(\mathbf{x}; \hat{\boldsymbol{\theta}}).$$

In general, for a specific network, a model is adopted and a likelihood calculated. The likelihood is clearly dependent on the set of parameters that are chosen in the model. And if the observed network changes, so does the likelihood function. For our purposes, we will view the likelihood as a function not only of the parameters, but also of the data \mathbf{x} .

One of the most difficult problems in network analysis is determining whether the modeled network \mathbf{X} contains the complete set of nodes, \mathcal{N}_c , and the complete set of edges, \mathcal{E}_c . So, let us assume that the complete node set is

$$\mathcal{N}_c = \mathcal{N}_o \cup \mathcal{N}_m$$

and the complete edge set is

$$\mathcal{E}_c = \mathcal{E}_o \cup \mathcal{E}_m$$

where \mathcal{N}_o and \mathcal{E}_o are the observed set of nodes and edges, respectively, while \mathcal{N}_m and \mathcal{E}_m are the missing (that is, unobserved) set of nodes and edges, respectively. If $\mathcal{N}_m = \emptyset$ ($\mathcal{E}_m = \emptyset$), then all nodes (edges) are observed in \mathbf{x} ; otherwise, some nodes (edges) are missing and the goal is to impute the missing nodes and/or edges. For applications, it is of interest to estimate the components of \mathcal{N}_m and \mathcal{E}_m .

We briefly describe one technique for estimating missing edges for a fixed set of nodes.

Estimating missing edges

We briefly outline the procedure we have used to estimate missing links. We note that the distribution of \mathbf{X} is based on an approximate multivariate normality assumption of particular graph statistics. Specifically, we select a number of graph statistics, which, as we describe in Steinley and Wasserman (2006), are approximately Gaussian after transformations. We use only these statistics and assume they are statistically independent (which of course is probably far from the truth, but we view this work as just a first attempt at this). We then get a joint probability function and from that, a function that we can maximize (with respect to one missing link at a time). We assume a link is missing, and we add the link that maximizes the function.

We of course are assuming that the statistics chosen are the sufficient statistics for some underlying probability distribution.

The details follow:

1. Posit the distribution, \mathcal{D} , of \mathbf{X}

2. Based on \mathcal{D} , compute the initial likelihood of $\mathbf{X} = \mathbf{x}$ using the current estimates of θ , $L(\mathbf{x}; \hat{\theta})$, denoted as L .
3. Set $j = 1$. Let $L_j = L$.
4. The total possible number of edges in \mathcal{E}_c is $\binom{g}{2}$. Let the number of observed edges be O , consequently resulting in the number of unobserved edges as $\binom{g}{2} - O$.
5. For the i^{th} ($i = 1, \dots, \binom{g}{2} - O$) unobserved edge in \mathcal{E}_m (denoted as $\mathbf{X}^{(i)}$), change $\mathbf{X}^{(i)} = 0$ to $\mathbf{X}^{(i)} = 1$, reestimate θ (denoted as $\hat{\theta}^{(i)}$), and compute the associated likelihood $L_j^{(i)}$.
6. Repeat step 3 for all $i \in \mathcal{E}_m$.
7. Choose $L_j = \max_i L_j^{(i)}$
8. If $L_j > L$, permanently change $\mathbf{X}^{(i)}$ from 0 to 1; set $j = j + 1$ and $L = L_j$; repeat steps 3-7.
9. If $L_j \leq L$ stop adding edges to \mathbf{X} .

The basic idea is to incrementally “test” each edge that has not been observed to see if the addition of the link to the edge set increases the likelihood of the observed network. Additions of edges continues until the likelihood of the network can no longer be increased or until the incremental increase is small.

The procedure described here is the simplest way to look for missing edges within a network; however, there are many modifications that can be made to search for links. One of the most obvious (and perhaps most worth considering) is the possibility of adding multiple links at the same time and evaluate the resulting likelihood. For instance, the likelihood could also be evaluated when all possible pairs of links are also added to the graph. In fact, it would be possible to add triplets, quadruplets, etc., up to the logical conclusion of all missing nodes being added simultaneously and computing the likelihood of the complete graph. This procedure would be a complete enumeration task, becoming combinatorically infeasible for networks with any reasonably sized node set.

In addition to adding links, it would also be possible to add nodes to the network. The proposed algorithm for adding links to a network would become augmented where a sequential addition of nodes is simultaneously evaluated. Thus, one could consider the likelihood of the network with g nodes and then consider the likelihood of a network with $g + 1$ nodes and its possible links. Prior to conducting this procedure, the researcher is required to determine the maximum possible number of nodes that will be added to the observed network. Finally, it should be recognized that the described procedures for detecting missing nodes and links can be augmented *ad infinitum* to adapt to specific network structures hypothesized by researchers. The only caveat is that additional structural properties imposed by the researcher can have effects (possibly adverse) on the final solution.

Acknowledgements

This research was supported by grants from the National Science Foundation and the US Office of Naval Research, and the Australian Research Council. We thank Stephen Fienberg for comments on this chapter.

References

- [1] Anderson, C.J., Wasserman, S., and Crouch, B. (1999). A p^* primer: Logit models for social networks. *Social Networks*. 21, 37–66.
- [2] Borner, K., Sanyal, S., & Vespignani, A. (in press). Network science: A theoretical and practical framework. In Blaise Cronin (Ed.), *Annual Review of Information Science & Technology*, Volume 4. Medford, NJ: Information Today, Inc./American Society for Information Science and Technology.
- [3] Carrington, P.J., Scott, J., and Wasserman, S. (eds.) (2005). *Models and Methods in Social Network Analysis*. New York: Cambridge University Press.
- [4] Frank, O., and Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*. 81, 832–842.
- [5] Goodreau, S. (in press). Applying advances in exponential random graph (p^*) models to a large social network. *Social Networks*.
- [6] Handcock, M.S. (2002). Statistical models for social networks: Degeneracy and inference. In Breiger, R., Carley, K., & Pattison, P. (eds.). *Dynamic Social Network Modeling and Analysis* (pp. 229-240). Washington DC: National Academies Press.
- [7] Holland, P.W., and Leinhardt, S. (1977). Notes on the statistical analysis of social network data.
- [8] Holland, P. W., and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*. 76, 33–65 (with discussion).
- [9] Hoff, P., Raftery, A., & Handcock, M. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97, 1090-1098.
- [10] Hunter, D.R. (In press). Curved exponential family models for social networks. *Social Networks*.
- [11] Hunter, D. & Handcock, M. (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*. 15, 565–583.
- [12] Park, J., & Newman, M. (2004). Solution of the 2-star model of a network. *Physical Review E*, 70, 066146.
- [13] Pattison, P. E., & Robins, G. L. (2002). Neighbourhood-based models for social networks. *Sociological Methodology*. 32, 301-337.
- [14] Pattison, P.E., and Wasserman, S. (1999). Logit models and logistic regressions for social networks: II. Multivariate relations. *British Journal of Mathematical and Statistical Psychology*. 52, 169–193.
- [15] Robins, G.L., Elliot, P., and Pattison, P.E. (2001). Network models for social selection processes. *Social Networks*. 23, 1–30.
- [16] Robins, G.L., Pattison, P.E., and Elliott, P. (2001). Network models for social influence processes. *Psychometrika*. 66, 161–190.
- [17] Robins, G.L., Pattison, P.E., Kalish, Y., & Lusher, D. (in press). An introduction to exponential random graph (p^*) models for social networks. *Social Networks*.
- [18] Robins, G.L., Pattison, P.E., and Wasserman, S. (1999). Logit models and logistic regressions for social networks, III. Valued relations. *Psychometrika*. 64, 371–394.

- [19] Robins, G.L., Pattison, P.E., & Woolcock, J. (2005). Social networks and small worlds. *American Journal of Sociology*. 110, 894-936.
- [20] Robins, G.L., Snijders, T.A.B., Wang, P., Handcock, M., & Pattison, P.E. (In press). Recent developments in exponential random graph (p^*) models for social networks. *Social Networks*.
- [21] Snijders, T.A.B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure*. 3, 2.
- [22] Snijders, T.A.B., Pattison, P.E., Robins, G.L., & Handcock, M. (2006). New specifications for exponential random graph models. *Sociological Methodology*.
- [23] Snijders, T.A.B., Steglich, C., & Schweinberger, M. (in press). Modeling the co-evolution of networks and behavior. To appear in van Monfort et al (Eds.), *Longitudinal Models in the Behavioral and Related Sciences*. New York: Erlbaum.
- [24] Strauss, D., & Ikeda, (1990). Pseudo-likelihood estimation for social networks. *Journal of the American Statistical Association*. 85, 204-212.
- [25] Strauss, D., and Ikeda, M. (1990). Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association*. 85, 204-212.
- [26] Wasserman, S., and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
- [27] Wasserman, S., and Pattison, P.E. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov random graphs and p^* . *Psychometrika*. 60, 401-426.
- [28] Wasserman, S., and Robins, G.L. (2005). An introduction to random graphs, dependence graphs, and p^* . In Carrington, P.J., Scott, J., and Wasserman, S. (eds.), *Models and Methods in Social Network Analysis*. New York: Cambridge University Press.