Glossary

General:

Dependence assumption: Theoretical assumption about dependencies among possible network ties; determines the type of parameters in the model.

Exponential random graph models: a model for a social network, expressing a probability distribution of graphs with an exponential form: see also *p**.

Homogeneity assumption: Assumption about which parameters to equate, to make the model identifiable.

Graph statistics: for homogeneous models these are counts of the configurations in the observed graph; more generally they may depend also on node-wise or dyad-wise covariates

Network Configuration: A small sub-graph that may be observed in the data and that is represented by parameters in the model: eg reciprocated ties, triangles.

Parameters: relate to specific network configurations that may be observed in the graph; a large positive parameter is interpreted as the presence of more of the configurations than might be expected from chance (given the other effects in the model); a large negative parameter signifies the relative absence of the configuration.

p:* the term for exponential random graph models introduced by Wasserman and Pattison (1996).

Edge and dyad independence models:

Dyad independence: assumes that dyads are independent of one another; the model includes edge and reciprocity parameters, and possibly also node or dyad attributes.

p1 models (Holland and Leinhardt): an early dyad independence model, including popularity and expansiveness effects.

p2 model: elaboration of *p*1 model, where popularity and expansiveness effects are random, and independent variables may be used to predict ties.

Simple random graphs, Bernoulli graphs, Erdos-Renyi graphs: assume that edges are independent of one another and are observed with a given probability.

Markov random graph models

alternating k-stars: a Markov parameter (and statistic) but used in the new higher order models; a particular combination of Markov *k*-star counts into the one statistic; equivalent to *geometrically weighted degree counts*; useful for modelling the degree distribution.

cyclic triad: a Markov graph configuration: in a directed network, ties *ij, jk* and *ki* are observed among actors *i, j,* and *k*.

degeneracy (or near-degeneracy): when a model implies that very few distinct graphs are probable, often only empty or complete graphs; degenerate models cannot be good models for social network data.

geometrically weighted degree counts: a statistic (and parameter) in higher order models: a sum of degree counts with geometrically decreasing weights; equivalent to *alternating k-stars.*

k-star: a Markov graph configuration: in a non-directed graph, *k* edges are expressed by the one actor.

k-in-star: a Markov graph configuration: in a directed graph, *k* arcs are directed to the one actor.

k-out-star: a Markov graph configuration: in a directed graph, *k* arcs are expressed by the one actor.

Markov dependence assumption: introduced by Frank and Strauss (1986), proposes that, conditional on the rest of the graph, two possible ties are dependent on one another when they share an actor.

mixed-star: a Markov graph configuration: a two path in a directed graph.

transitive triad: a Markov graph configuration: in a directed network, ties *ij, jk* and *ik* are observed among actors *i, j,* and *k.*

triangle: a Markov graph configuration: in a non-directed network, a clique of three actors, ties *ij, jk* and *ik* are observed among actors *i, j,* and *k*.

Higher order models

alternating independent-2-paths: a parameter (and statistic) in higher order models; a particular combination of *k-independent-2-path* counts into the one statistic; when this parameter is negative, together with a positive alternating k-triangle parameter, there is a tendency against 4-cycles in the network, unless those cycles include triangles (alternatively, the presence of many 2-paths between nodes is related to the formation of triangles.)

alternating k-triangles: a parameter (and statistic) in higher order models; a particular combination of *k-triangle* counts into the one statistic; expresses the tendency for many triangles to form together in the observed network; a positive parameter in the model suggests regions in the network of high triangulation, possibly core-periphery-type structures; a positive parameter, together with a negative alternating *k*-star parameter, suggests several smaller regions (possibly connected) of triangulation; equivalent to *weighted shared partners.*

Dyad-wise shared partners (dsp): a parameter (and statistic) in the higher order models; expresses the tendency in the observed network for dyads (whether tied or not) to have multiple shared partners; equivalent to *alternating independent 2-paths*.

Edge-wise shared partner distribution: Distribution of the number of dyads who are themselves related and who have a fixed number of shared partners.

Edge-wise shared partners (esp): a parameter (and statistic) in the higher order models; expresses the tendency in the observed network for tied nodes to have multiple shared partners; equivalent to *alternating k-triangles*.

k-triangle: a configuration in higher order models; in a non-directed graph, the combination of *k* triangles, each sharing the one edge (the base of the *k*-triangle).

k-independent-2-paths: configurations in the higher order models; equivalent to *k*-triangles but without the base.

partial dependence assumption (Pattison & Robins, 2002)*:* assumption for dependencies among possible ties created by the presence of other ties; permits models with higher order configurations than Markov configurations. For example, X_{ij} and X_{kl} are conditionally dependent if $x_{ik} = x_{jl} = 1$ or if $x_{il} = x_{jk} = 1$.

Estimation

Monte Carlo Markov Chain maximum likelihood estimation (MCMCMLE): Method of estimation based on computer simulation; more principled than pseudolikelihood; produces reliable standard errors.

Pnet: (Wang, Robins, & Pattison, 2005). Software that includes procedures for MCMCMLE for exponential random graph models – University of Melbourne, Australia.

Pseudo-likelihood estimation: an approximate method of estimation using logistic regression; does not produce reliable standard errors; properties are not well understood.

Statnet: (Handcock, Hunter, Butts, Goodreau, & Morris, 2005). A software package using *R*, including procedures for MCMCMLE for exponential random graph models – University of Washington.

SIENA: (Boer, Huisman, Snijders, & Zeggelink, 2003). A procedure within the StOCNET software package that includes provisions for MCMCMLE for exponential random graph models – University of Groningen, the Netherlands. (http://stat.gamma.rug.nl/StOCNET)