Reprint of

Snijders, T.A.B., and van Duijn, M.A.J.,
Simulation for statistical inference in dynamic network models.
In: Conte, R., Hegselmann, R., and Terna, P. (Eds.),
Simulating Social Phenomena.
Lecture Notes in Economics and Mathematical Systems, 456.
Berlin: Springer, 1997, 493-512.

Simulation for Statistical Inference in Dynamic Network Models

Tom Snijders and Marijtje van Duijn

ICS/Department of Statistics and Measurement Theory, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands.

Keywords: social networks, statistical modeling, actor-oriented model, continuous-time Markov chain, Robbins-Monro process.

Abstract
Actor-oriented models are proposed for the statistical analysis of longitudinal social network data. These models are implemented as simulation models, and the statistical evaluation is based on the method of moments and the Robbins-Monro process applied to computer simulation outcomes. In this approach, the calculations that are required for statistical inference are too complex to be carried out analytically, and therefore they are replaced by computer simulation.

The statistical models are continuous-time Markov chains. It is shown how the reciprocity model of Wasserman and Leenders can be formulated as a special case of the actor-oriented model.

1 Introduction

Social networks are structures consisting of dyadic relations, or ties, between individuals or other units (organisations, countries, etc.); a commonly used term for these units is *actors*, stressing their active role in the constitution of the networks. The relations between two individuals are not necessarily symmetric. Examples of such dyadic relations are friendship, esteem, cooperation, etc. An introduction to social network analysis is given by (Wasserman & Faust 1994).

This paper is concerned with *entire networks*, where all relationships within a given set of n actors are considered. Such a network can be represented by an $n \times n$ matrix $X = (X_{ij})$, where X_{ij} is a number or a vector representing the relation directed from actor i to actor j $(i, j = 1, ..., n)$. Self-relations are not considered, so that the diagonal values X_{ii} are meaningless. In this paper we only consider dichotomous relations: the relation (e.g., friendship) can be present from i to j, denoted $X_{ij} = 1$, or it can be absent, denoted $X_{ij} = 0$. The diagonal values are formally defined as $X_{ii} = 0$. The network can be represented by a directed graph, with X as its adjacency matrix.

The statistical analysis of social networks is difficult because each data point, X_{ij}, refers to two individuals rather than to just one. This leads to a complicated dependence structure of the elements of X. Some examples of interesting kinds of dependence are the following.

- ⋆ *Reciprocity* : the dependence between X_{ij} and X_{ji}. This type of dependence is so important that often these two reciprocal relations are considered jointly. The pair (X_{ij}, X_{ji}) is called a *dyad*.
- ⋆ The dependence between the elements of each *row*, corresponding to out-going relations of the same actor i. This actor is called the *sender* of these relations. The out-degree of an actor, $\sum_j X_{ij}$, is an indicator for the 'activity' of sender i.
- ⋆ The dependence between the elements of each *column*, corresponding to in-coming relations of the same actor i. This actor is called the *receiver* of these relations. The in-degree of an actor, $\sum_j X_{ji}$, is an indicator for the 'popularity' of receiver i.
- ⋆ *Transitivity*: "a friend of my friend is also my friend", which implies a dependence between triples of actors.
- ⋆ *Group formation*, which implies a dependence between sets of three or more actors.

Longitudinal data is much more informative about the studied phenomena in social networks than cross-sectional data, but creates additional problems for statistical analysis. An interesting collection of papers about longitudinal social networks is (Doreian & Stokman 1997).

Example. As an example, we use the Electronic Information Exchange System (EIES) data collected by (Freeman & Freeman 1979) and reproduced in (Wasserman & Faust 1994). More information can be found in these references. We use complete data on 32 researchers who participated in a study on the effects of electronic information exchange. Two measures of acquaintanceship are used, collected before and after the study (8 months apart). The data as reproduced by Wasserman & Faust were dichotomized: 1 ("positive relation") for having met or being a friend (or close friend) of the other, 0 ("null relation") for not knowing or at least not having met the other. In addition, a dichotomous individual-bound covariate w_i is used: the number of citations of the researcher's work in the *SSCI* in the year before the research started. This variable was dichotomized: $w_i = 0$ for 12 or less citations, $w_i = 1$ for more than 12 citations.

At the first measurement, 513 of the $n(n-1) = 992$ directed relations were positive, which leads to a density of $513/992 = 0.52$. Of these positive relations, 7 had changed to a null relation at the second measurement (this change from "knowing" to "not knowing" is rather unlikely, and happened very infrequently), while of the 479 null relations at the first time point, 147 had changed to a positive relation at the second time point. Thus the density at the second time point had increased to $653/992 = 0.66$. Below we shall study the structure of change that took place.

The present paper is about statistical procedures for time-series $X(t), t \in \mathcal{T}$ of social networks for a constant set $\{1, ..., n\}$ of actors, where the set of observation times is finite, $\mathcal{T} = \{t_1, ..., t_M\}$. The examples are longitudinal data with $M = 2$ observation times, but the methods treated can also be used for larger numbers of observation times.

Existing methods for longitudinal social networks can be roughly divided into two types, each with their own shortcomings: simulation models that are not a suitable basis for data analysis because either they are deterministic, or they are are not accompanied by methods for relating the model to observational data; and statistical models with unrealistic assumptions or a lack of flexibility.

If we wish to compare, on the basis of empirical evidence, several simulation models based on different assumptions or theories, then we need some kind of statistical framework. In our view it is preferable to use stochastic simulation models because they explicitly take into account the uncertain nature of observational data. The stochastic element has to be included in such a way that it can be interpreted as a source of unexplained variability, similar to the error term in linear regression analysis. If stochastic simulation models also include a number of unknown parameters that can be "fitted to data", then in principle it is possible to use them for statistical inference: parameter estimation, tests of parameters and of goodness of fit. Such an approach will eventually lead to more realistic simulation models and more theoretically relevant data analysis.

Realistic models for longitudinal social network data are necessarily quite complex. The lack of practical statistical methods for dealing with such models has been a severe restriction for the development of a methodology for longitudinal social networks analysis. Overviews of some statistical models for longitudinal social networks are given by (Frank 1991) and (Snijders 1995). Earlier statistical methods for change in networks succeeded in taking account of reciprocity, sender, and receiver effects, but not of more involved effects such as transitivity or group formation. The most promising earlier models are the continuous-time Markov chain models proposed by (Wasserman 1977, 1979, 1980) and (Leenders 1995a, 1995b). However, these models still assume *conditional dyad independence*, i.e., when t_1 and t_2 are consecutive observation times, they assume that, conditional on $X(t_1)$, the dyad $(X_{ij}(t_2), X_{ji}(t_2))$ is stochastically independent of all other dyads. This assumption effectively allows to change the analysis from the level of the network to the level of the dyad. This is computationally attractive, but does not leave much room for realistic statistical modeling. Effects such as transitivity, that lead to dependence in the relations between sets of three or more actors, cannot be represented by models with conditional dyad independence.

A solution to this problem was proposed by (Snijders 1996) in the form of so-called *stochastic actor-oriented models*. In these models non-deterministic rules are formulated that govern the behavior of actors in the network. Actors' behavior is defined as changing their own relations by choosing from several alternatives and is aimed at maximization of an objective function under constraints. The objective function may be regarded as a utility, or expected utility, function. The objective function and the constraints are actor-dependent. The position of the actor in the network is an important part of the constraints.

In contrast to many usual simulation models, the objective functions include a random element. This disturbance makes the model stochastic, and thereby allows deviations between predicted and observed outcomes.

The statistical character of the model enables the estimation of its parameters. The estimation uses the method of moments, implemented with the Robbins-Monro algorithm and computer simulation. Thus, since the parameters of the statistical model can be tested, the underlying theory as expressed in the objective function can be tested as well. The statistical model is presented in more detail in the next section.

This approach uses simulation models not as theoretical metaphors, but as statistical models for data. This means that the simulation models must be taken more seriously than is sometimes done, and that empirical data must be used to develop the statistical model in order to obtain an adequate fit between model and data.

This paper focuses, first, on the formulation of actor-oriented models for dichotomous relations, i.e., for social networks represented as directed graphs. In the model of (Snijders 1996), the actors' actions are propelled by utility functions, or objective functions, reflecting their evaluation of given network configurations. In the present paper, not only an objective function but also a "gratification function" is included. This function reflects the instantaneous evaluation of a change of the actor's relations. This creates a greater flexibility in modeling. The second focus of this paper is on the relation between the actor-oriented models and the dyadic independence models of Wasserman and Leenders. It will be shown that the gratification function is necessary to formulate the dyadic independence models as a special case of actor-oriented models.

The authors are working on the empirical application of these models, cf. (Van de Bunt, Van Duijn & Snijders 1995).

2 Stochastic actor-oriented models for change in networks

In this section we develop the principle of actor-oriented models for the standard social network data structure of directed graphs, i.e., a dichotomous relational variable X_{ij}. (In (Snijders 1996), these models were proposed for a more special data structure in view of an example application to the data set of (Newcomb 1961).) It is assumed here that the set of actors is fixed. The number of actors is denoted n.

2.1 Basic model ingredients

A directed graph X can be represented by its "sociomatrix", or adjacency matrix, i.e., an $n \times n$ matrix X_{ij} with elements that can be 0 or 1, depending on whether the given relationship is absent ($X_{ij} = 0$) or present ($X_{ij} = 1$) from i to j. The class of all sociomatrices, i.e., of all $n \times n$ matrices of 0-1 elements with a zero diagonal, is denoted by $\mathcal{X}$. We consider data consisting of a time series $X(t_m),\ m = 1, ..., M$ of directed graphs, and statistical models where such a time series is embedded in an (unobserved) continuous-time process $X(t)$ with $t_1 \le t \le t_M$. The reasons for assuming an underlying continuous-time process are, first, the fact that in reality change also takes place between the observation times and, second, that this assumption permits a simpler and more straightforward approach than discrete time modeling.

It is assumed that each actor "controls" his outgoing relations, which are collected in the row vector $(X_{i1}(t), ..., X_{in}(t))$. At stochastic times, with a distribution determined by the functions λ_i introduced below, the actors have the opportunity to change these outgoing relations. When an actor changes his outgoing relations, he is assumed to pursue two "goals": attaining a rewarding configuration for himself in the network; and instaneous gratification inherent in the action of a specific change. (The word "gratification" must be understood in a generalized sense; this component can stand for, e.g., minus the costs associated with making a given change.) These two goals are modeled in the functions f and g below. In addition, the actions of each actor are propelled by a random component, representing the actor's drives that are not explicitly modeled. This actor-oriented model represents the idea that actors pursue their own goals under the constraints of their environment, while they themselves constitute each others' changing environment (cf. Zeggelink, 1994).

The actors act independently, given the current network structure. At any single time point, at most one actor may change his outgoing relations. Furthermore, he may change only one relation at the time. Of course, many small changes between two observation times can result in a big difference between the two observed networks. The fact that the model specification focuses on changes of single relations is the major reason why continuous time modeling is relatively straightforward. (An example of a continuous-time model where more than one relation can change at one time point is given by (Mayer, 1984).)

The model specification is given by the following three families of functions, all depending on a K-dimensional statistical parameter θ that assumes values in an open set $\Theta \subset \mathbb{R}^K$. This θ plays the usual role of a statistical parameter, and methods will be proposed for estimating this parameter from the data.

1. A family of *rate functions*

$$\lambda_i(\theta, x)\,, \quad i = 1, ..., n, \quad x \in \mathcal{X}\,, \tag{1}$$

which indicate the rate at which actor i is allowed to change something in his outgoing relations.

2. A family of *objective functions* with respect to the network configuration,

$$f_i(\theta, x)\,, \quad i = 1, ..., n, \quad x \in \mathcal{X}\,, \tag{2}$$

which indicate the preference of actor i for the relational situation represented by x.

3. A family of *gratification functions*

$$g_i(\theta, x, j)\,, \quad i, j = 1, ..., n, \quad i \neq j, \quad x \in \mathcal{X}\,, \tag{3}$$

which indicate the instantaneous gratification experienced by actor i when, from the given network configuration x, element x_{ij} is changed into its opposite, $1 - x_{ij}$.

Whenever actor i has the opportunity to change his outgoing relations, he changes one relation, say x_{ij}. He can withdraw an outgoing tie to one of the actors to whom he has such a tie, or initiate an outgoing tie to one of the actors to whom he does not have a tie. The network that results when the single element x_{ij} is changed into $1 - x_{ij}$, is denoted by $x(i \leadsto j)$. When the current network is x, actor i has the choice between changes to $x(i \leadsto j)$ for all different j. The momentary total objective function maximized by i is the sum of the actor's preference for the new state, the gratification experienced as a result of the change, and a random element:

$$f_i(\theta, x(i \leadsto j)) + g_i(\theta, x, j) + U_i(t, x, j)\,. \tag{4}$$

The term $U_i(t, x, j)$ is a random variable, indicating the part of the actor's preference that is not represented by the systematic components f_i and g_i . In this paper, it is assumed that these random variables are independent and identically distributed for all i, t, x, j.

Markov chain with random utility component

These functions are used in the following way to define a continuous-time Markov chain $X(t)$ with the finite outcome space $\mathcal{X}$. (For an introduction to continuous time Markov chains, see (Norris, 1997)). Events, i.e., changes of the network structure, take place at discrete time points; in between these points, the network structure remains constant. The process is modeled as being right-continuous: if a change takes place from state x_0 to state x_1 at time t_0, then $X(t) = x_0$ for t sufficiently close to, but smaller than t_0, while $X(t) = x_1$ for $t = t_0$ and also for t sufficiently close to, but larger than t_0.

The n actors are acting independently, given the current state of the network. Each of them has the individual change rate $\lambda_i(x, \theta)$. At each time point t, the time until the next change by *any* actor has the negative exponential distribution with parameter

$$\lambda_+(\theta, x) = \sum_{i=1}^{n} \lambda_i(\theta, x) \ , \text{ where } x = x(t) \ . \tag{5}$$

The parameter of the negative exponential distribution is taken here as the reciprocal of the expectation, so the expected waiting time until the next change after time t is $1/\lambda_+(\theta, x(t))$. Given that an event occurs, the actor who may change his out-relations is actor i with probability

$$\frac{\lambda_i(\theta, x)}{\lambda_+(\theta, x)} \ . \tag{6}$$

Given that actor i may change his outgoing relations, he chooses to change his relation to that actor j ($j \neq i$) for whom the value of the momentary total utility function (4) is highest.

It is convenient to let the $U_i(t, x, j)$ have the type 1 extreme value distribution with mean 0 and scale parameter 1. This assumption is commonly made in random utility modeling in econometrics, cf. (Maddala, 1983). When this distribuition is used, and denoting the systematic part of the momentary objective function by

$$r(\theta, i, j, x) = f_i(\theta, x(i \rightsquigarrow j)) + g_i(\theta, x, j) \, ,$$

the probability of change j is given by the multinomial logit expression, cf. (Maddala 1983, 60),

$$p_{ij}(\theta, x) = \frac{\exp(r(\theta, i, j, x))}{\sum_{h=1, h\neq i}^{n} \exp(r(\theta, i, h, x))} \quad (j \neq i). \tag{7}$$

This expression will be used further in this paper.

From expression (4) or (7) it follows that this probability does not change when to $r(\theta, i, j, x)$ a term is added that does not depend on j. It is often more convenient to work with

$$r(\theta, i, j, x) = f_i(\theta, x(i \rightsquigarrow j)) - f_i(\theta, x) + g_i(\theta, x, j) \ . \tag{8}$$

The instantaneous effect g_i is a more general model component than the objective function f_i , since (8) itself could be used as the gratification function, without the need also to have an objective function. The reverse, however, is not true: a non-trivial gratification function cannot always be expressed as a difference between objective function values. The reason for not working with just the gratification function is that the objective function, attaching a value to each network configuration, is conceptually more attractive and better interpretable than the instantaneous gratification effect.

Intensity matrix

Transition distributions of continuous-time Markov chains are characterized by their *intensity matrix*, *infinitesimal generator*, or *generator matrix*, cf. (Norris, 1997). In our case, where relations are allowed to change only one at a time, this generator matrix can be represented[1] by the functions $q_{ij}(x)$, indicating the change rates of x to $x(i \rightsquigarrow j)$. These functions are defined as

$$q_{ij}(x) = \lim_{dt \downarrow 0} \frac{\mathrm{P}\{X(t+dt) = x(i \rightsquigarrow j) \mid X(t) = x\}}{dt} \quad i,j = 1,...,n,\ i \neq j\,. \tag{9}$$

It can be proven that this is given by

$$q_{ij}(x) = \lambda_i(\theta, x)\, p_{ij}(\theta, x)\,. \tag{10}$$

Computer simulation

With the specifications given here, a computer simulation of this stochastic process can be set up. It is convenient to construct the continuous-time Markov chain as the combination of its *holding times* and its *jump process*, cf. (Norris 1997, Section 2.6). The holding times are the times between consecutive changes, and have the negative exponential distribution with parameter (5). The jump process is the process of consecutive distinct states visited by the Markov chain. The simulation algorithm is as follows; S is the holding time and x the outcome of the jump process.

1. Set $t = 0$ and $x = X(0)$.
2. Generate S according to the negative exponential distribution with mean $1/\lambda_+(\theta, x)$.
3. Select randomly $i \in \{1, ..., n\}$ using probabilities (6).
4. Select randomly $j \in \{1, ..., n\}$, $j \neq i$ using probabilities (7).
5. Set $t = t + S$ and $x = x(i \rightsquigarrow j)$.
6. Go to step 2 (unless the stopping criterion is satisfied).

2.2 Specification of the model

The model is specified by the choice of the functions λ_i, f_i, and g_i and the way in which they depend on the K-dimensional parameter θ. In the mathematically simplest case, the change rates $\lambda_i(x)$ are constant, e.g., $\lambda_i(\theta, x) \equiv \theta_1$.

[1]For those who know the theory of continuous-time Markov chains, it will be clear that the generator matrix of this chain is a matrix with $2^{n(n-1)}$ rows and columns, filled at appropriate places with the elements q_{ij} as defined here, and with zeros elsewhere.

The change rates can also depend on the position in the group (e.g., actors who are dissatisfied with their relation might change faster than those who are more satisfied) or on actor-dependent characteristics (e.g., newcomers in the group may change faster than those who have been in the group longer). Some specifications of the change rate are mentioned in the section on the reciprocity model.

The functions f_i and g_i must contain the substantive ingredients of the model, including, e.g., actor attributes and structural properties of the directed graph. Since the actor has direct control only of his outgoing relations, it is irrelevant in this model to have components in f_i or g_i that are a function only of other parts of the directed graph.

A possible choice for f_i is to define it as a sum of some of the following terms, where the weights θ_2, θ_3, etc., are statistical parameters indicating the strength of the corresponding effect, controlling for all other effects in the model. For effects (3) and (4), it is assumed that an actor-bound covariate W, with values w_i, is available. All formulae indicate a contribution to the objective function of actor i, while the other actors to whom i could be related are indicated by j.

1. $\theta_2 \sum_j x_{ij}$: out-degree; θ_2 reflects the value of activity for actor i;
2. $\theta_3 \sum_j x_{ij} x_{ji}$: number of reciprocated relations; θ_3 reflects the value of reciprocated relations;
3. $\theta_4 \sum_j x_{ij} w_j$: the sum of the covariate over all actors to whom i has a relation; θ_4 reflects the aspiration of the actor to have relations with others who score high on W;
4. $\theta_5 \sum_j x_{ij} \,|\, w_i - w_j \,|$: the sum of absolute covariate differences between i and the others to whom he is related; θ_5 reflects the preference for similar others;
5. $\theta_6 \sum_j x_{ij} \sum_h x_{hj}$: the sum of the popularity (as measured by the in-degree $\sum_h x_{hj}$) of all actors to whom i is related; θ_6 reflects the preference for popular others;
6. $\theta_7 \; \#\{j \mid x_{ij} = 0, \, \sum_h x_{ih} x_{hj} > 0\}$: the number of actors i is indirectly related to (through one intermediary, i.e., at sociometric distance 2); θ_7 reflects the value of indirect relations;
7. $\theta_8 \sum_{j,h} x_{ij} x_{ih} x_{jh}$: the number of transitive patterns in i's relations (ordered pairs of actors (j, h) to whom i is related, while also j is related to h); θ_8 reflects the value of having relations to others who are related among themselves.

Examples of terms that can be included in the instantaneous effects $g_i(\theta, x, j)$ are the following. Note that the presence of a factor x_{ij} in a term in g_i indicates that this term refers to breaking off an existing relationship, while a factor $(1 - x_{ij})$ refers to establishing a new relation.

1. $\theta_9 \, x_{ij} x_{ji}$: indicator of a reciprocated relation; a negative value of θ_9 reflects the costs associated with breaking off a reciprocated relation.

2. $\theta_{10} \, (1 - x_{ij}) \sum_h x_{ih} x_{hj}$: the number of actors through whom i is indirectly related to j; a positive value of θ_{10} reflects that it is easier to establish a new relation to another actor j if i has many indirect relations to j via others who can serve as an introduction.

These lists can be extended with other components. Theoretical insights in the relational process and experience with modeling the type of data have to determine the effects that are included.

3 Statistical estimation

The functions λ_i , f_i , and g_i depend on a parameter θ that must be estimated from the data. The available data are the observed digraphs $x(t), \ t = t_1, ..., t_M$ $(M \geq 2)$; and covariates if these are included in the functions λ_i , f_i , or g_i . The estimation methods considered here condition on the observed value $x(t_1)$ at the first time point, and do not make the assumption of a stationary distribution for $X(t)$. The approach to estimation followed is the same as proposed in (Snijders 1996), where further elaboration can be found.

3.1 Method of moments

Assume first that $\lambda_i(x) = \theta_1$ and that

$$f_i(\theta, x) = \sum_{k=1}^{K-1} \theta_{k+1} \, s_{ik}(x) \ , \quad g_i \equiv 0 \ ,$$

where the $s_{ik}(x)$ are suitable digraph statistics such as those mentioned in the list of examples above.

The method of moments, cf. (Bowman & Shenton, 1985), is used to estimate θ. For this estimation method we need K statistics that carry information about the K parameters θ_k. For θ_1, a relevant statistic is the amount of change, measured by the sum of distances between successive observations,

$$C = \sum_{m=2}^{M} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} |X_{ij}(t_m) - X_{ij}(t_{m-1})| \ . \tag{11}$$

For θ_{k+1} with $k = 1, ..., K-1$, a relevant statistic is the sum over all actors i of the digraph statistics s_{ik}, as observed at the final observation moment:

$$S_k = \sum_{i=1}^{n} s_{ik}(X(t_M)) \ . \tag{12}$$

These statistics are collected in the vector $Z = (C, S_1, ..., S_{K-1})$. Stochastic variables and observations will be distinguished in the notation by indicating the former by capitals and the latter by lower case letters. Accordingly, z denotes the observed value for Z.

The moment estimate $\hat{\theta}$ is defined as the parameter value for which the expected value of the statistic is equal to the observed value:

$$\mathcal{E}_{\hat{\theta}} Z = z \ . \tag{13}$$

The statistical efficiency of this estimator depends, of course, on the choice of the statistics Z, cf. (Bowman & Shenton 1985) and (Snijders 1996).

For model specifications with more complex functions λ_i , and with non-zero g_i , the vector of statistics Z has to be chosen in other ways, so that it is informative about the parameter values. More research is needed about this choice; some illustrations are given below.

3.1.1 Robbins-Monro process

Equation (13) cannot be solved by analytical or the usual numerical procedures, because (except for some simple cases) the expected value $\mathcal{E}_\theta Z$ cannot be calculated explicitly. However, the solution can be approximated by the Robbins-Monro method (proposed by (Robbins & Monro 1951); for an introduction see, e.g., (Ruppert, 1991)). The Robbins-Monro method is a stochastic approximation algorithm that yields values $\hat{\theta}_N$ by an iterative simulation process. If certain conditions are satisfied, this sequence converges to the solution of the moment equation (13). The iteration step of the Robbins-Monro algorithm is defined as

$$\hat{\theta}_{N+1} = \hat{\theta}_N - \frac{1}{N} D_N^{-1}(z_N - z) \ , \tag{14}$$

where D_N is a suitable matrix. The optimal value of D_N is the derivative matrix $D_\theta = (\partial \mathrm{E}_\theta Z / \partial \theta)$. In adaptive Robbins-Monro procedures, this derivative matrix is estimated during the approximation process. In (Snijders 1996) it is proposed to compute D_N by estimating the derivative matrix using common random numbers. In practice it cannot be guaranteed in the applications considered here that the process will converge, and it is advised to check the approximate validity of (13) for the solution found by carrying out a number of simulation runs for the found value $\hat{\theta}$.

A loose description of the estimation algorithm using the Robbins-Monro algorithm is the following.

1. Choose a starting value $\hat{\theta}_1$ and a suitable matrix D_1. Set $N = 1$.

2. For $m = 1, ..., M - 1$:

   (i) Set $x = x(t_m)$;
   (ii) Simulate the continuous time Markov chain starting at $t = t_m$ and continuing until $t = t_{m+1}$, with parameter $\theta = \hat{\theta}_N$. (The algorithm for this simulation was given above.)

3. Compute from this simulation the vector of statistics Z and denote its outcome by z_N.

4. Define $\hat{\theta}_{N+1}$ by (14), set $N = N + 1$ and update D_N.

5. Go to step 2, unless the convergence criterion (cf. (Snijders, 1996)) is satisfied.

An alternative estimation method on the basis of simulations of the probabilistic model is the method of simulated moments proposed by (McFadden 1989) and by (Pakes & Pollard 1989), also see (Gouriéroux & Montfort 1996).

4 Dyadic Markov models for network change

Models for change in digraphs were proposed by (Wasserman 1977, 1979, 1980). In these models, the dyads are independent. The models are continuous-time Markov chains, like the actor-oriented models. Therefore they are completely specified by the infinitesimal generator matrix. Since these, too, are models where relations change only one at a time, they are determined completely by the matrix (q_{ij}) defined in (9).

The basic model in this class is the *reciprocity model*, where the change rate from $x_{ij} = 0$ to $x_{ij} = 1$ is defined by

$$q_{ij}(x) = \lambda_0 + \mu_0 x_{ji} \quad (x_{ij} = 0)\,, \tag{15}$$

while the change rate from $x_{ij} = 1$ to $x_{ij} = 0$ is defined by

$$q_{ij}(x) = \lambda_1 + \mu_1 x_{ji} \quad (x_{ij} = 1)\,. \tag{16}$$

The parameters $\lambda_0, \lambda_1, \mu_0, \mu_1$ are allowed to depend on dyad-bound covariates, e.g., variables indicating the similarity between actors i and j; see (Leenders 1995a). Since the change rates depend only on functions of the dyad, the dyads are independent in this model. A method for calculating maximum likelihood (ML) estimators for these parameters was given by (Wasserman 1977), with a correction by (Leenders 1995a).

The advantage of the reciprocity model is that ML estimators can be calculated numerically. The disadvantage is the restrictive assumption of dyad independence. This precludes the modeling of effects that involve three or more actors, such as transitivity or group formation. In (Leenders 1995b), a method is proposed to estimate a transitivity model on the basis of the reciprocity model, but the statistical treatment is of an approximate nature because calculations are still made as if the dyads change independently. (Wasserman 1977, 1980) also treats popularity and expansiveness models where the change rates depend on in- or out-degrees, respectively, and where the reciprocity effect is absent. This leads to adjacency matrices where columns or rows, respectively, are independent. Wasserman presents estimators for the parameters, but these can be derived only under the usually rather unrealistic assumption of a stationary process.

5 Formulation of dyadic models as actor-oriented models

What is the relation between the reciprocity model and the actor-oriented model for network change? Both are continuous-time Markov chain models, in which at a given time point t at most one element x_{ij} of the adjacency matrix may change. This implies that these models can be compared on the basis of the change rate functions $q_{ij}(x)$ defined in (9).

For the reciprocity model, the change rates are given above in (15), (16). For the actor-oriented model (option 1), the change rates are given in (10). This section elaborates the correspondence between these two models. (Note that λ_0 and λ_1 occurring in the reciprocity model should be distinguished from $\lambda_i(x, \theta)$ occurring in the actor-oriented model.)

5.1 Independent relations model

In the simplest model, all $n(n-1)$ relations x_{ij} change independently. In other words, the reciprocity effect is absent. This corresponds to $\mu_0 = \mu_1 = 0$ in (15), (16). To find an actor-oriented representation, define

$$f_i(\theta, x) = \theta_2\, x_{i+}\ , \tag{17}$$

where replacing an index by $+$ denotes summation over this index, and define $g_i \equiv 0$. Then it follows from (8) that adding a new outgoing relation adds θ_2 while withdrawing an existing relation subtracts θ_2 from the objective function, i.e.,

$$r(\theta, i, j, x) = f_i(\theta, x(i \rightsquigarrow j)) - f_i(\theta, x) = \begin{cases} \theta_2 & (x_{ij} = 0)\ , \\ -\theta_2 & (x_{ij} = 1)\ . \end{cases}$$

With (7), this yields the probabilities

$$p_{ij}(\theta, x) = \frac{(1 - x_{ij}) \exp(\theta_2)\ +\ x_{ij} \exp(-\theta_2)}{(n - x_{i+} - 1) \exp(\theta_2)\ +\ x_{i+} \exp(-\theta_2)}\ .$$

To obtain the reciprocity model (15), (16) with $\mu_0 = \mu_1 = 0$, we can take

$$\lambda_i(\theta, x) = \theta_1 \{(n - x_{i+} - 1)\exp(\theta_2) \; + \; x_{i+}\exp(-\theta_2)\} / (n-1) \; , \qquad (18)$$

and the parameters correspond according to

$$\begin{aligned} \lambda_0 &= (\theta_1/(n-1))\exp(\theta_2) \; , \\ \lambda_1 &= (\theta_1/(n-1))\exp(-\theta_2) \; . \end{aligned} \qquad (19)$$

This illustrates that under the independent relations model with, e.g., $\theta_2 > 0$, the actors with a smaller out-degree change their relations at a faster rate than the actors with a larger out-degree.

To obtain a model that includes as special cases the independent relations model as well as the actor-oriented model with a constant rate function, the objective function (17) is used while expression (18) for the change rate is replaced by

$$\lambda_i(\theta, x) = \theta_1 \{(n - x_{i+} - 1)\exp(-\theta_3) \; + \; x_{i+}\exp(\theta_3)\} / (n-1) \; . \qquad (20)$$

A constant rate function corresponds to $\theta_3 = 0$; the independent relations model is obtained for $\theta_3 = -\theta_2$.

Analytical properties

For the independent relations model, more properties can be calculated analytically than for the general reciprocity model. This provides opportunities for checking the results of simulation-based calculations.

In the independent relations model, we have $n(n-1)$ dichotomous variables X_{ij} that are independently carrying out "on – off" processes, with a rate λ_0 for going from 0 to 1, and a rate λ_1 for going from 1 to 0. Consider one such dichotomous variable $X_0(t)$, and denote by $\xi_x(t) = \mathcal{E}\{X_0(t) \mid X_0(0) = x\}$ its expectation, conditional on $X_0(0) = x$, for $x = 0, 1$. Then, by conditioning on $X_0(t)$, it can be derived that for small dt we have the difference equation

$$\xi_x(t + dt) - \xi_x(t) \approx (1 - \xi_x(t))(\lambda_0 \, dt) \; - \; \xi_x(t)(\lambda_1 \, dt) \, .$$

This leads to the differential equation

$$\xi_x'(t) = \lambda_0 - (\lambda_0 + \lambda_1)\xi_x(t)$$

with the solution

$$\xi_x(t) = \frac{\lambda_0}{\lambda_+} \; - \; \frac{1}{\lambda_+}\exp(-\lambda_+(t + c)) \; ,$$

where $\lambda_+ = \lambda_0 + \lambda_1$, and c depends on the initial condition. With the initial conditions $\xi_x(0) = x$ we obtain the solutions

$$\begin{aligned} \xi_0(t) &= \frac{\lambda_0}{\lambda_+}\{1 - \exp(-\lambda_+ t)\} \; , \\ \xi_1(t) &= \frac{1}{\lambda_+}\{\lambda_0 + \lambda_1\exp(-\lambda_+ t)\} \; . \end{aligned}$$

Note that this implies $0 < \xi_0(t) < \xi_1(t) < 1$.

Now consider observations on two stochastic networks $(X_{ij}(t_1))$ and $(X_{ij}(t_2))$ where the change between times t_1 and t_2 is governed by the independent relations model. Denote $T = t_2 - t_1$. Sufficient statistics are the four change counts

$$N_{hk} = \sharp\{(i,j) \mid X_{ij}(t_1) = h,\, X_{ij}(t_2) = k\}$$

for $h, k = 0, 1$. The totals $N_{0+} = N_{00} + N_{01}$ and $N_{1+} = N_{10} + N_{11}$ are the numbers of absent and present relations at time t_1. These are treated as given numbers.

It follows from the independence of the relations together with the results obtained above about the distribution of $X_0(t)$, that N_{01} and N_{11} are independent binomially distributed random variables, $N_{01} \sim \mathrm{B}(N_{0+}, \xi_0(T))$ and $N_{11} \sim \mathrm{B}(N_{1+}, \xi_1(T))$.

Estimation

The estimation of the independent relations model is elaborated here only for the case $M = 2$. Define the relative frequencies

$$\hat{p}_{01} = \frac{N_{01}}{N_{0+}} \ , \quad \hat{p}_{11} = \frac{N_{11}}{N_{1+}} \ .$$

The independent binomial distributions of N_{01} and N_{11} imply that the ML estimators are the values of λ_0 and λ_1 for which

$$\hat{p}_{01} = \xi_0(T) \ , \quad \hat{p}_{11} = \xi_1(T) \ .$$

Some algebra shows that the estimates are given by

$$\begin{aligned} \hat{\lambda}_0 &= -\frac{\hat{p}_{01}}{T(1 + \hat{p}_{01} - \hat{p}_{11})} \log(\hat{p}_{11} - \hat{p}_{01}) \ , \\ \hat{\lambda}_1 &= -\frac{1 - \hat{p}_{11}}{T(1 + \hat{p}_{01} - \hat{p}_{11})} \log(\hat{p}_{11} - \hat{p}_{01}) \ . \end{aligned}$$

These equations are valid only if $\hat{p}_{11} > \hat{p}_{01}$. If this condition is not satisfied, the basic consequence of the independent relations model, $\xi_0(T) < \xi_1(T)$, is not reflected by the data. This may suggest either that the model does not fit, or (if $\hat{p}_{11} - \hat{p}_{01}$ is negative but small) that the duration T of the period between the observations is too long for drawing reliable conclusions about the change rates.

For the actor-oriented approach, Section 3 proposes the moment estimator based on the statistics C and the total number of relations at time t_2, $S_1 = X_{++}(t_2)$ (cf. (12)). The ML estimator is the moment estimator for the statistics N_{01} and N_{11}. Since $C = N_{01} + N_{10}$ and $S_1 = N_{01} + N_{1+} - N_{10}$, while N_{1+} is considered a fixed number because it is a function of $x(t_1)$, it can be concluded that N_{01} and N_{11} are linear functions of C and S_1. Therefore the moment estimator obtained on the basis of the statistics C and S_1 also is identical to the maximum likelihood estimator.

If $\hat{p}_{11} \leq \hat{p}_{01}$, the ML estimator does not exist because the supremum of the likelihood function, although finite, is not assumed for a finite value of the parameters λ_0, λ_1. Under this condition, the moment estimators also do not exist.

For the actor-oriented model with constant rate function, the recipe of Section 3 can again be used, with the same statistics.

For the estimation of parameter θ_3 in the actor-oriented model specified by (17) and (20), it is necessary to use information about the relation between out-degrees and number of changes. Moment estimation for this model when $M = 2$ can be based on C, S_1, and C_{out} defined by

$$C_{\text{out}} = \sum_{i,j=1}^{n} X_{i+}(t_1)\, |X_{ij}(t_2) - X_{ij}(t_1)| \ . \tag{21}$$

Example: the independent relations model. For the EIES data introduced above, we have $\hat{p}_{01} = 0.307$ and $\hat{p}_{11} = 0.986$, which yields $\hat{\lambda}_0 = 0.370$, $\hat{\lambda}_1 = 0.0165$, corresponding to actor-oriented parameters $\hat{\theta}_1 = 2.418$, $\hat{\theta}_2 = 1.557$. Estimation by the Robbins-Monro algorithm (with 500 simulation steps) yielded estimates $\hat{\theta}_1 = 2.406$ (s.e. 0.22), $\hat{\theta}_2 = 1.557$ (s.e. 0.24). These estimates are, of course, stochastic. In view of this, the correspondence may be called excellent.

For the actor-oriented model specified by (17) and (20), the estimates were $\hat{\theta}_1 = 4.77$ (s.e. 0.40), $\hat{\theta}_2 = 1.69$ (s.e. 0.24), $\hat{\theta}_3 = -0.56$ (s.e. 0.34). It may be tentatively concluded that θ_3 is less strongly negative than $-\theta_2$, so researchers with a higher out-degree are less active in changing their relations, but this effect is not as strong as is implied by the independent relations model.

5.2 Reciprocity model

Now consider the model defined by (15), (16) with arbitrary values of μ_0 and μ_1 (subject only to the restriction that all change rates are positive). Consider the actor-oriented model with objective function

$$f_i(\theta, x) = \theta_2\, x_{i+} \, + \, \theta_3 \sum_j x_{ij} x_{ji} \ , \tag{22}$$

still with $g_i \equiv 0$. For this model, adding a relation yields an increase of the objective function equal to

$$\theta_2 \, + \, \theta_3 x_{ji} \ ,$$

while withdrawing a relation decreases the objective function by the same amount. The probability distribution for changes of x to $x(i \rightsquigarrow j)$ therefore is given by

$$p_{ij}(\theta,x) = \frac{1}{n(x,\theta)} \times \begin{cases} \exp(\theta_2) & \text{for } x_{ij} = x_{ji} = 0 \ , \\ \exp(\theta_2 + \theta_3) & \text{for } x_{ij} = 0,\, x_{ji} = 1 \ , \\ \exp(-\theta_2) & \text{for } x_{ij} = 1,\, x_{ji} = 0 \ , \\ \exp(-\theta_2 - \theta_3) & \text{for } x_{ij} = x_{ji} = 1 \ , \end{cases} \tag{23}$$

where

$$\begin{aligned} n(x,\theta) = \sum_{\substack{j=1 \\ j\neq i}}^{n} \{ & (1-x_{ij})(1-x_{ji}) \exp(\theta_2) \ + \ (1-x_{ij})x_{ji} \exp(\theta_2+\theta_3) \ + \\ & x_{ij}(1-x_{ji}) \exp(-\theta_2) \ + \ x_{ij}x_{ji} \exp(-\theta_2-\theta_3) \} \ . \end{aligned} \tag{24}$$

When we also define

$$\lambda_i(x,\theta) = \theta_1 \, \frac{n(x,\theta)}{n-1} \ , \tag{25}$$

it can be verified that the resulting actor-oriented model is identical to the reciprocity model with

$$\begin{aligned} \theta_1 &= (n-1)\sqrt{\lambda_0 \lambda_1} \ , \\ \theta_2 &= \tfrac{1}{2} \log \left(\frac{\lambda_0}{\lambda_1} \right) \ , \\ \theta_3 &= \log \left(\frac{\lambda_0 + \mu_0}{\lambda_0} \right) \ , \end{aligned} \tag{26}$$

and with the proportionality condition

$$\frac{\lambda_0 + \mu_0}{\lambda_0} = \frac{\lambda_1}{\lambda_1 + \mu_1} \ . \tag{27}$$

If the proportionality condition (27) is not satisfied, it is still possible to represent the reciprocity model as an actor-oriented model, but then it is necessary to include in the model an instantaneous gratification function g_i. Define

$$g_i(\theta,x,j) = -\theta_4 \, x_{ij} x_{ji} \ , \tag{28}$$

with the interpretation that breaking off a reciprocated relation leads to a cost (negative gratification, loss, or instantaneous pain) of θ_4. In this case, define

$$\theta_4 = \log \left(\frac{\lambda_0}{\lambda_0 + \mu_0} \right) - \log \left(\frac{\lambda_1 + \mu_1}{\lambda_1} \right) \ , \tag{29}$$

and replace (24) by

$$\begin{aligned} n(x,\theta) = \sum_{\substack{j=1 \\ j\neq i}}^{n} \{ & (1-x_{ij})(1-x_{ji}) \exp(\theta_2) \ + \ (1-x_{ij})x_{ji} \exp(\theta_2+\theta_3) \ + \\ & x_{ij}(1-x_{ji}) \exp(-\theta_2) \ + \ x_{ij}x_{ji} \exp(-\theta_2-\theta_3-\theta_4) \} \ . \end{aligned} \tag{30}$$

It can be checked that this specification yields an actor-oriented model formulation of the reciprocity model, valid without the condition (27).

Condition (27) can be regarded as a "conservation of utility" condition. If it is not satisfied, e.g., $(\lambda_0+\mu_0)/\lambda_0 > \lambda_1/(\lambda_1+\mu_1)$, then unilaterally breaking off a reciprocated relation entails a loss greater than the reward associated with starting a reciprocated relation.

This model also can be embedded in a model that also contains as a special case the model with constant change rates. This is achieved by replacing (25) by

$$\lambda_i(x,\theta) = \frac{\theta_1}{n-1} \sum_{\substack{j=1 \\ j\neq i}}^{n} \{(1-x_{ij})(1-x_{ji})\exp(\theta_5) + (1-x_{ij})x_{ji}\exp(\theta_5+\theta_6) \\ + x_{ij}(1-x_{ji})\exp(-\theta_5) + x_{ij}x_{ji}\exp(-\theta_5-\theta_6-\theta_7)\} \; . \tag{31}$$

Estimation

Again we consider only the case $M = 2$. For the reciprocity model, the ML estimator is given in (Wasserman 1977, 1979, 1980) with a correction by (Leenders 1995a). In the actor-oriented framework, the proposed method of Section 3 leads, for the model with objective function (22) and $g_i \equiv 0$, to the moment estimator on the basis of C and the two statistics

$$S_1 = X_{++}(t_2) \; , \tag{32}$$
$$S_2 = \sum_{i,j=1}^{n} X_{ij}(t_2)X_{ji}(t_2) \; .$$

When also the gratification function (28) is included in the model, a fourth statistic is necessary, relevant for the loss associated to breaking off reciprocated relations. The proposed statistic is

$$S_3 = \sum_{i,j=1}^{n} X_{ij}(t_1)X_{ji}(t_1)(1-X_{ij}(t_2)) \; . \tag{33}$$

Unlike in the independent relations model, these moment estimators do not coincide with the ML estimator. The reason is that the reciprocity model is a "curved exponential family", for which a sufficient statistic with as many dimensions (3 or 4) as the parameter vector does not exist.

Example: the reciprocity model. The ML estimates for the parameters of the reciprocity model for the EIES data, calculated by the method of Leenders (1995a), are $\hat\lambda_0 = 0.272$, $\hat\lambda_1 = 0.0907$, $\hat\mu_0 = 0.478, \hat\mu_1 = -0.0816$. This corresponds to actor-oriented parameters $\hat\theta_1 = 5.02$, $\hat\theta_2 = 0.549$, $\hat\theta_3 = 1.01$, $\hat\theta_4 = 1.29$.

The Robbins-Monro method yielded the estimates $\hat{\theta}_1 = 4.53$ (s.e. 1.02), $\hat{\theta}_2 = 0.550$ (s.e. 0.22), $\hat{\theta}_3 = 1.03$ (s.e. 0.20), $\hat{\theta}_4 = 0.74$ (s.e. 1.3). The correspondence is not exact, which is understable given the fact that the estimators are not equivalent.

Example: a more general actor-oriented model. A more general actor-oriented model was also estimated. Recall that the actor-bound variable w_j was equal to 0 or 1, respectively, for researchers with low or high citation rates. Some explorations with various effects led to the following model specification and estimates. The change rate is modeled by

$$\lambda_i(\theta, x) = \theta_1 \frac{\{(n - x_{i+} - 1)\exp(-\theta_2) \, + \, x_{i+}\exp(\theta_2)\}}{n-1} \; ,$$

the utility function by

$$\begin{aligned} f_i(x,\theta) \, = \, & \theta_3 \sum_j x_{ij} \, + \, \theta_4 \sum_j x_{ij} x_{ji} \, + \, \theta_5 \sum_j x_{ij} w_j \\ & + \theta_6 \sum_j x_{ij} \, | \, w_i - w_j \, | + \theta_7 \sum_{j,h} x_{ij} x_{ih} x_{jh} \; , \end{aligned}$$

and there is no gratification function. The estimated effects are as follows (with standard errors in parentheses).

θ_1	constant factor in rate	4.62	(0.99)
θ_2	effect of out-degree on rate	-0.69	(0.29)
θ_3	number of relations	-1.31	(0.58)
θ_4	reciprocity	0.92	(0.29)
θ_5	popularity of others with high citation rates	-0.51	(0.21)
θ_6	similarity to others with respect to citation rates	0.20	(0.24)
θ_7	transitivity	0.255	(0.075)

The interpretation is that researchers with a high out-degree tend to make less changes in their relationships; there are clear reciprocity and transitivity effects; others with high citation rates tend to be chosen less; and the tendency to choose others with the same citation rate (both high, or both low) is not significant.

References

Bowman, K.O. & Shenton, L.R. (1985). Method of moments. p. 467 – 473 in Kotz, S., Johnson, N.L. & Read, C.B. (eds.) *Encyclopedia of Statistical Sciences, vol. 5.* New York: Wiley.

Doreian, P. & Stokman, F.N. (eds.) (1997). *Evolution of Social Networks.* Amsterdam etc.: Gordon and Breach.

Frank, O. (1991). Statistical analysis of change in networks. *Statistica Neerlandica, 45*, 283 – 293.

Gouriéroux, C. & Montfort, A. (1996). *Simulation-based econometric methods.* Oxford: Oxford University Press.

Leenders, R.Th.A.J. (1995a). Models for network dynamics: a Markovian framework. *Journal of Mathematical Sociology, 20*, 1 – 21.

Leenders, R.Th.A.J. (1995b). *Structure and influence. Statistical models for the dynamics of actor attributes, network structure and their interdependence.* Amsterdam: Thesis Publishers.

Maddala, G.S. (1983). *Limited-dependent and qualitative variables in econometrics.* Cambridge: Cambridge University Press.

Mayer, T.F. (1984), Parties and networks: stochastic models for relationship networks. *Journal of Mathematical Sociology, 10*, 51 – 103.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica, 57*, 995 – 1026.

Newcomb, T. (1961). *The acquaintance process.* New York: Holt, Rinehart and Winston.

Norris, J.R. (1997). *Markov Chains.* Cambridge: Cambridge University Press.

Pakes, A. & Pollard, D. (1989). The asymptotic distribution of simulation experiments, *Econometrica, 57*, 1027 – 1057.

Robbins, H. & Monro, S. (1951). A stochastic approximation method. *Annals of Mathematical Statistics, 22*, 400 – 407.

Ruppert, D. (1991). Stochastic approximation. In Gosh, B.K. & Sen, P.K., *Handbook of Sequential Analysis.* New York: Marcel Dekker.

Snijders, T.A.B. (1995), Methods for longitudinal social network data. Pp. 211 – 227 in Tiit, E.-M., Kollo, T. & Niemi, H. (eds.), *New Trends in Probability and Statistics, Vol. 3: Multivariate Statistics and Matrices in Statistics.* Vilnius, Lithiania: TEV and Utrecht, The Netherlands: VSP.

Snijders, T.A.B. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology, 21*, 149 – 172.

Van de Bunt, G.G., Van Duijn, M.A.J. & Snijders, T.A.B. (1995). Friendship networks and rational choice. In *Proceedings International Conference on Social Networks, Volume 1: Methodology.* London: CASSM, University of Greenwich.

Wasserman, S. (1977). *Stochastic models for directed graphs.* Ph.D. dissertation, University of Harvard, Dept. of Statistics.

Wasserman, S. (1979). A stochastic model for directed graphs with transition rates determined by reciprocity. In Schuessler (ed.), *Sociological Methodology 1980.* San Francisco: Jossey-Bass.

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

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications.* New York and Cambridge: Cambridge University Press.

Zeggelink, E.P.H. (1994). Dynamics of structure: an individual oriented approach. *Social Networks, 16*, 295 – 333.