Statistical models for dynamics of social networks: inference and applications

Snijders, Tom A.B.[1]
University of Oxford, Department of Statistics
Nuffield College
New Road
Oxford OX1 1NF, United Kingdom
tom.snijders@ox.ac.uk

University of Groningen, Department of Sociology
Grote Rozenstraat 31
9712 GC Groningen, The Netherlands

Social Networks

Social networks are relational structures between social actors. The basic mathematical representation for a social network is a graph or directed graph; in this paper only directed relations will be studied. The graph represents a dyadic relation of some kind, e.g., friendship between individuals or buyer/supplier relations between companies, and the word 'tie' will be used to refer generically to the existence of a link according to this relation. For statistical modeling a convenient representation is the adjacency matrix $Y = (y_{ij})$ for $i, j \in \mathcal{N}$ where the elements of the node set $\mathcal{N}$ are the social actors (which could be individuals, companies, countries, etc.) and $y_{ij} = 1$ if there is a tie from i to j, also denoted $i \to j$, while $y_{ij} = 0$ is there is no such tie. The variables y_{ij} are called *tie variables.* Reflexive ties are excluded, so the diagonal elements y_{ii} may be regarded as structural zeros. We shall consider applications where the node set $\mathcal{N}$ is fixed by the research design, and the tie variables are random. The set of all directed graphs, or adjacency matrices, on a given node set $\mathcal{N}$ will be denoted by $\mathcal{Y}^{(\mathcal{N})}$.

More specifically, we consider situations where the data is longitudinal, consisting of repeated measures $Y(t_1), \ldots, Y(t_M)$ of a network on a constant node set, with $M \geq 2$ observations. In addition, there may be covariates which could be monadic (defined on $\mathcal{N}$) or dyadic (defined on $\{(i, j) \in \mathcal{N}^2 \mid i \neq j\}$). Dependence of functions on covariates will be often left implicit in the notation. As an example, one may think of a friendship network in a school class of fixed composition, observed at several points in time, spaced by a few months; data collection taking place by questionnaires filled in by the students.

The basic question then is to find rules modeling the dynamics of the network as it moves from $y(t_1)$ to $y(t_2)$, etc., and eventually to $y(t_M)$.

Actor-oriented models

Plausible models for network dynamics should represent typical network dependencies such as reciprocation of ties, the limitation of the number of ties that an actor could possibly entertain, transitivity ('friends of friends tend to become, or stay, friends'), etc. This paper is about a class of such models proposed by Snijders (2001), called *stochastic actor-oriented models.* This is a specification of the very general framework proposed by Holland and Leinhardt (1977), of which earlier models, limited to quite restricted network dependencies, were formulated by Wasserman (1980).

[1]Support is gratefully acknowledged from NIH grant 1R01HD052887-01A2, *Adolescent Peer Social Network Dynamics and Problem Behavior.* I thank Ruth Ripley and Krists Boitmanis for creating the RSiena package.

Basic properties of this model, conforming to the proposals of Holland and Leinhardt (1977), are the following.

1. The model is a continuous-time stochastic process $Y(t)$, $t_1 \le t \le t_M$, on the space of directed graphs, observed only at the discrete time points $t_1, t_2, \ldots, t_M$.
2. $Y(t)$ is a Markov process with stationary transition distribution (without any requirements about the stationarity of the marginal distribution).
3. At each time point t in the continuous-time model, at most one tie variable Y_{ij} can change. (This rules out direct modeling of coordination or strategic considerations.)

The assumption of change taking place continuously between observation moments has a clear intuitive validity. The combination of a continuous-time model with allowing changes only of single tie variables at any given moment provides a parsimonious model that represents dependencies as following from a feedback process, where existing network patterns and tie changes may lead to further tie changes.

The actor-oriented model (Snijders, 2001) takes the point of view that social actors control their outgoing ties, and network changes are regarded as being the consequence of choices made by the actors. The model is decomposed into two sub-models.

4. *Timing.* For each actor $i \in \mathcal{N}$ there is a positive *rate function* $\lambda_i(\alpha, \rho, y)$. The waiting time until the next opportunity for change is exponentially distributed with parameter $\sum_i \lambda_i(\alpha, \rho, y)$. Given that an opportunity for change occurs, this is an opportunity for actor i with probability
$$\frac{\lambda_i(\alpha, \rho, y)}{\sum_j \lambda_j(\alpha, \rho, y)} \ .$$
5. *Choice.* For each actor $i \in \mathcal{N}$ there is an *objective function* $f_i(\beta, y', y)$, interpreted as a measure for the tendency of actor i to move to y, given that a move takes place from the current state y'. The set of networks that can be obtained according to rule (3) as the results of an opportunity for change for actor i starting from network state y', under the proviso that it is permitted not to make a change, is denoted by
$$\begin{aligned} \mathcal{C}_i(y') \subset \big\{y'\big\} \cup \big\{ y \in \mathcal{Y}^{(\mathcal{N})} \ \mid \ & \text{there is exactly one } j \neq i \text{ with } y_{ij} = 1 - y'_{ij}, \\ & \text{and } y_{hk} = y'_{hk} \text{ for all other } (h,k) \big\}. \end{aligned}$$
This formula uses the subset ($\subset$) rather than equality sign to obtain greater generality; in some cases, some tie variables y_{ij} may be defined to be unchangeable, and outcomes where these variables become different then are left out of the set $\mathcal{C}_i(y')$. If actor i has an opportunity for change, the probability of changing to the new state $y \in \mathcal{C}_i(y')$ is defined by
$$(1) \qquad \frac{\exp\big(f_i(\beta, y', y)\big)}{\sum_{y'' \in \mathcal{C}_i(y')} \exp\big(f_i(\beta, y', y'')\big)} \ .$$

These two points are not assumptions, because given conditions (1-3) they can always be satisfied for an appropriate choice of the functions λ_i and f_i. The assumptions of the actor-oriented model consist of the specification of these two functions. They are specified as generalized models, depending on statistics $r_{ik}(y)$ for the rate functions and $s_{ik}(y)$ for the objective function, as follows.

6. For $t_m \le t < t_{m+1}$, the rate functions are defined as
$$\lambda_i(\alpha, \rho, y) = \rho_m \, \exp\Big(\sum_{k=1}^{L} \alpha_k \, r_{ik}(y) \Big) \ .$$

7. The objective functions are defined as

$$f_i(\beta, y', y) = \sum_{k=1}^{K} \beta_k \, s_{ik}(y', y) \ .$$

The statistics r_{ik} and s_{ik} are called *effects.* They must be specified by the user, depending on research interests as well as theories and empirical knowledge regarding the subject matter. In this paper we restrict attention to models where $L = 0$, i.e., rates of change are constant across actors in the time periods between observation moments; and $s_{ik}(y', y) = s_{ik}(y)$, i.e., relative probabilities of change depend only on the new state and not on the old state. More general models are treated in Snijders (2001). The effects $s_{ik}(y)$ will depend in practice on the 'local network neighborhood' of actor i – this concept is left vague here but is illustrated by the following examples. All these effects are statistics reflecting the network 'around' actor i.

1. Out-degree, $s_{i1}(y) = y_{i+} = \sum_j y_{ij}$.
2. Number of reciprocated ties, $s_{i2}(y) = \sum_j y_{ij} \, y_{ji}$.
3. Number of transitive triplets, $s_{i3}(y) = \sum_{j,h} y_{ij} \, y_{jh} \, y_{ih}$.
4. Number of three-cycles, $s_{i4}(y) = \sum_{j,h} y_{ij} \, y_{jh} \, y_{hi}$.
5. If an attribute W is defined on the set of actors, with categorical values w_i: the number of ties to actors with the same attribute, $s_{i5}(y) = \sum_j y_{ij} \, I\{w_i = w_j\}$, where $I\{A\} = 1$ if A is true, and 0 otherwise.

More examples are given in Snijders (2001) and Snijders, van de Bunt and Steglich (2010).

Simulation

The actor-oriented model can be simulated in a straightforward way. The following algorithm simulates the model from t_m to t_{m+1}, with "$\leftarrow$" denoting the assignment operator.

1. Set $y \leftarrow y(t_m)$, $t \leftarrow t_m$.
2. Let Δt be the outcome of an exponentially distributed random variable with parameter $\sum_i \lambda_i(y)$. Set $t \leftarrow t + \Delta t$.
3. Let i^* be the outcome of a random variable with values i in $\mathcal{N}$ and probabilities $\lambda_i(y) / \big(\sum_j \lambda_j(y)\big)$.
4. Define $y^{(\pm ij)}$ for $i \neq j$ as the adjacency matrix with $y^{(\pm ij)}_{hk} = y_{hk}$ for $(h, k) \neq (i, j)$ and $y^{(\pm ij)}_{ij} = 1 - y_{ij}$; and define $y^{(\pm ii)} = y$.
   Let j^* be the outcome of a random variable with values j in $\mathcal{N}$ and probabilities
   $$\frac{\exp\big(f_{i^*}(\beta, y, y^{(\pm i^* j)})\big)}{\sum_{h \in \mathcal{N}} \exp\big(f_{i^*}(\beta, y, y^{(\pm i^* h)})\big)} \ .$$
   If $j^* \neq i^*$, set $y_{i^* j^*} \leftarrow 1 - y_{i^* j^*}$.
5. If $t \geq t_m$ terminate; else go to (2.).

Each choice of $\Delta t, i^*$, and j^* according to this sequence of steps is called a *micro step.*

Method of Moments Estimation

In many practical cases the changing network $x(t)$ is observed only at a limited finite number of time points $t_1, \ldots, t_M$. Such cases can be regarded as situations with a lot of missing data. The parameters to be estimated are ρ and β. Due to the Markov property, it is straightforward to simulate the model starting with a given observation. This was exploited in the following method of moments (or estimating equations) estimator proposed by Snijders (2001).

For each period (t_m, t_{m+1}), this estimator considers the probability distribution of $Y(t_{m+1})$, given that $Y(t_m) = y(t_m)$. In particular, this means that the first observation, $y(t_1)$, is conditioned upon in the estimations. The statistics used for this estimation procedure are, for parameter ρ_m, the Hamming distance

$$(2) \quad H\big(y(t_m), y(t_{m+1})\big) = \sum_{i,j \in \mathcal{N}} \mid y_{ij}(t_{m+1}) - y_{ij}(t_m) \mid ;$$

and for parameter β_k, the sums of the effects

$$(3) \quad S_k\big(y(t)\big) = \sum_{i \in \mathcal{N}} s_{ik}\big(y(t)\big) \; .$$

The moment equations are defined by

$$\begin{aligned}
&(4a) \quad E_{\rho,\beta}\Big\{H_m\big(y(t_m), Y(t_{m+1})\big) \mid Y(t_m) = y(t_m)\Big\} &&= H_m\big(y(t_m), y(t_{m+1})\big) \;, \; m = 1, \ldots, M-1; \\
&(4b) \quad \sum_{m=1}^{M-1} E_{\rho,\beta}\Big\{S_k\big(Y(t_{m+1})\big) \mid Y(t_m) = y(t_m)\Big\} &&= \sum_{m=1}^{M-1} S_k\big(y(t_{m+1})\big) \;, \; k = 1, \ldots, K.
\end{aligned}$$

This system of equations can be solved approximately, e.g., by stochastic approximation (Robbins and Monro, 1951; Pflug, 1996). The iteration step of this algorithm is as follows; the number N is the index for the iteration step. Denote the current provisional parameter estimate by $\big(\rho^{(N)}, \beta^{(N)}\big)$. Then for each $m = 1, 2, \ldots, M-1$ the network evolution is simulated using these parameters, going from $Y(t_m) = y(t_m)$ to $Y(t_{m+1})$. Here the algorithm of the preceding section is used. Denote these simulated values of $Y(t_{m+1})$ by $y^{(N)}(t_{m+1})$. Then the parameters are updated according to

$$(5) \quad \begin{pmatrix} \rho_1^{(N+1)} \\ \vdots \\ \rho_{M-1}^{(N+1)} \\ \theta^{(N+1)} \end{pmatrix} = \begin{pmatrix} \rho_1^{(N)} \\ \vdots \\ \rho_{M-1}^{(N)} \\ \theta^{(N)} \end{pmatrix} - a_N \, D^{-1} \begin{pmatrix} H\big(y(t_1), y^{(N)}(t_2)\big) - H\big(y(t_1), y(t_2)\big) \\ \vdots \\ H\big(y(t_{M-1}), y^{(N)}(t_M)\big) - H\big(y(t_{m-1}), y(t_M)\big) \\ \sum_{m=1}^{M-1} \Big\{S_k\big(y^{(N)}(t_{m+1})\big) - S_k\big(y(t_{m+1})\big)\Big\} \end{pmatrix} .$$

Here a_N is a sequence of positive constants tending to 0; D is a suitable matrix, approximating the matrix of partial derivatives of the expected values of (2) and (3) with respect to the elements of ρ and θ. Details and further specification of the algorithm are discussed in Snijders (2001, 2005). This algorithm is very robust and has been implemented in the R package *RSiena* (Ripley, Snijders and Preciado, 2011).

Likelihood-based estimation

Maximum likelihood and Bayesian estimation for this model were developed, respectively, by Snijders, Koskinen, and Schweinberger (2010) and Koskinen and Snijders (2007). Both can be based on data augmentation, augmenting the observed data by the sequences of micro steps taking $y(t_m)$ to $y(t_{m+1})$ $(m = 1, \ldots, M-1)$ according to the simulation algorithm described above. The two mentioned publications describe Metropolis Hastings schemes for sampling from the augmented data, conditional on the observed data. In Snijders, Koskinen, and Schweinberger (2010) the holding times Δt are integrated out, so that the augmented data consist only of the sequences of variables (i^*, j^*), where each (i^*, j^*) corresponds to a micro step.

Various algorithms may be used to approximate the ML estimates, using this kind of data augmentation. The mentioned paper uses the Markov Chain Stochastic Approximation (MCSA) algorithm of Gu and Kong (1998), which rests on the missing information principle (Orchard and Woodbury, 1972; Louis, 1982). The principle of this algorithm is as follows. Denote the data by $y = \big(y(t_1), \ldots, y(t_M)\big)$, and its probability function conditional on $Y(t_1) = y(t_1)$ by $p_Y(y; \rho, \theta)$. Denote the sequence of coordinates (i^*, j^*) of micro steps (including those with $i^* = j^*$) by v, which are regarded as the outcome of a random vector V. Denote the joint probability function of (Y, V) conditional on $Y(t_1) = y(t_1)$ by $p_{YV}(y, v; \rho, \theta)$. Further denote the observed data score function $\partial \log(p_Y(y; \rho, \theta))/\partial(\rho, \theta)$ by $S_Y(\rho, \theta; y)$ and the total data score function $\partial \log\big(p_{YV}(y, v; \rho, \theta)\big)/\partial(\rho, \theta)$ by $S_{YV}(\rho, \theta; y, v)$. Then (Orchard and Woodbury, 1972)

$$(6) \quad E_{\rho,\theta}\{S_{YV}(\rho, \theta; y, V) \mid Y = y\} = S_Y(\rho, \theta; y) \ .$$

Therefore, the likelihood equation can be expressed under differentiability conditions as

$$(7) \quad E_{\rho,\theta}\{S_{YV}(\rho, \theta; y, V) \mid Y = y\} = 0,$$

which equation therefore can be used to determine the ML estimate.
The MCSA algorithm is a 'simulations within simulations' algorithm: the outer loop is a stochastic approximation algorithm with steps analogous to (5), applied to the score function for the augmented data as estimated from simulated data $v^{(N)}$,

$$S_{YV}(\rho^{(N)}, \theta^{(N)}; y, v^{(N)}) \ ;$$

the inner loop is composed of Metropolis Hastings steps for simulating a random draw $v^{(N)}$ from the distribution of V given the observed data $Y = y$.

Multivariate networks

The actor-oriented model can also be applied for multivariate observations. In Snijders, Steglich and Schweinberger (2007) this was done for networks and individual variables. Here we propose a model for multivariate networks. The formulation is given here for bivariate networks, but this can be extended straightforwardly to more than two dependent networks.

Suppose that on the same set of actors $\mathcal{N}$ two dependent networks are defined, denoted by $y_{[1]}$ and $y_{[2]}$. An example is friendship and collaboration relations for the same set of individuals. The actor-oriented model now is defined similarly as above, but for the state variable $Y(t) = \big(Y_{[1]}(t), Y_{[2]}(t)\big)$. The points defining the actor-oriented model for this state variable are the following, referring to points (1-7) above, used to define the actor-oriented model for a single network.

1, 2. Identical to (1,2) as above.

3. At each time point in the continuous-time model, at most one single variable $Y_{[\ell]ij}(t)$ (i.e., a variable for one triple (ℓ, i, j)) can change.

4. Rate functions $\lambda_{[\ell]i}$ are defined separately for each dependent network ℓ. The waiting time until the next opportunity for change is exponentially distributed with parameter $\sum_\ell \sum_i \lambda_{[\ell]i}(\alpha, \rho, y)$. Given that an opportunity for change occurs, this is an opportunity for dependent network ℓ and for actor i with probability
$$\frac{\lambda_{[\ell]i}(\alpha, \rho, y)}{\sum_{l,j} \lambda_{[l]j}(\alpha, \rho, y)} \ .$$

5. Objective functions $f_{[\ell]i}$ also are defined separately for each dependent network ℓ. The set of networks $y_{[\ell]}$ that can be obtained as a result from an opportunity for change for dependent network ℓ and actor i, given the current state $y' = \big(y'_{[1]}, y'_{[2]}\big)$, is denoted $\mathcal{C}_{[\ell]i}(y')$. In such an event, the probability of changing network ℓ to the new state $y_{[\ell]} \in \mathcal{C}_{[\ell]i}(y')$ is given by
$$\text{(8)} \qquad \frac{\exp\big(f_{[\ell]i}(\beta, y', y_{[\ell]})\big)}{\sum_{y''_{[\ell]} \in \mathcal{C}_{[\ell]i}(y')} \exp\big(f_{[\ell]i}(\beta, y', y''_{[\ell]})\big)} \ .$$
In certain model specifications, the sets of permitted new values for one network may depend on the other network. For example, below an example is presented for 'like' and 'dislike' relations that are assumed to be mutually exclusive.

6. For $t_m \le t < t_{m+1}$, the rate functions are defined as
$$\lambda_{[\ell]i}(\alpha, \rho, y) = \rho_{[\ell]m} \ \exp\Big(\sum_{k=1}^{L} \alpha_{\ell k}\, r_{\ell ik}(y)\Big) \ .$$

7. The objective functions are defined as
$$f_{[\ell]i}(\beta, y', y_{[\ell]}) = \sum_{k=1}^{K} \beta_{\ell k}\, s_{\ell ik}(y', y_{[\ell]}) \ .$$
Thus, parameters for the rate and objective functions are specific to the dependent network, unless otherwise specified.

As examples, again we consider only the situation where rate functions are constant across actors within time periods for each dependent network (i.e., $L = 0$), and relative probabilities of change depend only on the new state (of all the networks). Some effects expressing dependencies between the networks are the following. Here the dependent network is the one with index $\ell = 1$.

1. Direct dependence, $s_{1ia}(y_{[1]}, y_{[2]}) = \sum_j y_{[2]ij}\, y_{[1]ij}$.
2. Cross-network reciprocity, $s_{1ib}(y_{[1]}, y_{[2]}) = \sum_j y_{[2]ji}\, y_{[1]ij}$.
3. Cross-network closure, $s_{1ic}(y_{[1]}, y_{[2]}) = \sum_{j,h} y_{[2]ih}\, y_{[2]hj}\, y_{[1]ij}$.
4. Mixed-network closure, $s_{1id}(y_{[1]}, y_{[2]}) = \sum_{j,h} y_{[2]ih}\, y_{[1]hj}\, y_{[1]ij}$.

These are examples of algebraic network effects (cf. Pattison, 1993), being based on composition of relations. For example, if network $y_{[1]}$ is friendship and $y_{[2]}$ is collaboration, then direct dependence means that collaborators tend to become, or remain, friends; cross-network dependence means that collaboration tends to be reciprocated by friendship; cross-network closure means that collaborators of collaborators tend to become, or remain, friends; and mixed-network closure means that friends of collaborators tend to become, or remain, friends.

For the method of moments estimation, the statistics defined by (3) cannot be directly used for the between-network dependencies. For example, direct dependence of $Y_{[1]}$ on $Y_{[2]}$ then would lead to the same statistic as direct dependence of $Y_{[2]}$ on $Y_{[1]}$, and thereby to collinearity in the estimation algorithm. The same problem was addressed in Snijders, Steglich, and Schweinberger (2007) for estimating parameters for models for the co-evolution of a network and an individual variable. A similar solution can be used here. For estimating a between-network effect, e.g., of $Y_{[2]}$ on $Y_{[1]}$, defined by $s_{1ik}(y_{[1]}, y_{[2]})$, the proposed statistic is

$$(9) \quad S_{1k}\big(y_{[1]}(t_{m+1}), y_{[2]}(t_m)\big) = \sum_{i \in \mathcal{N}} s_{1ik}\big(y_{[1]}(t_{m+1}), y_{[2]}(t_m)\big)$$

and the corresponding components of the moment equations (4b) are replaced by

$$(10) \quad \sum_{m=1}^{M-1} E_{\rho,\beta}\Big\{S_{1k}\big(Y_{[1]}(t_{m+1}), y_{[2]}(t_m)\big) \mid Y(t_m) = y(t_m)\Big\} = \sum_{m=1}^{M-1} S_{1k}\big(y_{[1]}(t_{m+1}), y_{[2]}(t_m)\big) \ .$$

Note the use of the time ordering to represent causality: $Y_{[2]}$ here is the explanatory and $Y_{[1]}$ the dependent variable.

Example: like and dislike networks

As an example, we consider the Sampson (1968) monastery data. This is a small longitudinal network data set distributed with the Pajek (de Nooy, Mrvar, and Batagelj, 2004) program, available from `http://vlado.fmf.uni-lj.si/pub/networks/data/esna/sampson.htm`. The data used here are like and dislike relations between 18 novice monks in a cloister, during a period of upheaval in the Roman Catholic Church which also was reflected in this monastery. We use the 'three best liked' and 'three most disliked' nominations at times 2, 3, and 4 in the file Sampson.net obtained from this website. The request to mention three was followed by the monks only approximately. The number of 'like' nominations per individual always is three or four. Two monks gave no dislike nominations at any of the three time points, and one monk mentioned no dislikes at time 4. The number of 'dislike' nominations by the others ranges from two to four.

Like and dislike are mutually exclusive relations. This is respected in the actor-oriented model. Indicate the 'like' network by $\ell = 1$ and 'dislike' by $\ell = 2$. Accordingly, the set of permitted 'like' networks after a micro step in the first network by actor i is

$$\mathcal{C}_{[1]i}\big(y'_{[1]}, y'_{[2]}\big) \;=\; \big\{y'_{[1]}\big\} \;\cup\; \Big\{y_{[1]} \in \mathcal{Y}^{(\mathcal{N})} \mid \text{ there is exactly one } j \neq i \text{ with } y_{[1]ij} = 1 - y'_{[1]1ij}, \\ \text{satisfying } y'_{[2]ij} = 0, \text{ and } y_{[1]hk} = y'_{[1]hk} \text{ for all other } (h,k)\Big\},$$

and similarly for the changes in the 'dislike' networks. In other words, for both ties of tie, a change in the tie $i \to j$ is permitted only if the corresponding tie in the other network is not present.

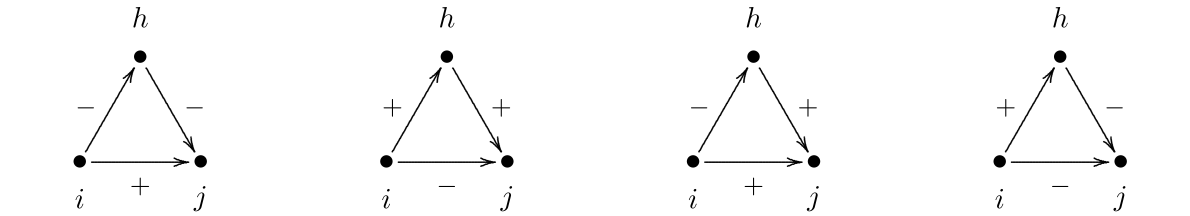
For these longitudinal network data, the following model specification was initially employed.

1. For both networks, the effects were included denoted above as out-degree, reciprocated ties, transitive triplets, and three-cycles.

2. As between-network effects were included: cross-network reciprocity, cross-network closure, and mixed-network closure. Expectations for the parameters for the triadic between-network effects can be based on balance theory, cf. Heider (1946), Davis (1967), and Doreian and Mrvar (1996). The four closure effects are represented in Figure 1. Briefly, the predictions of balance theory are that the triadic structures with a single 'dislike' and two 'like' ties are unstable and therefore will have negative parameters in the actor-oriented model, while the triadic structures with two

'dislike' ties and a single 'like' tie are stable and will have positive parameters. Cross-network closure means for 'like' that if i dislikes h who dislikes j, then i will tend to like j, for which a positive parameter is expected; and it means for 'dislike' that if i likes h who likes j, then i will tend to dislike j – here a negative parameter is expected. Mixed-network closure means for 'like' that if i dislikes h who likes j, then i will tend to like j, for which a negative parameter is expected; and for 'dislike' that if i likes h who dislikes j, then i will tend to dislike j, for which a positive parameter is expected.

Figure 1. Between-network closure effects for like (+) and dislike (–).



Cross-network closure for 'like'. | Cross-network closure for 'dislike'. | Mixed-network closure for 'like'. | Mixed-network closure for 'dislike'.

3. To represent the low value of the dispersion of observed out-degrees, which was due to the request to mention the three best liked and three most disliked individuals, the squared out-degree effect was included, defined by $s_{\ell i6}(y) = \big(\sum_j y_{[\ell]ij}\big)^2$. Note that a component in the objective function given by the quadratic function of the out-degree

$$\beta_{\ell 1} \sum_j y_{[\ell]ij} + \beta_{\ell 6} \Big(\sum_j y_{[\ell]ij}\Big)^2 ,$$

where $\beta_{\ell 1}$ and $\beta_{\ell 6}$ both are large in absolute value, with $\beta_{\ell 6} < 0 < \beta_{\ell 1}$, will influence the out-degrees strongly toward values of about $\beta_{\ell 1}/(2\,|\beta_{\ell 6}|)$.

4. A particular feature of the 'dislike' network is that for some monks the out-degrees are zero at all three time points, and the only other out-degree of zero occurs for the last time point. The other out-degrees for 'dislike' are equal to three or four, except for one value of two. Thus, the zero values seem special, and individuals who once give zero 'dislike' nominations continue doing so at later points in time. To represent this, an effect is added to the objective function for the 'dislike' network defined by

$$\beta_{2,7}\, I\Big\{ \sum_j y_{[2]ij} = 0 \Big\}$$

with $\beta_{2,7}$ fixed at the value +100. This represents that a zero out-degree of a given actor in the 'dislike' network is a sink, and this property of the network state, once reached, will continue with a very high probability. This perhaps amounts to overfitting, but it is a direct way of representing this particular pattern in the data.

Parameters were estimated by the method of moments. It turned out that the mixed reciprocity effects, the three-cycle effects, and the transitive triplets effect for 'dislike' were quite weak and non-significant. Therefore we present results in which these four effects are dropped from the model.

	parameter estimate	standard error
Like		
Out-degree	29.54	(11.62)
Squared out-degree	–4.83	(1.68)
Reciprocity	1.40	(0.30)
Transitive triplets	0.51	(0.13)
Cross-network closure	0.28	(0.29)
Mixed-network closure	–0.60	(0.47)
Dislike		
Out-degree	7.21	(4.79)
Squared out-degree	–1.42	(0.80)
Reciprocity	0.74	(0.33)
Out-degree 0	–100.00	(—)
Cross-network closure	–0.00	(0.34)
Mixed-network closure	0.68	(0.22)

Table 1. Method of moments estimates for dynamics of bivariate network (likes and dislikes) in Sampson's (1968) data set of 18 monks.

The conclusions from this table are the following. The effects of functions of out-degrees are in line with the characteristics of the out-degrees in the data set: large absolute values, negative parameters for the squared out-degrees, and the ratio of out-degree parameter to absolute parameter for squared out-degree close to the expected value of 6. There is reciprocation of likes as well as dislikes, but stronger for likes. There is evidence for transitivity of liking. The only closure effect between networks for which there is clear evidence is mixed-network closure for dislike: the monks tended to agree in their dislikes with those whom they liked.

REFERENCES

Davis, J.A. (1976). Clustering and structural balance in graphs. *Human Relations*, 20: 181–187.

de Nooy, W., Mrvar, A., and V. Batagelj. (2004). *Exploratory Social Network Analysis with Pajek.* Cambridge: Cambridge University Press.

Doreian, P., and A. Mrvar (1996). A partitioning approach to structural balance. *Social Networks*, 18: 149–168.

Gu, M.G., and F.H. Kong. (1998). A stochastic approximation algorithm with Markov chain Monte-Carlo Method for incomplete data estimation problems. *Proceedings of the National Academy of Sciences, U.S.A.*, 95: 7270–7274.

Heider, F. (1948). Attitudes and cognitive organization. *Journal of Psychology*, 21: 107–112.

Holland, P., and S. Leinhardt. (1977). A dynamic model for social networks. *Journal of Mathematical Sociology*, 5: 5–20.

Koskinen, J.H., and T.A.B. Snijders. (2007). Bayesian inference for dynamic network data. *Journal of Statistical Planning and Inference*, 13: 3930–3938.

Louis, T.A. (1982). Finding observed information when using the EM algorithm. *Journal of the Royal Statistical Society* Series **B** 44: 226–233.

Orchard, T. and M.A. Woodbury. (1972). A missing information principle: theory and applications. *Proceedings 6th Berkeley Symposium on Mathematical Statistic and Probability*, 1: 697–715. Berkeley: University of California Press.

Pattison, P. (1993). *Algebraic Models for Social Networks.*

Pflug, G.Ch. (1996). *Optimization of Stochastic Models.* Boston: Kluwer.

Ripley, R.M., Snijders, T.A.B., and P. Preciado Lopez. (2011). *Manual for SIENA version 4.0.* Oxford: University of Oxford, Department of Statistics; Nuffield College. http://www.stats.ox.ac.uk/siena/

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

Sampson, S.F. (1968). *A Novitiate in a Period of Change. An Experimental and Case Study of Social Relationships.* Cornell University: PhD thesis.

Snijders, T.A.B. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology–2001*, 31: 361–395.

Snijders, T.A.B. (2005) Models for longitudinal network data. In *Models and Methods in Social Network Analysis*, edited by Carrington, P.J., Scott, J., and S. Wasserman, pp. 215–247. New York: Cambridge University Press.

Snijders, T.A.B., J.H. Koskinen, and M. Schweinberger. (2010). Maximum Likelihood Estimation for Social Network Dynamics. *Annals of Applied Statistics*, 4: 567–588.

Snijders, T.A.B., Steglich, C.E.G., and M. Schweinberger. (2007). Modeling the co-evolution of networks and behavior. In *Longitudinal models in the behavioral and related sciences*, edited by van Montfort, C.A.G.M., Oud, J.H.L., and A. Satorra, pp. 41–71. Mahwah, NJ: Lawrence Erlbaum.

Snijders, T.A.B., van de Bunt, G.G., and C.E.G. Steglich. (2010). Introduction to actor-based models for network dynamics, *Social Networks*, 32, 44–60.

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