Longitudinal Methods of Network Analysis *

Tom A.B. Snijders †

Contents

*Preprint of publication in the *Encyclopedia of Complexity and System Science* (editor-in-chief Bob Meyers) in the *Social Networks* section (section editor John Scott), Springer Verlag, 2009; pp. 5998-6013.

†University of Oxford and University of Groningen. Collaboration with Christian Steglich, Michael Schweinberger, and Mike Pearson in the research leading to this chapter is gratefully acknowledged. Thanks are due to Patrick West and Mike Pearson for the permission to use the data of the *Teenage Friends and Lifestyle Study*.

Glossary

Actors The social actors who are represented by the nodes of the network, and indicated by a label denoted i or j in the set $1, \ldots, n$.

Behavior An umbrella term for changing characteristics of actors, considered as components of the outcome of the stochastic system: e.g., behavioral tendencies or attitudes of human actors, performance, etc. Each behavior variable Z_h is assumed to be measured on an ordinal discrete scale with values $1, 2, \ldots, M_h$ for some $M_h \geq 2$. The value of behavior variable Z_h for actor i is denoted Z_{ih}.

Change determination process The stochastic model defining the probability distribution of changes, conditional on the event that there is an opportunity for change.

Change opportunity model The stochastic process defining the moments where tie indicators can change. This can be either *tie-based*, meaning that an ordered pair of actors (i, j) is chosen and the possibility arises that the tie variable from i to j is changed; or *actor-based*, meaning that an actor i is chosen and the possibility arises that one of the outgoing tie variables from actor i is changed.

Covariates Variables which can depend on the actors (actor covariates) or on pairs of actors (dyadic covariates), and which are considered to be deterministic, or determined outside of the 'stochastic system' under consideration.

Effects Components of the objective function.

Influence The phenomenon that change probabilities for actors' behavior depend on the network positions of the actors, usually in combination with the current behavior of the other actors.

Markov chain A stochastic process where the probability distribution of future states, given the present state, does not depend on past states.

Method of moments A general method of statistical estimation, where the parameters are estimated in such a way that expected values of a vector of selected statistics are equal to their observed values.

Network A simple directed graph representing a relation on the set of actors with binary tie indicators X_{ij} which can be regarded as a state which can change, but will normally change slowly.

Objective function Usually denoted by f_i; the informal description is that this is a measure of how attractive it is to go from an old to a new state. More formally, when there is an opportunity for change, the probability of the change is assumed to be proportional to the exponential transform of the objective function.
The objective function has a similar role as the linear predictor in generalized linear models in statistics, and is specified here as a linear combination of *effects*.

Rate function Usually denoted by λ, the expected number of opportunities for change per unit of time.

Selection The phenomenon that change probabilities for network ties depend on the behavior of one or both of the two actors involved.

Tie indicator A variable X_{ij} indicating by the value $X_{ij} = 1$ that there is a tie $i \to j$, and by the value 0 that there is no such tie. Also called tie variables.

I. Definition

Social networks represent the patterns of ties between social actors. To analyze empirically the mechanisms that determine creation and termination of ties, especially if several mechanisms that may be complementary are studied simultaneously, statistical methods are needed. This chapter is aimed at the case that network panel data are available to the researcher, and treats recently developed statistical models for such data, with corresponding estimation methods. To represent the feedback processes inherent in network dynamics, it is helpful to regard such panel data as momentary observations on a continuous-time stochastic process on the space of directed graphs. Tie-oriented and actor-oriented stochastic models are presented, which can reflect endogenous network dynamics as well as effects of exogenous variables. These models can be regarded as agent-based models, and they can be implemented as computer simulation models. To estimate the parameters of the model, stochastic approximation methods can be used. Social networks are especially interesting because they are important influences on individual behavior – and in turn the network ties are influenced by individual behavior. This two-way influence can be represented by models for the co-evolution of networks and changeable actor attributes. Such models, and statistical methods to analyze panel data on networks and behavior, are also treated. An extensive example is discussed about friendship among teenagers in a school setting.

II. Introduction

When we think of social networks, it is quite natural to think of them as being dynamic. Ties are established, gain in strength, they can blossom and decay, and they may wither or be terminated with a bang. This applies to all kinds of relations – friendship or collaboration between humans, joint ventures between companies, bilateral agreements between countries, and so on. (Note that all these are examples of 'positive' ties; 'negative' ties are not dealt with in this chapter). It is also natural to think of such changes as being dependent on characteristics of the nodes – inclinations and abilities of humans, resources of companies, locations and capacities of countries – and on characteristics of pairs of nodes such as similarity or spatial proximity, as well as on the existing network structure – reciprocation of friendships, transitive closure of friendships (which is the case when friends of friends become friends), group formation of companies or countries. Finally, we are not surprised when such network changes have repercussions – friendships are often thought to have good or bad influences on the individuals concerned, agreements between companies and between countries will have consequences for the performance of the companies and for how the countries fare. Indeed, in the cases of the companies and countries, the links often are created with the purpose of having beneficial consequences for the companies or countries, respectively.

This sets the stage for this chapter, which is concerned with inferential data analysis for network dynamics. The focus is on social networks, and the nodes, or vertices, in the network will be referred to as (social) actors. Data analysis means that methods will be presented for analyzing empirical data on network change. Inference means that the aim is to have methods for testing hypotheses about mechanisms that may drive the network dynamics, and for estimating parameters that figure in such mechanisms. As usual in statistical inference, the 'mechanisms' will be expressed as probability models, also called stochastic models. For hypothesis testing it must be possible to play off one theory against another, which means that when we have theories, or mechanisms, T_1 and T_2, that are not contradictory but could occur together, we need models which can express both mechanisms simultaneously and reflect each mechanism in a set of parameters. The mechanism operates if some of its parameters are nonzero. Then we can test the null hypothesis that T_1 operates but not T_2, against the alternative hypothesis that both mechanisms operate; in other words, we can test for T_2 while controlling for T_1. This requires flexibility of the stochastic models being used: what was stated above implies that it must be possible to specify the models in diverse ways, e.g., with some parameters for effects of actor characteristics and others for effects of various aspects of the existing network structure, and these parameters being estimable from observed data. For data consisting of independent observations the elaboration of these principles of statistical inference in the linear regression model and its generalizations is well known. Network data have the complicating feature that they are not composed of independent observations: the occurrence, creation, or termination of one tie is highly dependent on the existence of other ties. Therefore more complex statistical models are needed, giving an adequate representation of the mutual dependence between the existence of various ties between the actors in the network. But readers who are acquainted with generalized linear statistical models will see that, although the models treated in this chapter have this greater complexity, many elements known from generalized linear models do play a role.

Social network analysis is concerned with diverse types of networks which can be represented by diverse data structures: simple graphs and their generalizations. In line with the current state of the statistical methodology for network dynamics, this chapter is restricted to data structures where the changing network is a changing simple directed graph, where the arcs represent social ties that can be regarded as *states* rather than events. This means that, although changeable, the ties have inertia, a tendency to endure. Friendship between humans and agreements between companies are examples of states; conversations and momentary transactions are events. For networks where ties are states, the dependence between ties can be represented by assuming that changes in the network are dependent upon the existing network structure. In mathematical terms, this is to say that it is reasonable to assume that – given the available 'independent' or 'explanatory' variables – the network is a *Markov chain*. A Markov chain is a stochastic process where the probability distribution of future states, given the present state, does not depend on the past states. Such models were first proposed for network dynamics by [22] and elaborated by [59]. For some sociological applications outside the realm of networks, see [2, 10]. For a network of events (e.g., the network of conversations ongoing at each given moment) the assumption of a Markov chain would be untenable. For a network of states (e.g., the network of joint ventures between companies) the Markov assumption is usually not totally realistic but can be used as a first approximation, and in many cases is the best assumption one can make given the limitation of the available data. This assumption often can be made more plausible by using relevant explanatory variables.

In the start of this introduction, ties were portrayed as possibly gaining in strength or decaying. It would be attractive to reflect this by measuring the ties on an ordinal scale. This is not considered here: we are dealing with simple and not with valued directed graphs, and the *tie indicators*, which are the variables X_{ij} indicating how actor i is tied to actor j, are constrained to having the values 0 or 1, indicating, respectively, absence and presence of the tie $i \to j$. The restriction to binary tie variables is in line with traditional network analysis, but extensions to valued ties are important and are the subject of current work. The Markov assumption will often be more reasonable for valued than for binary ties.

The assumption that we are dealing with a network of relational states must be reflected by the way in which the ties are measured. Measurement in social network analysis is a subject which tends to receive too little attention, and this chapter is no exception. In practice, ties will have to be measured in such a way that observing a tie is a good indicator for the relational state being investigated, such as friendship or collaboration. When the relation under study is a type of communication, which usually has an ephemeral nature, it will be necessary to aggregate the communication over a sufficiently long time interval, so that the resulting variable can be regarded as indicative of a relational state. For example, [25] aggregated email communication to the binary tie variable defined by at least one email being sent over a 60-day period. In other situations, shorter or longer periods may be relevant.

III. Stochastic models for network dynamics

This section presents stochastic models for use in statistical modeling. The inferential aspects (parameter estimation and testing of hypotheses) are treated in Section IV. These models were applied, e.g., to the testing of theories about dynamics of friendship networks [12, 55, 57], of trust networks in firms [56], of artistic prestige [13], and of ties between venture capital firms [8].

The network is represented by the node set $\{1, \ldots, n\}$ with tie variables x_{ij}, where $x_{ij} = 1$ or 0 indicates whether the tie $i \rightarrow j$ is present or absent. The tie variables are collected in the $n \times n$ adjacency matrix $x = \big(x_{ij}\big)$. Self-ties are excluded, so that $x_{ii} = 0$ for all i. The concepts of network (directed graph) and matrix (its adjacency matrix) will be used interchangeably, depending on what is most convenient. In accordance with the usual notation in probability and statistics, random variables will be indicated by capitals; and observations, or other non-random variables, by small letters. The ties are assumed to be outcomes of time-dependent random variables, denoted by $X_{ij}(t)$ and collected in the time-dependent random matrix $X(t)$.

In addition to the network $X(t)$, which can be regarded as the dependent variable of the model, there can be other variables regarded as independent or explanatory variables in the sense that their values are not modeled but accepted as given, and they may influence the network. Such variables are called *covariates* and when depending on the actors they are denoted v_i, while if they depend on pairs of actors (*dyads*) the notation is w_{ij}. Examples are the age of actors (actor variable) and their spatial proximity (dyadic variable).

Basic model definition

The following basic assumptions are made.

1. Time, denoted by t, is a continuous variable.
   This does not mean that it is assumed that observations are made continuously; in most practical cases, observations are made at a number (perhaps a small number) of discrete time moments. However, it is natural and mathematically convenient to assume that there is an underlying process $X(t)$ (which may be observed only partially) which proceeds in continuous time.

2. $X(t)$ is a Markov process.
   This means that the conditional distribution of future states depends on the past only as a function of the present. In other words, to predict the future it is sufficient to know the present state of the network, and knowledge of past states will not improve predictability. This assumption was discussed above. It can be expressed by saying that the network represents a *state*, and usually goes together with inertia, i.e., the tendency of ties to remain in existence unless something special happens.

3. At any given moment t, no more than one tie variable $X_{ij}(t)$ can change.
   This assumption, first proposed by [22], means that changes of ties are not directly coordinated, and ties are mutually dependent only because tie changes will depend on the current total configuration of ties. This is an important simplifying condition and excludes, for example, partner swapping and the coordinated formation of groups. Changes of several ties are decomposed as sequences of changes of single ties.

These assumptions still allow an extremely wide array of probability models, and further specification is necessary. In network change, two aspects can be distinguished: the frequency of tie change, which may depend on the actors involved; and the network structures that tend to be formed by the tie changes – the 'direction' of change. Examples of the former are that younger individuals might change their friendship ties more frequently than older individuals, or that more central actors might change their ties more frequently than peripheral actors. Examples of the latter are tendencies toward tie reciprocation, and toward transitive network closure. These two aspects will be represented by distinct components of the model and distinct parameters, which allow the inference about the one aspect to be relatively undisturbed by inference about the other aspect. The first component is the *change opportunity* process, the second the *change determination* model.

Change opportunity process.
Two specifications of the change opportunity process are given. They use the concept of a *Poisson process*, which is a stochastic process of events occurring at a certain rate λ, which means that the probability that an event occurs in the time interval from t to $t+\epsilon$, where ϵ is a small positive number, is given (in the limit for ϵ tending to 0) by $\lambda\epsilon$. One could say that the rate is the probability of occurrence per unit of time (in short time intervals).

Two specifications of the opportunity process are given here.

1. *Tie-based change opportunities.*

For each tie variable X_{ij}, opportunities for change occur according to a Poisson process with rate λ_{ij}.

2. *Actor-based change opportunities.*

For each actor i, opportunities to establish one new outgoing tie $i \to j$, or dissolve one existing tie $i \to j$, occur according to a Poisson process with rate λ_i.

The rates λ can be constant, or depend on covariates or functions of the current state of the network; if they are not constant, they are called *rate functions*. Tie-based rates λ_{ij}, for example, could depend on the proximity between actors or on their joint embeddedness such as the current number of common friends $\sum_h X_{ih}(t)\, X_{jh}(t)$. Actor-based rates λ_i could depend on actor variables or on positional variables such as actor i's current outdegree $\sum_j X_{ij}(t)$.

Tie-based change opportunities were proposed by Robins and Pattison (personal communication), and correspond to the Gibbs sampling and Metropolis-Hastings procedures for simulating exponential random graph models, see [42]. Actor-based change opportunities were proposed by [46].

When an opportunity for change occurs, there is, in each opportunity model, a set of potential new networks that could be the result of the change. Denoting the current network by x^0, for the tie-based opportunity model this set can be denoted by $\mathcal{C}_{ij}(x^0)$. This is the set of the two possible matrices x where all elements other than x_{ij} are equal to those in the current matrix x^0, and where x_{ij} itself can be either 0 or 1. In the actor-based opportunity model actor i, when confronted with an opportunity for change, chooses one of his outgoing tie variables and changes this into its opposite value, changing 0 to 1 (creating a new tie) or changing 1 to 0 (terminating an existing tie). Therefore the set of potential new networks here is the set composed of x^0 itself together with the $n-1$ matrices which are equal to x^0 except for exactly one non-diagonal element in line i which is replaced by its opposite, $x_{ij} = 1 - x^0_{ij}$.

The set of new possible states is denoted in shorthand applicable to either case by $\mathcal{C}(x^0)$. Since it is allowed that the current situation is continued, it always holds that $x^0 \in \mathcal{C}(x^0)$.

Change determination model.
The choice of the new state of the network is dependent on what is called the *objective function*, which is a function $f_i(x^0, x, v, w)$ depending on the current state of the network x^0, the potential new state x, the actor i, and the covariates summarized here as v (actor covariates) and w (dyadic covariates). The objective function can be interpreted informally as a measure of how attractive it is for actor i to change from state x^0 to state x.

> When actor i has the opportunity to change some outgoing tie variable X_{ij}, given that currently $X(t) = x^0$,
> the set of possible new states of the network is denoted $\mathcal{C}(x^0)$.
> All $x \in \mathcal{C}(x^0)$ differ from x^0 by at most one element x_{ij} for some j.
> When there is an opportunity for change from the current state x^0,
> the probabilities of the values of the next state $x \in \mathcal{C}(x^0)$
> are proportional to $\exp\big(f_i(x^0, x, v, w)\big)$.

The models can be summarized as follows. For the tie-oriented model, when an opportunity for change occurs, it refers to some pair (i, j); opportunities for changing X_{ij} occur at a rate λ_{ij} for each pair (i, j). When such an opportunity occurs, the probability that x^0 changes to the different state x is given by

$$\begin{aligned} &\mathrm{P}\{X(t) \text{ changes to } x \mid (i,j) \text{ has a change opportunity at time } t, X(t) = x^0\} \\ &= p_{ij}(x^0, x, v, w) = \frac{\exp\big(f_i(x^0, x, v, w)\big)}{\exp\big(f_i(x^0, x^0, v, w)\big) + \exp\big(f_i(x^0, x, v, w)\big)} \ , \end{aligned} \tag{1}$$

where x and x^0 are identical except for $x_{ij} = 1 - x^0_{ij}$.

For the actor-oriented model, opportunities for change occur for actors i. Opportunities for actor i to change one of the outgoing tie variables X_{ij} ($j = 1, \ldots, n; j \neq i$) occur at a rate λ_i. The set of permitted new states, following on a given current state x^0, is $\mathcal{C}(x^0)$. The probability that the new state is x, provided that x is permitted (i.e., $x \in \mathcal{C}(x^0)$), is given by

$$\begin{aligned} &\mathrm{P}\{X(t) \text{ changes to } x \mid i \text{ has a change opportunity at time } t, X(t) = x^0\} \\ &= p_i(x^0, x, v, w) = \frac{\exp\big(f_i(x^0, x, v, w)\big)}{\sum_{x' \in \mathcal{C}(x^0)} \exp\big(f_i(x^0, x', v, w)\big)} \ . \end{aligned} \tag{2}$$

The two model components can be put together by giving the transition rate matrix, also called Q-matrix, of which the elements are defined by

$$q_{x^0,x} = \lim_{dt \downarrow 0} \frac{\mathrm{P}\big\{X(t+dt) = x \mid X(t) = x^0\big\}}{dt} \qquad (x \neq x^0)$$

(see textbooks on continuous-time Markov chains, such as [34]). Note that the assumptions imply that

$$q_{x^0,x} = 0 \text{ whenever } x_{ij} \neq x^0_{ij} \text{ for more than one element } (i,j).$$

For digraphs x and x^0 which differ from each other only in the element with index (i,j), the elements of the Q-matrix are given for the tie-based opportunity process by

$$q_{x^0,x} = \lambda_{ij}(x^0, v, w)\, p_{ij}(x^0, x, v, w) \tag{3}$$

and for the actor-based opportunity process by

$$q_{x^0,x} = \lambda_i(x^0, v, w)\, p_i(x^0, x, v, w)\ . \tag{4}$$

Tie-based models with constant change rates and objective functions defined as $f(x, v, w)$ (not depending on the preceding state x^0 or on the actor i) can be regarded as Metropolis-Hastings dynamics (cf. [34]) for obtaining random draws from the digraph probability distribution with probability function

$$c \exp\big(f(x, v, w)\big)$$

where c is a normalizing constant. In statistical mechanics $f(x, v, w)$ then will be called a potential function, see [32]. When $f(x, v, w)$ is a linear combination as in (6) below, the distribution is an exponential random graph model for which the Metropolis-Hastings algorithm is treated in [42].

The actor-based opportunity model was proposed in [46] and, for a different data structure, in [45], as a stochastic actor-oriented model. In this model, the network dynamics is regarded as being driven by the social actors. Actors are assumed to control their outgoing ties, subject to inertia and the current network structure. This point of view is in accordance with the methodological approach of structural individualism [54, 60], where actors are assumed to be purposeful and to behave subject to structural constraints. The purposes and constraints of the actors are summarized in the objective functions. One way to obtain the probabilities (2) is to assume that at each opportunity for change, actor i myopically optimizes the objective functions plus a random term, under the constraint that only one tie can change at a time. The myopia means that the actor only optimizes the state of the network that will be the immediate result of this change, without considering later network structures that might result in the future further ahead. The random term expresses otherwise unmodeled purposes and constraints. When the random terms have independent Gumbel distributions (the precise mathematical form of which does not matter for the present exposition), the choice probabilities are given by (2) (cf. [30, 45, 46]).

Simulation.
The Markov process defined above can be iterating by the following algorithm. The various steps in the algorithm can be derived using basic properties of Poisson processes and conditional probabilities (see [34]).

1. The process starts with a given time t and current state $X(t) = x^0$.

2. For the tie-based opportunity process define $\lambda = \sum_{ij} \lambda_{ij}$, and for the actor-based opportunity process $\lambda = \sum_i \lambda_i$.
   Let U be an independently drawn random number, uniformly distributed between 0 and 1, and let $\Delta\, t = -\ln(U)/\lambda$. Note that $\Delta\, t$ has the exponential distribution with parameter λ.
   Change t into $t + \Delta\, t$.

3a. In the case of the tie-based opportunity process, choose a random pair (i, j) (with $i \neq j$) with probabilities λ_{ij}/λ. With probability given by (1), change $X_{ij}(t)$ into $1 - x^0_{ij}$.

3b. In the case of the tie-based opportunity process, choose a random actor i with probabilities λ_i/λ.
   To have a way for denoting the permitted new digraphs which are elements of $\mathcal{C}_i(x^0)$, define by $x^0(i \rightsquigarrow j)$ for $j \neq i$ the digraph which is equal to x^0 except only that $\big(x^0(i \rightsquigarrow j)\big)_{ij} = 1 - x^0_{ij}$; define $x^0(i \rightsquigarrow i) = x^0$. Then choose a random j with probabilities

$$p_i(x^0, x^0(i \rightsquigarrow j), v, w) = \frac{\exp\big(f_i(x^0, x^0(i \rightsquigarrow j), v, w)\big)}{\sum_{x' \in \mathcal{C}(x^0)} \exp\big(f_i(x^0, x', v, w)\big)}\,, \qquad (5)$$

   which is just the same as (2). If $j \neq i$, change $X_{ij}(t)$ into $1 - x^0_{ij}$.

4. Go to step 1.

The stochastic process on the space of digraphs can be defined by this simulation algorithm just as well as by the Q-matrices (3) and (4), respectively.

Specification of the change determination process

The changes can be regarded for tie-based opportunities as determinations of new values for X_{ij} according to a binary logistic regression model (see [23]); and for actor-based opportunities, according to a multinomial logistic regression model (see [29]). In the further elaboration the parallel with logistic regression is followed because the objective function f_i is specified as a linear combination

$$f_i(x^0, x, v, w) = \sum_k \beta_k \, s_{ki}(x^0, x, v, w) \tag{6}$$

where the functions s_{ki} are so-called *effects* driving the network dynamics while the weights β_k are parameters indicating the force of these effects and which can be estimated from the data.

The specification of the model will be the choice of a limited set of such effects for use in (6). A list of some effects is the following. In the formulae, replacing an index by a $+$ means that a sum is taken over this index. Many examples of effects do not depend on x^0 or on v or w, and these arguments are then dropped from the notation.

Outdegree effect $s_{1i}(x) = x_{i+} = \sum_j x_{ij}$

This effect models the tendency to have ties at all; this tendency will also be influenced by all other effects, and therefore the interpretation of its parameter is conditional on the further selection of effects included in the model. In many models this is the only effect to which those actors j contribute who have no reciprocal link to i nor any links with any others to whom i is linked. In such models, its weight β_1 can be interpreted as the 'value' for actor i of a tie to such an other actor who is further completely isolated from i's personal network.

Reciprocity effect $s_{2i}(x) = \sum_j x_{ij} \, x_{ji}$

This is the number of reciprocated ties for actor i. It models the tendency toward reciprocation of choices. Thus, a higher value for its parameter β_2 will imply a higher tendency to forming reciprocated ties.

Degree distribution. The following three effects are related to modeling the dynamics of the degree distribution. Since the data are directed graphs, three distinct aspects of the degree distribution are the variability of indegrees, variability of outdegrees, and association between in- and outdegrees. Sometimes the term of preferential attachment is used [1, 40] for the increased attractiveness of ties to nodes that already have high degrees. In our model these three aspects of preferential attachment can be expressed by including in the objective function terms depending on in- and out-degrees. The precise functional form for the effects is determined also by the requirement that the resulting model be amenable to statistical inference. Experience shows that using the square root of the degrees often leads to more stable estimation of the parameters than using the raw (untransformed) degrees. This suggests that, for many empirical networks, the models with squared roots of the degrees are better descriptions of reality than the models with untransformed degrees. Therefore only the models using the square roots are presented here.

Popularity effect (square root measure) $s_{3i}(x) = \sum_j x_{ij}\,\sqrt{x_{+j}}$

This effect is defined by the sum of the square roots of indegrees of the others to whom i is tied. In other words, popularity of other actors is measured by the square root of their indegree. The root-popularity effect models the tendency to form ties to those actors who have high indegrees already (the Matthew effect in networks; see the chapter on *GRAPH THEORETICAL APPROACHES TO SOCIAL NETWORK ANALYSIS*). This will be reflected by the dispersion of the indegrees.

Activity effect (square root measure) $s_{4i}(x) = \sum_j x_{ij}\,\sqrt{x_{j+}}$

This effect is defined by the sum of the square roots of outdegrees of the others to whom i is tied. The effect models the tendency to form ties to those actors who have high outdegrees already. This will be reflected by the association between indegrees and outdegrees.

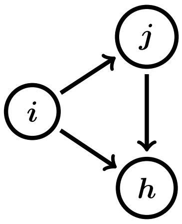
Own outdegree, power 1.5, effect $s_{4i}(x) = \sum_j x_{ij}\,\sqrt{x_{i+}} = x_{i+}^{1.5}$

The first expression show that each new tie has a ‘value’ equal to the actor’s outdegree, comparing to a unit ‘value’ for the outdegree effect. This effect expresses that actors who already have many outgoing ties have a higher propensity to establish new ties. This will lead to a greater dispersion of outdegrees.

Network closure. The following two effects are related to network closure, also called transitivity or clustering: for friendship networks, this expresses that friends of friends tend to become friends. In addition to the two effects mentioned here, tendencies toward structural balance [7] and tendencies to avoid geodesic distances equal to 2 can also be implemented as effects that will lead to transitivity; also see [46, 47].

Transitive triplets effect $s_{5i}(x) = \sum_{j,h} x_{ij}\, x_{ih}\, x_{jh}$

This formula represents the number of transitive patterns in i's ties as indicated in the figure below. A transitive triplet for actor i is a configuration (i, j, h) in which all three of the ties $i \to j, j \to h, i \to h$ are present (and irrespective of whether there are also other ties between these three actors).



Transitive triplet

This models the tendency toward network closure, where (for a positive parameter) formation of the tie $i \to h$ becomes increasingly likely when there are more indirect connections ('two-paths') $i \to j \to h$.

Transitive ties effect $s_{6i}(x) = \sum_h x_{ih}\ \max_j(x_{ij}\, x_{jh})$

This is the number of actors j to whom i is directly as well as indirectly tied, i.e., for which there exists at least one h such that (i, j, h) is a transitive triplet. The transitive triplets and transitive ties effects are two distinct ways of modeling tendencies toward network closure. For the effect on the probability of forming the tie $i \to h$, the number of two-paths $i \to j \to h$ makes no difference for the transitive ties effect, as long as there is at least one indirect connection; for the transitive triplets effect the contribution to the objective function increases linearly with the number of two-paths.

Three-cycle effect $s_{7i}(x) = \sum_{j,h} x_{ij}\, x_{jh}\, x_{hi}$

This is the number of three-cycles $i \to j \to h \to i$ in which actor i is involved. This effect models the tendency toward forming three-cycles, which is the simplest form of generalized exchange [3, 27] (i gives to j, j gives to h, and h gives to i) and which is opposed to hierarchy.

Covariate effects. Actor covariates v can influence the propensities to form or terminate ties in different ways, because this propensity might be influenced by the value for the sender or the receiver of the tie, or by some combination of these two values.

v-related popularity $s_{8i}(x, v) = \sum_j x_{ij}\, v_j$

The v-related popularity effect is defined by the sum of the covariate over all actors to whom i is tied. Positive parameter values will imply that ties to actors with high v values are more attractive. This will lead to a tendency toward a correlation between v_i and the indegree of i.

v-related activity $s_{9i}(x, v) = v_i\, x_{i+}$

In words, this is v_i times i's outdegree. Positive parameter values will imply that actors with high v values tend to make more ties. This will lead to a tendency toward a correlation between v_i and the outdegree of i.

v-related similarity $s_{10,i}(x, v) = \sum_j x_{ij}\, \text{sim}_v(i, j)$

Here $\text{sim}_v(i, j)$ indicates the similarity between actors i and j defined by $\text{sim}_v(i, j) = 1 - (|v_i - v_j|/\Delta)$, where $\Delta = \max_{h,k} |v_h - v_k|$ is the observed range of the covariate v. Thus, the effect is is the sum of similarities between i and the others to whom he is tied. Positive parameter values will imply that there is a preference for ties between actors with similar values of v_i and v_j.

v ego-alter interaction $s_{11,i}(x, v) = \sum_j x_{ij}\, v_i\, v_j$

This product interaction is an alternative to the similarity effect for expressing that the propensity for a tie to exist depends on the combined values of the v for the sending actor ('ego', i) and the receiving actor ('alter', j).

Main effect of w $s_{12,i}(x, w) = \sum_j x_{ij}\, w_{ij}$

For a dyadic covariate w, this is defined by the sum of the values of w_{ij} for all other actors j to whom i is tied.

It is clear that the list can be extended indefinitely, and that researchers have to make a limited choice reflecting theoretical and content-matter knowledge and interest. The potential complexity of network dynamics justifies to have many candidate effects that may be used to model the network dynamics. For instance, when a researcher wishes to estimate a model for testing whether there is a preference for choosing network partners with similar v values, while controlling for the tendency to have ties, and the tendencies for reciprocation and transitive closure, then effects s_1, s_2, s_5 and/or s_6, and s_{10} should be included in the model (6). Since transitive closure could be expressed by effects s_5 as well as s_6, there may be no strong prior arguments for choosing between these two 'control' effects, and empirical grounds could be used to choose either one or both. If there are grounds to suspect that the v_i values may also be associated with in- or outdegrees, the popularity and activity effects s_8 and s_9 may also be included.

Specification of the change opportunity process

For the model specification it should be noted that the 'social time' which determines the speed of change of the network is not necessarily the same as the physical time elapsing between consecutive observation moments. Given the absence of an extraneous definition of this 'social time', it is not a restriction to set to 1 the total time elapsed between each pair of consecutive observations. If there are $M \geq 3$ observation moments, it is advisable to specify distinct rate parameters ρ_m governing the frequency of opportunities for change between t_m and t_{m+1}, and allow ρ_1, ρ_2, etc., to be different. If the change rate further is constant (independent of actors), ρ_m then represents the expected number of opportunities for change between t_m and t_{m+1}. This is the expected number per ordered pair (i, j) in the case of tie-based opportunities, and per actor i in the case of actor-based opportunities. The symbol ρ will denote the vector $(\rho_1, \ldots, \rho_{M-1})$.

In the more general case the rate function can be defined, e.g., for the actor-based model, by a function depending on actor covariates and positional characteristics of the actors. When, for example, a dependence on one covariate v_i and the current out-degree x^0_{i+} is considered, a logarithmic link function could be used giving a model such as

$$\lambda_i(x^0, v) \;=\; \rho_m \,\exp\big(\alpha_1 v_i + \alpha_2 x_{i+}\big).$$

In general the symbol α will be reserved for parameters indicating dependence of the rate function on covariates and network characteristics.

IV. Statistical estimation and testing

The most usual type of longitudinal network data is panel data, where for $M \geq 2$ time points, an observation $x(t_m)$ is available of the network on the same set $\{1, \ldots, n\}$ of actors.

These models can be simulated on computers in rather straightforward ways (the algorithm is written out in [47]). Parameter estimation, however, is more complicated, because the likelihood function or explicit probabilities can be computed only for uninteresting models. This section presents the Method of Moments (MoM) estimates proposed in [46]. Maximum Likelihood (ML) estimators are presented in [48]. In the current implementation, ML estimators are more time-consuming than MoM estimators, and can be used only for relatively small data sets. For the more straightforward models (dynamics of networks only, no endowment functions), the MoM method is hardly less efficient than the ML method. For more complicated models, where it is important to squeeze every bit of information out of the data, it can be useful to employ ML methods. In the following description of the estimation method, the parameter vector (ρ, α, β) is denoted by θ.

It is undesirable to make the restrictive assumption that the distribution of the process is stationary. Instead, for each observation moment t_m $(m = 1, \ldots, M - 1)$ the observed network $\mathbf{x}(t_m)$ can be used as a conditioning event for the distribution of $\mathbf{X}(t_{m+1})$. The Method of Moments requires that a vector of statistics $U_{m+1} = U\big(\mathbf{X}(t_m), \mathbf{X}(t_{m+1})\big)$ is utilized, such that the expected value

$$\mathrm{E}_\theta\big\{U\big(\mathbf{X}(t_m), \mathbf{X}(t_{m+1})\big) \mid \mathbf{X}(t_m) = \mathbf{x}(t_m)\big\}$$

is sensitive to the parameter θ. Given the conditioning on the preceding observation, the moment equations, or estimating equations, can then be written as

$$\sum_{m=1}^{M-1} \mathrm{E}_\theta\big\{U\big(\mathbf{X}(t_m), \mathbf{X}(t_{m+1})\big) \mid \mathbf{X}(t_m) = \mathbf{x}(t_m)\big\} \;=\; \sum_{m=1}^{M-1} U\big(\mathbf{x}(t_m), \mathbf{x}(t_{m+1})\big)\;. \tag{7}$$

It turns out that suitable statistics are the following. The number of changed ties between consecutive observations,

$$\sum_{i,j} |X_{ij}(t_{m+1}) \;-\; X_{ij}(t_m)|\;,$$

is especially sensitive to the rate of change ρ_m. A vector of statistics sensitive especially to β is the sum of the individual objective functions

$$\sum_i f_i\big(\mathbf{X}(t_{m+1})\big)\;.$$

To solve the estimating equation (7), in the absence of ways to calculate analytically the expected values, stochastic approximation methods can be used. Variants of the Robbins-Monro [9, 41] algorithm have been used with good success. This is a stochastic iteration method which produces a sequence of estimates $\theta^{(N)}$ which is intended to converge to the solution of (7), and which works here as follows. For a given provisional estimate $\theta^{(N)}$, the model is simulated so that for each $m = 1, \ldots, M - 1$, a simulated random draw is obtained from the conditional distribution of $\mathbf{X}(t_{m+1})$ conditional on $\mathbf{X}(t_m) = \mathbf{x}(t_m)$. This simulated network is denoted $\mathbf{X}^{(N)}(t_{m+1})$. Denote $U_m^{(N)} = U\big(\mathbf{x}(t_m), \mathbf{X}^{(N)}(t_{m+1})\big)$ and

$U^{(N)} = \sum_{m=1}^{M-1} U_m^{(N)}$, and let u^{obs} be the right-hand side of (7). Then the iteration step in the Robbins-Monro algorithm for obtaining the Method of Moments estimate is given by

$$\theta^{(N+1)} = \theta^{(N)} - a_N D^{-1} \left(U^{(N)} - u^{\text{obs}} \right) , \tag{8}$$

where D is a suitable matrix and a_N a sequence of positive constants tending to 0. Tuning details of the algorithm, including the choices of D and a_N, are given in [46]. The experience with the convergence of this algorithm is quite good. The standard errors can be computed using the standard formulae of standard errors for the Method of Moments, based on the delta method, and applying simulation methods; such simulation methods are discussed in [44]. Bayesian estimators for these models are presented in [24] and Maximum Likelihood estimators in [48].

V. Example: dynamics of adolescent friendship

As an example, the adolescent friendship network is considered of a year cohort at a secondary school in Glasgow (Scotland), studied in the Teenage Friends and Lifestyle Study [36, 38]. This data set was collected at three measurement points t_1, t_2, t_3 in 1995–1997, at intervals of roughly one year, starting when the pupils were 12-13 years old. Here the network is studied that is formed by the 129 (out of 160) pupils who were present at all three measurement waves. Sex (boys scored as 1, girls as 2) and drinking behavior are used as actor variables, drinking (alcohol consumption) being measured on a 5-point scale ranging from 1 (not at all) to 5 (more than one a week). Both variables are centered around the mean, which is 1.43 for sex and 2.60 for drinking (averaged over t_1 and t_2). The data set is used as an illustration; more wide-ranging analyses are presented in [37, 52, 53].

Three models are presented. Rate parameters are assumed constant within periods between observation moments, and the duration of the periods is (arbitrarily but without loss of generality) set at 1. The first model contains only the most basic dyadic and triadic effects: outdegree, reciprocity, transitive triplets, and three-cycles. The second adds to this the three effects to model more precisely the degree distribution. The third model adds to these structural effects the effects of two covariates: gender and alcohol consumption.

Parameter estimates are approximately unbiased and normally distributed; for this assertion there is no mathematical proof yet, but it is supported by simulation studies. Therefore, effects can be tested by referring the studentized estimates (or t-ratios, i.e., estimate divided by standard error) to a standard normal distribution. When the t-ratio exceeds 2 in absolute value, the effect can be interpreted as being significant at the significance level of 5 %.

Table 2: Parameter estimates for modeling evolution of friendship network, Glasgow school cohort. Standard errors between parentheses.

	Model 1		Model 2		Model 3	
Effect	par.	(s.e.)	par.	(s.e.)	par.	(s.e.)
Rate 1	11.82	(1.00)	12.05	(1.15)	12.19	(1.13)
Rate 2	9.23	(0.83)	9.05	(0.79)	9.09	(0.77)
Outdegree	–2.697	(0.047)	–0.16	(0.34)	–0.18	(0.41)
Reciprocity	2.38	(0.10)	2.48	(0.12)	2.22	(0.12)
Transitive triplets	0.459	(0.033)	0.569	(0.036)	0.544	(0.034)
Three-cycles	–0.57	(0.10)	–0.55	(0.11)	–0.44	(0.11)
Popularity (sq. root)	–		0.223	(0.095)	0.184	(0.093)
Activity (sq. root)	–		–0.89	(0.14)	–0.92	(0.17)
Own outdegree, power 1.5	–		–0.560	(0.084)	–0.587	(0.094)
Sex (F) popularity	–		–		–0.16	(0.11)
Sex (F) activity	–		–		0.12	(0.12)
Sex similarity	–		–		0.904	(0.097)
Drinking popularity	–		–		–0.026	(0.030)
Drinking activity	–		–		–0.086	(0.038)
Drinking ego $\times$ alter	–		–		0.107	(0.024)

The table present first the estimated rate parameters. These indicate that the pupils had about 12 opportunities for changes in the first period ($t_1 - t_2$), and about 9 in the second period ($t_2 - t_3$).

The parameters for the objective function, which the table shows next, are more important for the interpretation. There are strong tendencies toward reciprocity and transitivity, and a tendency away from three-cycles. This is the case in all three models, although the parameter estimates are slightly different. This indicates that the structural features of transitive closure and of local hierarchy (the interpretation of the negative three-cycle effect) cannot be 'explained away' by tendencies in tie formation and dissolution that are associated to degrees, sex, or alcohol consumption.

The degree effects in Models 2 as well as 3 show that there is a positive popularity effect – with a borderline significance at the 5% level in Model 3; there is a negative activity effect; and a negative effect of own outdegree raised to the power 1.5. The positive popularity effect suggests that differential values in in-degrees tend to be self-sustaining, leading to rather strongly dispersed in-degrees. The other two, negative, degree-related effects indicate that differences in out-degrees, as well as correlations between in- and out-degrees, tend to be self-correcting, leading to relatively low dispersions of out-degrees and low in-degree – out-degree correlations.

Of the three gender-related effects, only the similarity effect is significant. In this age range, a strong preference for same-sex friendships is to be expected (the interest which perhaps exists in the other sex is not reported as friendship). For alcohol consumption, the interaction effect shows that those who drink more themselves have a higher preference for friends who also drink more; and the activity effect shows that those drinking more tend to mention less friends, for friends of average drinking habits (where the contribution of the interaction is nil). This is most clearly expressed by jointly considering the three contributions related to drinking behavior that can be made by the tie $i \to j$. In a formula, this is represented by the coefficient of the variable y_{ij} in the objective function. Recall that the actor variables are centered around the mean. Using the formulae for effects $s_8,\ s_9$, and s_{11} and filling in the coefficients in Table 1 yields

$$-0.026\,(v_j - \bar{v}) \;-\; 0.086\,(v_i - \bar{v}) \;+\; 0.107\,(v_j - \bar{v})(v_i - \bar{v})\;,$$

with $\bar{v} = 2.60$. The following picture illustrates the contributions to the objective function made by ego's drinking behavior v_i and alter's drinking behavior v_j.

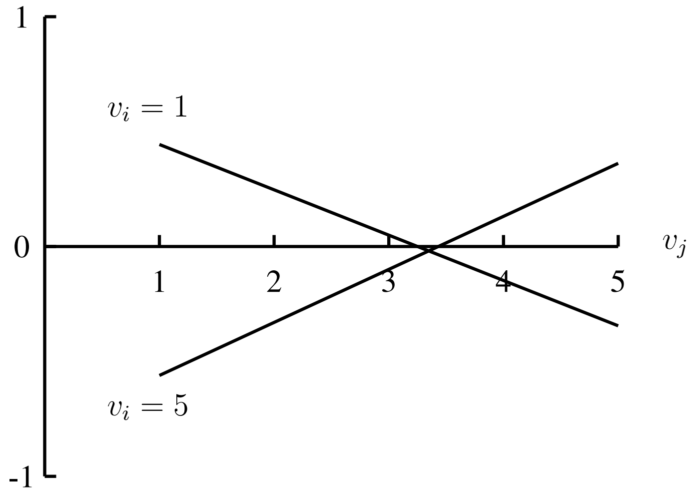


Figure 1. Contributions of drinking behaviors $v_i,\ v_j$ to the objective function for friendship.

It can be concluded that those who do not drink alcohol ($v_i = 1$), prefer friends who drink no or little alcohol, while the reverse is true for those who drink a lot of alcohol. The range of this contribution is 0.8 for those with $v_i = 1$ and 0.9 for those with $v_i = 5$, which is comparable to the value 0.9 of the gender similarity effect. Thus, for those with the highest and the lowest values of alcohol use, the importance of the alcohol use of potential friends, when comparing potential friends with the minimum and the maximum alcohol use, is approximately as great as the importance of their gender. On the other hand, for those with medium values of alcohol use, the alcohol use of potential friends plays virtually no role at all.

VI. Models for the co-evolution of networks and behavior

The importance of networks derives, for an important part, from the effects of networks on the behavior and performance of the actors. This can be, for example, because of influence between friends or between collaborating partners [16], because of exchange of resources [28], or because of structural advantages [4].

A few examples drawn from the many studies in this field are concerned with influence of friends of adolescents on smoking behavior [14, 37, 52], effects of acquaintances on labor market outcomes [17], competitive advantage of firms [18], effects of friends on delinquency [6, 19], effects of position in patent citation networks on growth rates of companies [39], and job performance [51].

This type of changing individual attributes, which might range from, e.g., behavioral tendencies and attitudes of individuals to performance of companies, will be briefly referred to as 'behavior'. When the ties in the network are influenced by the behavior and the behavior, in turn, is influenced by the network, a mutual feedback arises between the network and the behavior. The network structure of the ties between the actors together with their behavior constitute the endogenously changing environment for each of the actors [61].Thus, this approach is well-suited to study macro-micro-macro questions of the type discussed by [11].

The vector of attributes for actor i at time t is denoted

$$z_i(t) \;=\; \big(z_{i1}(t), \ldots, z_{iH}(t)\big) \;,$$

where $z_{ih}(t)$ denotes the h^{th} attribute of actor i. For the n actors these are stacked in the matrix $z(t)$, which is regarded as the outcome of a random matrix $Z(t)$.

To model the co-evolution of networks and behavior, where the network and the behavior influence one another dynamically, the stochastic process $\big(X(t), Z(t)\big)$ is considered. This is treated in quite the same way as the stochastic process $X(t)$ was treated above. It is assumed that $z_i(t)$ represents an enduring, but changeable, state of actor i rather than a momentary behavior, so that it is a sensible approximation to assume that $\big(X(t), Z(t)\big)$ is a Markov process in continuous time. Now the change probabilities of $X(t)$ will depend on the current state of $X(t)$ as well as $Z(t)$. The fact that network change is co-determined by behavior is called *behavior-dependent selection*, and will often be referred to briefly as *selection*. Similarly, the change probabilities of $Z(t)$ will depend on the current state of $Z(t)$ as well as $X(t)$, and this will be called *influence*. This terminology was also used, e.g., by [14].

To remain close to the framework for network modeling, it is assumed that all of the behavior variables z_{ih} are measured on a discrete ordinal scale, with values coded as consecutive integers $\{0, 1, \ldots, M_h\}$ for some $M_h \geq 1$. The analogue of the simplifying assumption (3) made above for network change is the following. This decomposes the change between consecutive observations $\big(x(t_m), z(t_m)\big)$ and $\big(x(t_{m+1}), z(t_{m+1})\big)$ into a sequence of the smallest possible steps.

3'. At any given moment t, no more than one of all the variables $X_{ij}(t), Z_{ih}(t)$ can change. When $Z_{ih}(t)$ changes, at any given instant it can change only to an immediate neighboring value, i.e., by a decrease or increase of 1 (permitted only if this step does not take Z_{ih} outside of the permitted range from 0 to M_h).
This means that there is no direct coordination between changes in ties and changes in behavior, and the dependence between networks and behavior is brought about because both react to each other.

The model for the network is just like it was above, the rate function and objective function now being denoted by λ^X and f_i^X, and allowed to depend on the current behavior $Z(t)$, to represent behavior-dependent selection. For each of the dependent behavior variables Z_h there also is a rate function λ_i^{Zh} driving the frequency of changes and an objective function f_i^{Zh} defining the probabilities of behavior changes when there is an opportunity of change. Their dependence on the current network will represent influence.

For each actor i, opportunities to change behavior Z_{ih} occur according to a Poisson process with rate λ_i^{Zh}.

When actor i has the opportunity to change behavior Z_{ih} given that currently $\big(X(t), Z(t)\big) = (x^0, z^0)$, there are three possible new states (x^0, z), where either $z = z^0$ or the only difference between z and z^0 is that $z_{ih} = z^0_{ih} - 1$ or $z_{ih} = z^0_{ih} + 1$; unless one of these values is outside the range $\{0, \ldots, M_h\}$, in which case there are only the two remaining possible new states. The probabilities of going from state (x^0, z^0) to state (x^0, z) are proportional to $\exp\big(f_i^{Zh}(z^0, z, x^0, v, w)\big)$.

The change probabilities in this model are given by

$$
\begin{aligned}
&\mathrm{P}\{Z(t) \text{ changes to } z \mid i \text{ has an opportunity to change } Z_{ih} \text{ at time } t, \\
&\qquad\qquad\qquad\qquad\qquad\qquad X(t) = x^0, Z(t) = z^0\} \\
&\qquad = \frac{\exp\big(f_i^{Zh}(z^0, z, x^0, v, w)\big)}{\sum_{z' \in \mathcal{C}^{Zh}(z^0)} \exp\big(f_i^{Zh}(z^0, z', x^0, v, w)\big)} \ ,
\end{aligned}
\tag{9}
$$

where $\mathcal{C}^{Zh}(z^0)$ is the set of two or three permitted new values for Z_h.

The effects in the objective function f_i^X for network changes now are denoted by s_{ki}^X with weights β_k^X, and likewise the objective function for behavior changes is assumed to be expressed as a linear combination of effects weighted by parameters,

$$
f_i^{Zh}(z^0, z, x^0, v, w) = \sum_k \beta_k^{Zh} \, s_{ki}^{Zh}(z^0, z, x^0, v, w) \ . \tag{10}
$$

A set of possible effects that could be included in the objective function for behavior are the following. For simplicity of notation, it is assumed that there is only one behavior variable, so that the index h can be dropped from the notation. The first two effects are used to define the shape of the objective function as a function of z_i, and other terms depending on z_i could be added.

Linear shape effect $\quad s_{i1}^Z(z) = z_i$

Squared shape effect $\quad s_{i2}^Z(z) = z_i^2$

The linear and squared shape effects together define what could be regarded as a quadratic preference function on the behavior, $\beta_1^Z z_i + \beta_2^Z z_i^2$. The word ‘preference function’ is used with some reluctance, because it is used here only as an easy shorthand term for the combined short-term result of preferences and constraints, depending only on the actor’s behavior z_i itself, net of the other terms in the behavior objective function. If $\beta_2 < 0$ this is a unimodal function of z_i. If $\beta_2 > 0$, on the other hand, and if the minimum of the quadratic function is assumed within the range of the behavior variable, then the behavior is drawn to the extremes of the range, with actors already low on z_i being drawn to low values and actors already high on z_i being drawn to high values. This can represent, e.g., addictive behavior.

Other nonlinear functions of z_i could, of course, also be included.

Indegree effect $\quad s_{i3}^Z(z, x) = z_i \, x_{+i}$

This represents that actors with a high indegree (‘popular’ actors) have a higher tendency toward high values of the behavior.

Outdegree effect $s^Z_{i4}(z,x) = z_i\, x_{i+}$

This represents that actors with a high outdegree ('active' actors) have a higher tendency toward high values of the behavior.

Total similarity effect $s^Z_{i5}(z,x) = \sum_j x_{ij}\, \text{sim}_z(i,j)$

The dyadic similarity sim_z is as defined above, now applied to the dependent behavior Z. The total similarity effects adds the similarity values between i and the actors toward whom i has a tie. This is the primary representation of social influence: the preference for behavior which is close to that of one's network members.

Average alter effect $s^Z_{i6}(z,x) = z_i\, x_{i+}^{-1} \sum_j x_{ij}\, z_j$

The coefficient of z_i is the average behavior of the actors to whom i is tied (i's 'alters', j). The coefficient is defined as 0 when the outdegree x_{i+} is 0. This is another representation of social influence: the preference for behavior depends on the average behavior of one's network members.

The parameter estimation for this model is discussed in [49].

VII. Example: co-evolution of adolescent friendship and alcohol use

As an example for the co-evolution of networks and behavior the data of the Glasgow school cohort is used again, but now the alcohol consumption is used as a dependent, or endogenous, variable. In the treatment in Table 2 the alcohol consumption was used as an exogenous variable, i.e., its values were accepted as if determined by processes independent of the network. Now we follow a co-evolution approach where it is assumed that the dynamics in alcohol consumption can be co-determined by the network just as the network dynamics can be co-determined by the alcohol consumption.

Table 4: Parameter estimates for modeling co-evolution of friendship network and alcohol consumption, Glasgow school cohort

Effect	par.	(s.e.)
Friendship dynamics		
Rate 1	12.13	(1.26)
Rate 2	9.09	(0.97)
Outdegree	–0.22	(0.43)
Reciprocity	2.19	(0.12)
Transitive triplets	0.529	(0.036)
Three-cycles	–0.42	(0.10)
Popularity (sq. root)	0.193	(0.086)
Activity (sq. root)	–0.92	(0.15)
Own outdegree, power 1.5	–0.581	(0.099)
Sex (F) popularity	–0.16	(0.10)
Sex (F) activity	0.12	(0.13)
Sex similarity	0.90	(0.11)
Drinking popularity	–0.026	(0.043)
Drinking activity	–0.118	(0.057)
Drinking ego $\times$ alter	0.166	(0.043)
Alcohol consumption dynamics		
Rate 1	1.48	(0.25)
Rate 2	2.22	(0.37)
Linear shape	0.36	(0.42)
Squared shape	–0.34	(0.14)
Indegree	0.07	(0.12)
Outdegree	–0.08	(0.17)
Sex (F)	–0.02	(0.23)
Average alter	0.83	(0.35)

To obtain interpretations of the numerical parameter values, it must be noted that this table refers to the centered actor variables. The raw variable ranges from 1 to 5, with averages at the three observations rising from 2.5 through 2.7 to 3.1. The overall mean, 2.8, is subtracted from the observations. We shall denote by z_i the raw value of actor i's alcohol consumption scored 1–5, by $\bar{z}$ the average equal to 2.8, and by $\tilde{z}_i$ the average of the alcohol consumption of i's friends.

The interpretation of the friendship dynamics is quite the same as in the analysis where alcohol was treated as an exogenous variable. For the dynamics of alcohol consumption, the indegree, outdegree, and sex of the actor do not seem to have an important influence. The effect of the average drinking behavior of the friends of the focal actor does have a significant effect, however, with a t-value of $0.83/0.35 = 2.4$. When we ignore the small and non-significant effects of outdegree, indegree, and sex, the remaining part of the objective function for drinking behavior is

$$\begin{aligned} &0.36\,(z_i - \bar{z}) \,+\, 0.83(\tilde{z}_i - \bar{z}))(z_i - \bar{z}) \,-\, 0.34(z_i - \bar{z})^2 \\ &= \big(0.36 \,+\, 0.83(\tilde{z}_i - \bar{z})\big)(z_i - \bar{z}) \,-\, 0.34(z_i - \bar{z})^2 \,. \end{aligned}$$

This is a quadratic function, unimodal, with a maximum at

$$z_i \,=\, \bar{z} \,+\, \frac{0.83}{2 \times 0.34}\,(\tilde{z}_i - \bar{z}) \,=\, -0.62 \,+\, 1.22\,\tilde{z}_i \,.$$

Taking account of the integer values of z_i, this implies that those with friends with the smallest possible average of smoking behavior $\tilde{z}_i = 1$ are drawn towards the 'preferred' value of 1 themselves, while those having friends who have the highest possible average $\tilde{z}_i = 5$ are themselves also drawn toward this value 5 as a 'preferred' value. It can be concluded that the data provide support for the existence, in the co-evolution of friendship and drinking tendencies, of selection (tendency to choose friends with similar behavior) as well as influence (tendency to change behavior in the direction of friends' behavior).

VIII. Extensions

The basic model specifications defined above can be extended in various ways. Above we already mentioned the possibility to let the rates of change depend on covariates or on current network structure. Another possibility is to introduce an asymmetry between the values of ties when they are formed and their values when they are lost. E.g., for friendship dynamics, there is theoretical and empirical evidence that the additional 'value' of a tie added by its being reciprocated is higher when considering a potential loss of the tie than when considering the potential new formation of the tie. This asymmetry can be modeled by endowment functions, see [49]. Technically, this means that the effects $s_{ki}(x^0, x, v, w)$ used in (6) depend not only on the new state x but, unlike the examples given above, also on the preceding state x^0. For example, the endowment effect of a tie being reciprocal is expressed by

$$s_{ki} = \sum_j x^0_{ij}\, x_{ij}\, x_{ji} \ ,$$

which is sensitive to the number of i's reciprocal ties only when they are candidates to being terminated and not when they are candidates to being created.

Similarly, going upward on a behavior variable might be not the opposite of going downward, which can be modeled by endowment effects in the objective function for behavior.

IX. Future Directions

Although the statistical modeling of network dynamics started already with [22] and [59], this area has been in rapid development only since recent years. Much work remains to be done, however, to extend these methods to other types of network data and to study their properties. The lists of effects that can be included in the objective functions illustrate the flexibility of this model and its adaptability to research questions and network data. The availability of methods for analysis of network panel data has been a stimulus also for the further collection of such data.

Plausible models and good methods for parameter estimation and testing have now been developed, as summarized in this chapter, and they are available in the *SIENA* (*Simulation Investigation for Empirical Network Analysis*) program. This program is available as freeware with additional material on the website `http://www.stats.ox.ac.uk/~snijders/siena` and has an extensive manual [50]. The examples presented here were analyzed using this program.

The tests used in the examples, based on studentized parameter estimates, can be regarded as Wald-type tests. Some limited simulation studies have supported the validity of these tests. Score-type tests associated to the Method of Moments estimators were developed in [43]. The examples in this paper underline the need for methods to assess fit of models, and also to compare non-nested models. This could be done formally based on estimated likelihoods, or informally based on the comparison of observed and expected values of relevant statistics that are not used for parameter estimation.

The open question of assessing fit also invites speculation about the robustness of the results against the use of models of which the fit is not beyond doubt. There is an inherent tension between the complexity of processes of network dynamics, and the limited amount of data that can in practice be observed concerning these processes. One issue is that the models proposed here are Markov processes. For two-wave data sets there are no clear alternatives to making such an assumption, but the assumption is certainly debatable. Including more information in the state space (by using covariates, by considering valued rather than dichotomous ties, etc.) may relax the doubts concerning such an assumption.

Another issue is the difference between the tie-oriented and actor-oriented models. Which type of model is to be preferred is a matter both of social science theory and of empirical fit. It will be important to know, supported by simulation studies and/or mathematical results, the extent to which results based on particular models for network dynamics are robust to deviations from the precise assumptions made. In addition, it will be useful to develop still other models, e.g., models accounting for actor heterogeneity (like were developed for non-longitudinal network data, e.g., by [35, 20, 58]) or measurement error.

Other open questions are about mathematical properties of the estimators and tests proposed. Simulation studies support the conjecture that the Method of Moments estimators have asymptotically normal distributions, but this has not been proven. It is unknown if the solution to the moment equation (7), under certain conditions, is unique. Similar questions can be asked about the Maximum Likelihood estimators. All this indicates that there is ample scope for future work on methods of statistical inference for network dynamics.

References

[1] Albert, R., Barabási, A.-L. (2002) Statistical mechanics of complex networks, *Reviews of Modern Physics*, 74, 47–97.

[2] Bartholomew, D.J. (1982) *Stochastic Models for Social Processes* (3rd ed.) New York: Wiley.

[3] Bearman, P.S. (1997) Generalized exchange, *American Journal of Sociology*, 102, 1383–1415.

[4] Burt, R.S. (1992) *Structural Holes*. Cambridge, MA: Harvard University Press.

[5] Brass, D.J., Galaskiewicz, J., Greve, H.R., Tsai, W. (2004) Taking stock of networks and organizations: a multilevel perspective, *Academy of Management Journal*, 47, 795–817.

[6] Burk, W.J., Steglich, C.E.G., and Snijders, T.A.B. (2007) Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. *International Journal of Behavioral Development*, 31, 397–404.

[7] Cartwright, D., and F. Harary. 1956. Structural Balance: A Generalization of Heiders Theory. *Psychological Review*, 63, 277–292.

[8] Checkley, M., and Steglich, C.E.G. (2007) Partners in Power: Job Mobility and Dynamic Deal-Making. *European Management Review* 4, 161–171.

[9] Chen, H.-F. (2002) *Stochastic Approximation and its Applications*. Dordrecht etc.: Kluwer.

[10] Coleman, J.S. (1964) *Introduction to Mathematical Sociology*. New York etc.: The Free Press of Glencoe.

[11] Coleman, J.S. (1990) *Foundations of Social Theory*. Cambridge/London: Belknap Press of Harvard University Press.

[12] de Federico de la Rua, A. (2003) La dinamica de las redes de amistad. La eleccion de amigos en el programa Erasmus. *REDES* 4.3, `http://revista-redes.rediris.es`

[13] de Nooy, W. (2002) The dynamics of artistic prestige. *Poetics*, 30, 147–167.

[14] Ennett, Susan T., and Karl E. Bauman. 1994. The Contribution of Influence and Selection to Adolescent Peer Group Homogeneity: The Case of Adolescent Cigarette Smoking. *Journal of Personality and Social Psychology* 67, 653–63.

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

[16] Friedkin, N.E. (1998) *A Structural Theory of Social Influence*. Cambridge: Cambridge University Press.

[17] Granovetter, M.S. (1973) The strength of weak ties. *American Journal of Sociology* 78, 1360–1380.

[18] Gulati, R., Nohria, N., and Zaheer, A. (2000) Strategic Networks. *Strategic Management Journal*, 21, 203–215.

[19] Haynie, Dana L. 2001. Delinquent Peers Revisited: Does Network Structure Matter? *American Journal of Sociology* 106, 1013–57.

[20] Hoff, P.D., Raftery, A.E., and Handcock, M.S. (2002) Latent Space Approaches to Social Network Analysis, *Journal of the American Statistical Association*, 97, 1090–1098.

[21] Holland, P.W., Leinhardt, S. (1975) Local structure in social networks, *Sociological Methodology - 1976*, 1–45.

[22] Holland, P.W., Leinhardt, S. (1977) A dynamic model for social networks, *Journal of Mathematical Sociology*, 5, 5–20.

[23] Hosmer, D., Lemeshow S. (2000) *Applied logistic regression* (2nd ed). New York: Wiley-Interscience.

[24] Koskinen, J.H., Snijders, T.A.B. (2007) Bayesian inference for dynamic social network data, *Journal of Statistical Planning and Inference*, 137, 3930–3938.

[25] Kossinets, G., and Watts, D.J. (2006) Empirical Analysis of an Evolving Social Network, *Science* 311, 88–90.

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

[27] Lévi-Strauss, Claude (1969; orig. 1947) *The Elementary Structures of Kinship*, Boston: Beacon Press.

[28] Lin, N., Cook, K., Burt, R.S. (Eds.) (2001) *Social Capital. Theory and Research*, New York: Aldine de Gruyter.

[29] Long, J.S. (1997) *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications.

[30] Maddala, G.S. (1983) *Limited-dependent and Qualitative Variables in Econometrics* (third ed.), Cambridge: Cambridge University Press.

[31] Newcomb, Theodore M. (1962) Student Peer-Group Influence. In Nevitt Sanford (Ed.): *The American College : A Psychological and Social Interpretation of the Higher Learning*. New York, Wiley.

[32] Newman, M.E.J., Barkema, G.T. (1999) *Monte Carlo methods in Statistical Physics*, Oxford: Clarendon Press.

[33] Newman, M. E. J., Watts, D. J., Strogatz, S. H. (2002) Random graph models of social networks, *Proceedings of the National Academy of Sciences USA*, 99, 2566–2572.

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

[35] Nowicki, K., Snijders, T.A.B. (2001) Estimation and prediction for stochastic blockstructures, *Journal of the American Statistical Association*, 96, 1077–1087.

[36] Pearson, M.A., and Michell, L. (2000) Smoke Rings: Social network analysis of friendship groups, smoking and drug-taking. *Drugs: education, prevention and policy*, 7, 21–37.

[37] Pearson, Michael, Steglich, Christian, and Snijders, Tom (2006) Homophily and assimilation among sport-active adolescent substance users. *Connections* 27(1), 47–63.

[38] Pearson, M., and West, P. (2003) Drifting Smoke Rings: Social Network Analysis and Markov Processes in a Longitudinal Study of Friendship Groups and Risk-Taking. *Connections*, 25(2), 59–76.

[39] Podolny, J.M., Stuart, T.E. and Hannan M.T. (1997) Networks, knowledge, and niches: competition in the worldwide semiconductor industry, 1984–1991. *American Journal of Sociology*, 102, 659–689.

[40] Price, D. de Solla (1976) A general theory of bibliometric and other advantage processes. *Journal of the American Society for Information Science* 27, 292–306.

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

[42] Robins, G.L., Woolcock, J., Pattison, P. (2005) Small and other worlds: Global network structures from local processes, *American Journal of Sociology*, 110, 894–936.

[43] Schweinberger, M. (2007) Statistical modeling of digraph panel data: Goodness-of-fit, Submitted for publication.

[44] Schweinberger, M., Snijders, T.A.B., (2006) Markov models for digraph panel data: Monte Carlo-based derivative estimation, *Computational Statistics and Data Analysis*, 51, 4465–4483.

[45] Snijders, T.A.B. (1996) Stochastic actor-oriented dynamic network analysis, *Journal of Mathematical Sociology*, 21, 149–172.

[46] Snijders, T.A.B. (2001) The statistical evaluation of social network dynamics, in: Sobel, M., Becker, M. (Eds.), *Sociological Methodology*, Boston and London: Basil Blackwell, 361–395.

[47] Snijders, T.A.B. (2005) Models for longitudinal network data, in: Carrington, P.J., Scott, J., Wasserman, S. (Eds.), *Models and Methods in Social Network Analysis*. New York: Cambridge University Press.

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

[49] Snijders, Tom A.B., Steglich, Christian E.G., and Schweinberger, Michael. 2007. Modeling the co-evolution of networks and behavior. In *Longitudinal models in the behavioral and related sciences*, edited by Kees van Montfort, Han Oud and Albert Satorra, pp. 41–71. Mahwah, NJ: Lawrence Erlbaum.

[50] Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M., Huisman, M. (2007) *Manual for SIENA version 3*, Groningen: ICS, University of Groningen; Oxford: Department of Statistics, University of Oxford, `http://www.stats.ox.ac.uk/~snijders/siena`.

[51] Sparrowe, R.T., Liden, R.C., Wayne, S.J., Kraimer, M.L. (2001) Social Networks and the performance of individuals and groups. *Academy of Management Journal*, 44, 316–325.

[52] Steglich, C.E.G., Snijders, T.A.B., Pearson, M. (2010) Dynamic networks and behavior: Separating selection from influence, *Sociological Methodology*, 40, 329–393.

[53] Steglich, Ch.E.G., Snijders, T.A.B., and West, P. (2006) Applying SIENA: An Illustrative Analysis of the Coevolution of Adolescents' Friendship Networks, Taste in Music, and Alcohol Consumption. *Methodology*, 2, 48–56.

[54] Udehn, L. (2002) The changing face of methodological individualism, *Annual Review of Sociology*, 8, 479–507.

[55] van de Bunt, G.G., van Duijn, M.A.J., and Snijders, T.A.B. (1999) Friendship networks through time: An actor-oriented statistical network model. *Computational and Mathematical Organization Theory,* 5, 167–192.

[56] van de Bunt, G.G., Wittek, R.P.M., and de Klepper M.C. (2005) The Evolution of Intra-Organizational Trust Networks; The Case of a German Paper Factory: An Empirical Test of Six Trust Mechanisms. *International Sociology* 20, 339–369.

[57] van Duijn, M.A.J., Zeggelink, E.P.H., Huisman, M., Stokman, F.N., and Wasseur, F.W. (2003) Evolution of Sociology Freshmen into a Friendship Network. *Journal of Mathematical Sociology* 27, 153–191.

[58] van Duijn, M.A.J., Snijders, T.A.B., Zijlstra, B.H. (2004) p_2 : a random effects model with covariates for directed graphs, *Statistica Neerlandica* 58, 234–254.

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

[60] Wippler, R. (1978) The structural-individualistic approach in Dutch sociology, *Netherlands Journal of Sociology*, 4, 135–55.

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

FURTHER READING

For further reading, basic concepts of continuous-time Markov processes can be found in [34]. The basic definition of the model presented here and of the statistical estimation methods for network dynamics based on panel data can be studied in [46, 47]. The approach to the co-evolution of networks and behavior is presented in [49, 52]. Some examples of the methods presented in this chapter can be found in [6, 53, 55, 56, 57].