Networks of action and events over time

Analytic designs for continuous-time longitudinal network data

Paper presented at the Sunbelt XXVIII International Social Network Conference, January 22-27, 2008, St. Petersburg, Florida.

Wouter de Nooy
Communication Science, University of Amsterdam
w.denooy@uva.nl

Abstract

Current models for the dynamics of social networks focus on repeated measurements of a network. They seek to predict or explain the appearance, maintenance, and disappearance of ties between consecutive measurements of a network both from previous network structure and properties of actors in the network. These models presuppose that (1) the measured relation is manifest and potentially enduring, and (2) any actor is able and likely to consider changing its tie to any other actor in the network. Friendship within a set of interacting people is an example of a social relation meeting these criteria.

Lots of longitudinal network data, however, do not meet these criteria, especially data on events or actions from one actor towards another. For instance, the person who is going to throw a birthday party is the only person who can invite other people. Turns in group discussions are usually short but the order of turns is likely to be driven by previous contributions to the discussion. For actions and events that are incidental rather than enduring, analytic designs developed for panel data are not appropriate. This paper proposes to use event history models, especially a discrete-time model using multilevel regression analysis for taking advantage of the fact that we know who acted towards whom at what moment.

Introduction

In social network analysis, little attention has been paid to the analysis of network data measured in continuous time, that is, chronologically ordered series of lines among vertices. Initially, this kind of data was not available (Katz and Proctor 1959; Runger and Wasserman 1979: 144). The focus was on enduring social relations such as friendship within a social group and data was collected by means of repeatedly administered surveys. As a consequence, models for longitudinal network data have been and are still predominantly being developed for panel wave data, including recent advances such as the MCMC approach to longitudinal data implemented in SIENA (Snijders 2005).

It is, however, possible to construct continuous-time networks from archives, documents, and observations (Burt and Lin 1977). The digitization of archives and documents greatly simplifies this process. As a consequence, continuous-time network data are readily available: networks of actions or events. It is time to develop analytic designs for this type of data. As I will argue in this paper, event history models can be used and in some cases well-known regression models are applicable provided that the data is organized is a special way and a multilevel design is used. The general multilevel regression design is developed in Section 2. Time is incorporated in two ways: as an exogenous factor (Section 3) and as an endogenous factor to be explained within the model (Section 4).

Conceptualizing longitudinal networks as series of instantaneous actions and events rather than more or less enduring social ties, which is implicated in panel designs, raises questions about the nature of social relations. In the case of informal enduring social ties, are we studying social relations or perceptions? What is the link between the two? These questions will be dealt with in the concluding section (Section 5). Let us first summarize the ways in which time has been incorporated in social network analysis until now.

1. A very brief history of time in social network analysis

Time has been incorporated in different ways in social network analysis. The change of network structure over time has received most attention, e.g., in special issues of the *Journal of Mathematical Sociology* – issue 1-2 of volume 21 (1996) and issue 2-3 of volume 27 (2003) – and edited books (Weesie and Flap 1990). However, this is not the only way time may play a role in social network analysis. I distinguish between three broadly defined practices.

The first and oldest approach uses arcs to represent time order. Genealogies are the prime example representing arcs pointing from parents to children or in the reverse direction. Analyzing genealogies in this way has been common in anthropology for a long time and probably constitutes the oldest type of SNA (Freeman 2004). Another application of this principle is found in citation networks representing publications (vertices) citing (arcs) previous publications. In social network analysis, Hummon and Doreian (Doreian 1990; Hummon and Carley 1993; Hummon and Doreian 1989; Hummon et al. 1990) developed special techniques for analyzing this type of data. The concept of the Spell Precedence Graph for life histories also belongs to this category (Butts and Pixley 2004).

The second approach uses a static network while vertex' attributes change over time. In social network analysis, this manner of including time in the analysis is probably introduced by diffusion studies, which tried to predict the adoption of an innovation or a disease from a person or organization's position in a social network. Through network ties, one is exposed to the influence of peers, so this is typically an example of the social influence model (Friedkin and Johnsen 1990; Robins et al. 2001b): (how) does network structure and attributes of network neighbours affect vertex' attributes?

In the course of time, event history models have increasingly been applied to this type of data because they adequately deal with the passing of time before adoption or another event takes place. Marsden and Podolny introduced event history designs for social influence models in the social network analysis community (Marsden and Podolny 1990). Inside and outside this community event history models using characteristics of a person's or organization's network position as a predictor for an action or event are widely used although it should be noted that network position sometimes merely measures whether one is married or the number of children one has. The impact of the size and composition of a person's social network on health or mortality is a popular topic (Adams et al. 2002; Bygren et al. 1996; Gustafsson et al. 1998; Kang and Bloom 1993; Litwin and Shiovitz-Ezra 2006; Patterson et al. 1996; Payette et al. 2000; Samuelsson and Dehlin 1993; Trovato 1998; Villingshoj et al. 2006) and so is the impact of exposure to adopters of an innovation or practice through social or geographical network ties (Bogart 2007; Bohman 2006; Chaves 1996; Davis and Stout 1992; Edling and Sandell 2001; Lipp and Krempel 2001; McKeown 1994; Mintrom and Vergari 1998; Soule 1997; Strang and Tuma 1993; Van den Bulte and Lillien 2001). The relevance of network ties for the 'hazard' of finding a job has also been studied (Bernasco et al. 1998; Brandt 2006; Yakubovich and Kozina 2000) as well as the hazard of leaving a field or organization (Clarysse et al. 1996; McPherson et al. 1992; Mossholder et al. 2005; Sutton and Chaves 2004).

Note that these studies usually have static networks for practical reasons rather than substantive arguments; collecting network data is costly. Indeed, event history models do allow for time-varying measures of network position, as I will discuss extensively in Section 4.

The third approach to time uses dynamic network data: one or more relations on a set of vertices measured at different moments. The goal here usually is the investigation of social selection processes (Robins et al. 2001a), explaining changes in network structure or network emergence rather than effects of network structure on attributes of the vertices. In social network analysis, the earliest reported example is the article by Katz and Proctor (Katz and Proctor 1959) introducing some features that characterize a lot of work on longitudinal networks done afterwards. First, they focused attention to processes at the dyadic level assuming that network change should be studied as change in (ordered) pairs rather than change at the level of the entire network. This still is the dominant perspective. Second, they introduced Markov Chain models for repeated measurements of network relations, analyzing the transition probabilities between different types of ties: mutual, asymmetric, null.

Later developments for discrete-time Markov Chain models (Robins and Pattison 2001; Runger and Wasserman 1979) and continuous-time Markov Chain models (Holland and Leinhardt 1977; Snijders 1995) gradually extended the options for testing the effects of covariates on the formation and dissolution of ties. Characteristics of vertices, such as gender, properties of pairs, e.g. (dis)similarities on social characteristics, and local network structure, e.g., the amount of reciprocity or transitivity created by a tie, can now be incorporated in a model to explain the maintenance or breaking of a tie. The latest models even allow for a simultaneous estimation of network change (social selection model) and change in vertex' attributes (social influence model), which are successfully applied to empirical networks of friendship relations in combination with, for instance, smoking or (other) substance abuse (Pearson et al. 2006; Snijders et al. 2007).

Current models for repeated measurements or panel wave data explain network change from previous network structure and characteristics of the actors. Thus, a wide range of theories can be tested. So why not aggregate continuous-time network data into panel waves? The usual arguments against doing that apply here (Blossfeld et al. 1989:14). First of all, aggregating data discards a lot of information on the temporal order of events. For action and events that evolve rapidly, e.g., a flame on a discussion board, the timing of events must be preserved. In addition, the researcher must make a priori decisions on the length of the period to be aggregated into one wave. This decision may have severe consequences for the results. Finally, many types of action or events take place irregularly, so it is not necessarily the case that there are observations for all cases during a fixed period. It is for these reasons that I now explore and present analytic designs for actions that are measured in continuous time.

2. *A multilevel regression approach to event or action data in continuous time*

This section presents the basic multilevel regression model that is central to this paper. I restrict my discussion to the social selection model, predicting the presence and/or characteristics of lines in the network. As a consequence, the case or unit of analysis is the ordered pair of vertices (*f*, *g*). Throughout, I will assume that the network is directed, that is, a distinction can be made between the (potential) actor or tail (*f*) and the one acted upon or head (*g*). The argument, however, can be applied easily to undirected networks.

The dependent variable is a characteristic of the ordered pair (*f*, *g*) at time *t*, denoted by *Y(fgt)*. Usually it is the presence or absence of an arc from the first to the second member of the pair. In that case, *Y(fgt)* is a binomially distributed dichotomy with presence and absence of an arc coded as 1 and 0 (Equation 1). Logistic regression analysis then predicts (the log odds of) the presence of an arc. If the arc can take one of several categories, e.g., be positive, negative, or absent, a multinomial regression analysis must be used. If it can take a value on a scale, the standard regression model can be used. However, the standard regression approach assumes that the absence of an arc (*Y(fgt)* = 0) is hardly different from an arc with a very low value (with very detailed measurement, the exact value 0 may even never occur), which may not be a very realistic assumption in SNA. In this paper, I assume *Y(fgt)* to be a dichotomy indicating the presence versus absence of an arc within an ordered pair of vertices, unless stated differently.

$$Y(fgt)_{ij(kl)} \sim \text{Binomial}(\text{denom}_{ij(kl)}, \pi_{ij(kl)}) \qquad (1)$$

$$\text{logit}_{ij(kl)} = \beta_{0j(kl)}\,\text{constant} + \beta_1 X(fgt)_{ij(kl)} + \beta_2 X(fg)_{j(kl)} + \beta_3 X(f)_k + \beta_3 X(g)_l \qquad (2)$$

$$\beta_{0j(kl)} = \beta_0 + v_{0l} + v_{0k} + u_{0j(kl)} \qquad (3)$$

$$[w_{0l}] \sim N(0, \Omega_w) : \Omega_w = [\sigma^2_{w0}] \qquad (4)$$

$$[v_{0k}] \sim N(0, \Omega_v) : \Omega_v = [\sigma^2_{v0}] \qquad (5)$$

$$[u_{0\,j(kl)}] \sim \mathrm{N}\,(0, \Omega_u) : \Omega_v = [\sigma^2_{u\ 0}] \tag{6}$$
$$\mathrm{var}(Y(fgt)_{ij(kl)} \mid \pi_{\,ij(kl)}) = \pi_{\,ij(kl)}\,(1 - \pi_{\,ij(kl)}) \,/\, \mathrm{denom}_{\,ij(kl)} \tag{7}$$

Following the approach proposed by Wasserman and Pattison (Wasserman and Pattison 1996) and extending it to a multilevel model, the well-known structural effects in SNA can be included in the regression model in the following manner. Activity and popularity effects of attributes or structural properties of vertices are included in the model by simply adding the attributes of the tail or head as independent variables (*X(f)* and *X(g)* in Equation 2). Their effect parameters signal their contribution to (the log odds of) the presence of an arc.

These are covariates at the level of the vertex. Because each vertex is a member of several ordered pairs, the latter are nested within the former, so we must use a multilevel model with at least two levels: a level for ordered pairs (*j*) and higher levels for vertices (*k* and *l*). Note that subscripts denote levels in Equations 1 thru 7, which specify a variance components multilevel model with random effects only for the constant. Random slopes can be added, e.g., differential effects for attributes at a higher level, but I will not discuss that here. Vertices appear both as tail and head in the ordered pair, so the multilevel model is cross-classified, hence the parentheses around *k* and *l*. In a 1-mode network, vertices may and indeed do appear both as tail and head. If their role as a sender of an arc is related to their role as a receiver, a multiple membership or multiple roles model should be used, assuming that the random effects are correlated for senders and receivers that refer to the same person (Snijders and Bosker 1999: 161-162).

Homophily effects of attributes of vertices, e.g., sex homophily, are parameterized by including a dummy variable signalling whether the tail and head of the pair have the same value on an attribute (1) or not (0) ($X(fg)_{\,j(kl)}$ in Equation 2). This is a dyadic covariate, more precisely, a characteristic of the unordered pair. Strictly speaking, the ordered pairs are nested within the unordered pair (each unordered pair containing two ordered pairs), so one should add a level for unordered pairs between the level of ordered pairs and the higher level of vertices. In practice, however, it isn't to be expected that this extra level will account for a substantive amount of variance because there are very few observations for each unordered pair (two as the maximum) whereas there are many distinct unordered pairs. For the sake of simplicity, no extra level is specified for unordered pairs in Equations 1-7 although it is easy to add it. Homophily effects are specified at the level of ordered pairs (*j*).

Deference effects of vertex' attributes, e.g., a tendency for people to initiate ties with people from higher social classes, can be modelled by dummy variables taking the value 1 if the tail has a lower value on the attribute than the head and 0 otherwise. Obviously, deference or inequality effects require the vertex' attribute or structural property to be measured minimally at an ordinal scale and the relation to be asymmetric. Now, the independent variable truly is a characteristic of the ordered pair, so it is correctly specified at that level ($X(fg)_{\,j(kl)}$ in Equation 2).

In a multiple relations network, ties on the relation of interest may be predicted by the occurrence of ties on another (exogenous) relation, e.g., the action of seeking advice may be more likely to occur between people who are tied by a friendship relation. Effects of ties on other network relations are measured at the level of the ordered or unordered pair, so they are included at that level (*j* in Equations 1-7). We may refer to these as exogenous network effects.

In contrast, endogenous network effects refer to the effects of ties on the relation of interest that exist before the time of the action. In principle, both characteristics of overall network structure, e.g., network centralization, and properties of local network structure, that is, the structure of the immediate or extended network neighbourhood of the ordered pair, may be taken into account. I focus exclusively on local network structure because it is both theoretically and practically more likely that actors can and do take note of their network neighbourhood.

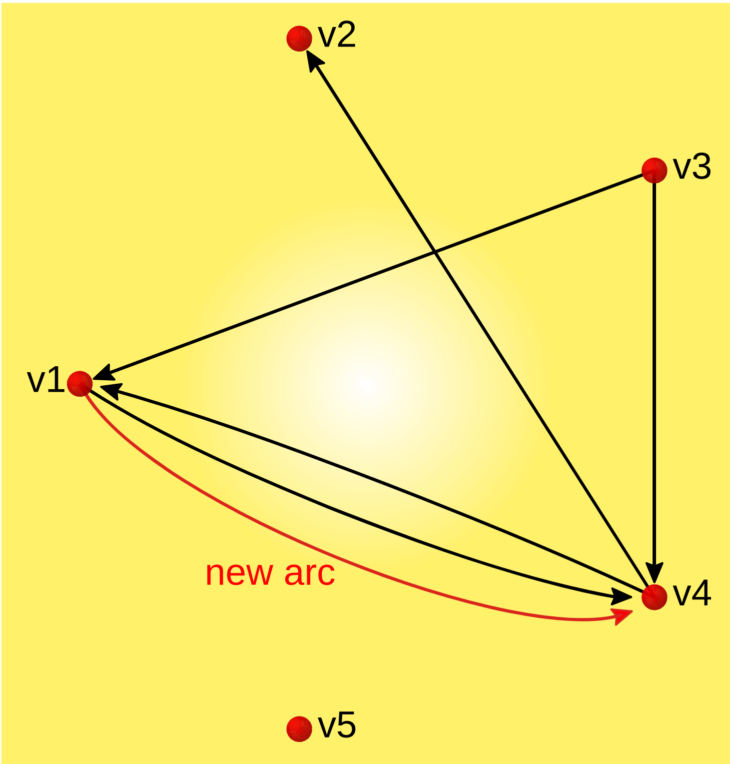

Figure 1 - A new arc (red) in its current network context.

Within the focal ordered pair, that is, the pair for which we want to explain the presence or properties of a line, two endogenous network effects can be distinguished: conformity and reciprocity. Conformity refers to the tendency to replicate previous choices; previous actions from the tail to the head raise the probability of new action. Reciprocity refers to the opposite situation: previous action from the head of the focal ordered pair towards the tail raises the probability of an action from the tail to the head. Figure 1 shows the structure of arcs established within a small network and the red arc represents a new action undertaken by vertex v1 towards vertex v4 within the focal ordered pair (v1, v4). Because there are previous (black) arcs from v1 to v2 and from v4 to v1, the new arc is predicted both by conformity and reciprocity. In the multilevel logistic regression model, parameter estimates for variables signalling the presence of previous arcs in the same or opposite direction respectively test these effects. The variables are included at the level of the (un)ordered pairs (*j*). It is up to the researcher to decide whether the variables are dummies, just indicating whether or not there are previous arcs, or a count variable indicating the number of previous arcs.

Note that we can distinguish between previous and current arcs in continuous time data. As a consequence, we do not need the unordered pair as the level of analysis, which must be done with analytic designs for cross-sectional data such as the p2 model and the multilevel approach described in (Snijders and Bosker 1999; Snijders and Kenny 1999). I return to the topic of dependencies among observations below.

The conformity and reciprocity effects exemplify the general principle of including effects of previous local network structure on the formation or contents of ties. We define the theoretically expected local configurations and assess whether or count the number of these configurations that would be established by an arc (of a particular type) within the focal pair (f, g). This count or dummy variable ($X(fg)_{j(kl)}$ in Equation 2) is included at the level of the (un)ordered pair (j). If theory, for instance, predicts that people establish those ties that maximize the size of their 2-neighbourhood (direct and indirect contacts), the independent variable expresses the size of the 2-neighbourhood of the tail if it would establish an arc to the head. The higher this number, the more likely this tie will be established and the stronger this tendency, the higher the parameter estimate of the variable expressing the local structural tendency.

Alternatively, the independent variable could measure the preponderance of theoretically expected local configurations over theoretically unexpected local configurations created by an arc. The tendency towards transitivity, for instance, could be measured by the preponderance of transitive triples over intransitive triples created by an arc. The covariate expressing this tendency would simply show the difference between the number of transitive triples and the number of intransitive triples created by an arc.

In addition to the presence or absence of an arc, a characteristic of the arc can also be used. Balanced semicycles offer a well-known example (Cartwright and Harary 1956). In a signed directed network, a tendency towards balance is captured by a variable showing the preponderance of balanced semicycles created by a positive arc over the number of balanced semicycles created by a negative arc. Alternative specifications are possible, e.g., counting the preponderance of balanced semicycles over unbalanced semicycles created by a positive arc.

A wide range of structural effects can be modelled taking into account ties outside the (focal) ordered pair in the way just described. Note, however, that structural effects that extend beyond the ties in which the actor (tail) is directly involved, require that ties between its neighbours (or their neighbours…) are visible to the actor. Ties must be manifest or broadcasted (Friedkin and Johnsen 1990) in one way or another.

Let us now turn to the lowest level in the multilevel model (i in Equations 1, 2 and 7), which I haven't discussed yet. This level is reserved for repeated observations on the ordered pair. With network data measured in continuous-time, it is very likely that ordered pairs do or may act at more than one time point. Thus, we have repeated measurements within each ordered pair. Multilevel models accommodate for this by introducing a level nested under the ordered pairs level. Dyadic covariates for an ordered pair that change between observations of (potential) action – time-varying covariates $X(fgt)_{ij(kl)}$ – must be entered at the level of the action instead of the level of the ordered pair. This will usually be the case for covariates expressing local network context in which we are primarily interested. Therefore, we are more likely to treat multiple observations within an ordered pair as just different observations rather than a sequence of observations that may display an interesting trend in time as in many designs for longitudinal data (Snijders and Bosker 1999: 181).

The well-known problem of (logistic) regression analysis of cross-sectional network data is the fact that observations are not independent (Wasserman and Robins 2005). If there is a complete triad in a directed network, we are bound to encounter six ordered pairs that are part of a complete triad. One could say that the dependent and independent variables are hopelessly confounded: the complete triad explains the presence of the arcs because they yield high transitivity and the complete triad exists thanks to the arcs. It is no use pretending that the six observations are collected through independent sampling. As a consequence, the estimation procedures and standard errors of regression analysis are unreliable.

With continuous time data, the time order between actions (arcs) removes the circularity between local configurations and the presence or characteristics of arcs. We can exactly identify the (local) structure of the network as it is just before a new arc appears. Each new arc can be treated as a new observation. It is analyzed in a local context that is unique as it is changed afterwards by the new arc itself. With a suitable sampling frame, one could imagine analyzing a random sample of ordered pairs, collecting data on the network neighbourhood at the time the arc appears to construct the relevant independent variables. Of course, such a sampling frame will usually not be available and the collection of data on the local context is time consuming, so we usually prefer studying the entire network. Nevertheless, the possibility of random sampling illustrates the idea of independence.

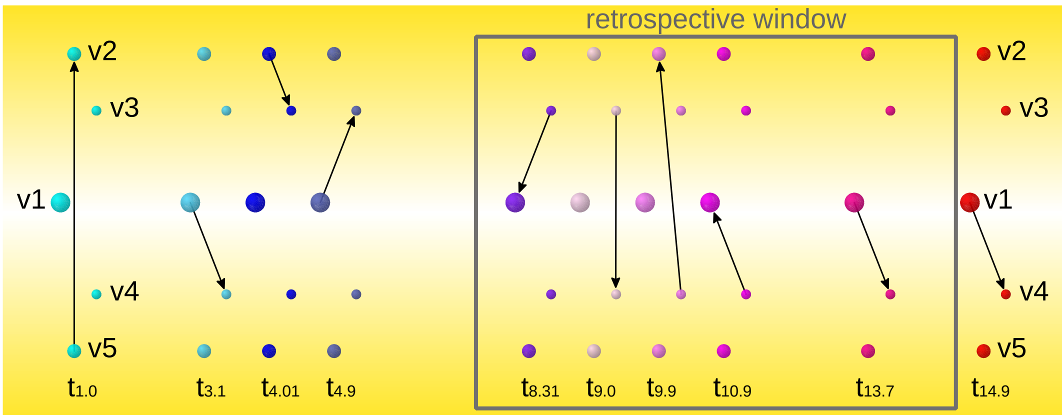

Figure 2 - A network in measured in continuous time and a retrospective window.

As long as we distinguish clearly between the new arc to be explained and local network structure as it existed just before the new arc appears, we may treat the observations as independent and apply the regression model. Then how do we delimit existing network structure? Some choices must be made.

First, the size of the local network must be chosen. This is a substantive choice: what do we expect the actor to know about the network? It is very likely that an actor knows its proper actions and the actions directed towards it. Perhaps, it is also aware of other actions in which all actors were involved. Extending the local network context beyond the actor’s neighbours’ neighbours (the 2-neighbourhood), however, requires very much operational knowledge of the network among the actors. Moreover, the number of theoretically expected local configurations usually explodes when extending the local network context beyond this realm. The 2-neighbourhood of the focal pair seems to be a reasonable choice but the investigator may have to experiment to find the optimal local network size in an empirical application.

The second choice concerns the issue of how far back in time we include actions in local network structure. Note that the nature of continuous-time data requires that we make assumptions about how actors mentally construct the structural context from actions that occur at specific moments in time. There is no first order Markov Chain assumption possible as in repeated measurements designs. If actors never forget, we should have complete data from the entrance of an actor to the social system that we are studying. In all but exceptional cases, this is unlikely to be the case in empirical applications. If actors do forget, we can use a retrospective window and only include the actions that occurred within a fixed period of time before the (potential) action that we want to explain. This is the only feasible option in most empirical applications. It is up to the investigator to experiment with and decide on the length (l) of the retrospective period. Note that it is only possible to calculate the independent variables for structural effects when we have data on the entire retrospective period, that is, data that happen more than l later than the start of the period that we are studying. The longer the retrospective window, the longer the period that we cannot study.

The third choice: do we weigh arcs according to the time passed? If we do, we must specify a decay function. With a decay function, the choice of the length of the retrospective window (l) becomes less important because actions that happened longer ago weigh less. We must, however, make additional assumptions about how we reckon with the weights when assessing local configurations, e.g., do we average the arc weights when we find a transitive triple or do we take the minimum arc weight?

The fourth and last choice: do we take into account multiple arcs or do we only consider the last arc? Multiple lines are no problem for counting theoretically expected local configurations but the numbers can go up quickly if we take into account each combination of multiple lines. Substantively, one may argue that humans are likely to take the last action as the one that is most salient, but again, this is just speculation and particular types of social relations may be modelled better using all previous actions.

All in all, a multilevel regression model seems to be fit for analyzing network change when applied to data measured in continuous time. The major problem of dependence among observations is solved on the one hand by the strict temporal distinction between the characteristic of the ordered pair that is to be explained and local network structure explaining it, and on the other hand by using a multilevel model accounting for repeated observations within pairs and nesting of pairs within vertices. However, due to the fact that we are now dealing with single actions occurring at a specific moment in time rather than an enduring tie or attitude, which we can elicit or measure at any moment, new questions arise, which are related to the role of events (in history) as exogenous occasions at which action may, can not, or must happen. Can anybody act at any time towards any other actor? This is the topic of the next section.

3. *Exogenous timing: who acts towards whom in which way?*

Should we predict the time at which an action takes place or should we predict which action is taken if action must be taken? Events featuring action may arise from the network dynamics studied or they may be caused by factors outside this process. Timing of action results from network structure itself, for instance, if an unfriendly act provokes an unfriendly reaction. In contrast, inviting people for a birthday party is triggered by an event related to the date of birth of a person. It is not possible to predict the timing of this event from our network model. Timing of the event and the action is exogenous to the model. At the same time, the person who is going to act (decide invitations) is exogenously determined in this example. In this Section, I discuss the design for analyzing continuous-time network data if the timing of action is exogenous to the model. In doing so, I will dwell upon the ways in which we can handle restrictions on who can act at the given moment. In the next section, I will discuss modelling endogenous timing using event history models (Section 4). In both cases, we will see that it is mainly a matter of selecting or constructing the right cases in the data set.

What are the cases in our design? The previous section stated that ordered pairs at particular moments are the cases but which ordered pairs do we take into consideration at what times? Applying the logistic regression model to cross-sectional network data, Wasserman and Pattison selected all ordered pairs in the network (Wasserman and Pattison 1996). For each ordered pair, a line may be present or absent. They did not have to worry about time because they dealt only with one time point. With longitudinal data measured in continuous time, however, we need to decide about time.

In the case of exogenous timing, the time problem is easily resolved. Because we know at which time action can or must be taken, it is natural to create cases for each ordered pair at each moment at which action may take place within this ordered pair. Because we know the moments of potential action, we can take them as our observation times and merely look back in time to construct the values of the independent variables at that moment. This is a purely retrospective approach: action has/could have happened at this moment, so what was the situation leading to the action?

In the least constrained case, any vertex may act towards any other vertex at the moment of action, say, at the occurrence of a particular event. Here, the analyst creates a case for each ordered pair at each time point at which action is or can be taken, regressing the (log odds of the) occurrence of an arc on the relevant independent variables. It is, however, very difficult to find a good example for this situation. If we think of an event at which all people can be expected to act, e.g., exchanging best wishes on New Year's Day, people do not act simultaneously. At a more precise temporal level of measurement, we will see a 'spontaneous' time ordering which is probably best understood from the network dynamics itself (Section 4). Some people may be expected to take the initiative for establishing a contact, thus precluding the contact in the reverse direction. Similarly, instructing pupils to pair up for group work in class does obliges all pupils to take action but we'll see that first some pairs form, then some more, and so on. Because the pairs that form later depend on prior pairs in the trivial sense that the members of prior pairs are no longer eligible as partners, we actually have a form of endogenous timing which is suited for event history analysis (Section 4).

Not being able to come up with a clear example of exogenously giving time at which all pairs act simultaneously, it is my conjecture that exogenous timing of action always goes together with restrictions on who can act. Invitations for a birthday party offer an example of an event constraining who can act: it is the soon to be birthday person who must invite other people. In this example, we only need to construct cases for $(n - 1)$

ordered pairs in which the tail is fixed to the person whose birthday it is going to be. If our data consists of invitations to birthday parties among a group of people, we have a series of events (birthday parties) with (n – 1) ordered pairs for each event. This clearly is an example in which the event fixes both the timing and the tail. It is probably relevant in many empirical social networks, e.g., social contacts after break up of a relationship; to whom do you turn for help in case of an emergency, etcetera.

Examples of events fixing timing and the head are available as well. Harrison White's well-known data on vacancy chains (White 1970), for instance, may be conceptualized as a 2-mode network of (m) ministers and (n) congregations. External events such as retirement and internal job changes primarily dictate where and when job vacancies appear. In principle, all ministers may apply for each new job (unless there are formal restrictions specified with a vacancy). Job application and its success may be analyzed with a regression model taking all (m) ordered pairs of ministers and the particular vacancy as its units of analysis. Each job vacancy in the studied period adds a number of cases to the data set that is equal to the number of ministers available at that moment.

Finally, the event may constrain the timing of action, who acts and whom is acted upon, e.g., if social norms or institutional practices dictate action by particular persons on specific events. Examples are birthday congratulations among next of kin in several cultures, coverage of major news events by tv networks and newspapers, and reviews of newly released major Hollywood movies in magazines or papers. Note that we have a similar situation if we want to predict the contents of a choice instead of the timing of a choice, e.g., once you decide to buy something new, which brand or make are you going to buy (Taris 2000)?

If the event constrains timing, tail and head of the action, we have observations for the ordered pairs containing the fixed tails and head(s) at the specified time. Because there must be action, we cannot model the presence versus absence of action. We have to take the type of action as our dependent variable, e.g., is the birthday congratulation expressed in a visit, postal delivery of flowers or a present, a post card, a telephone call, or an e-mail? Here, the structure of the underlying graph is not explained but network structure, defined as the graph and additional information such as the valence of lines, is modelled, so I believe that we are doing social network analysis even in this case.

It is important to note that our thinking about events outside the model triggering or facilitating action has made us aware of constraints on who may act or whom may be acted upon. Investigating concrete acts over time, we realize that not everybody may act at any time. This is a consequence of moving away from enduring ties to action as incidents in time. It suggests that the difference between a panel design and a continuous-time design is not just a technical matter but also a substantive one. I'll return to this in the Conclusion (Section 5) but let us first consider the situation in which the timing of events is not exogenously fixed but is free to vary, e.g., depending on the development of network structure or vertex' attributes.

4. Endogenous timing: (when) do events happen?

Let us return to a previous example: New Year's visits. Each visit reduces the set of visits that can happen afterwards. This yields a special dependency between observations. In addition, visits may still happen after the termination of the data observation, e.g., New

Year's visits that are delayed until January 2 or later. For these pairs, we have partial information: we know that they haven't visited yet at the end of January 1, but we don't know whether a visit will take place and if so, at what time. Both aspects are consequential to the estimation of parameters.

Event history models provide a solution for both problems by modelling the hazard function, that is, the hazard that an event will occur to an actor as a function of time passed since the actor's entrance in the system or since the previous event. Time dependencies among observations are taken care for by the hazard function and observations for which the event has not yet happened are marked as right-censored, implying that they are part of the risk set during the observation period.

Event history models allow for the inclusion of covariates, e.g., attributes of the persons or families that are relevant to whether and how quickly they will pay a visit to other families. Therefore, they are suitable for analyzing structural and non-structural effects on the (hazard of) establishing ties. There are, however, some complications with network data measured in continuous time. Perhaps as a result of this, event history models for network structure (social selection model) are hard to find. Kim and Higgins, for example, analyze the hazard of biotechnical firms forming an alliance rather than predicting which firms become allies (Kim and Higgins 2007). Similarly, Tsai predicts the hazard of intraorganizational linkages (Tsai 2000). Robinson and Smith-Lovin investigate interaction within a group by modelling who takes a speaker turn assuming that the rest of the group is always the addressee, so they do not need to look at the level of ordered pairs (Robinson and Smith-Lovin 2001). As to date, Krempel's study of contacts among university freshman is the only example I have found of an event history model for the interaction between pairs of persons (Krempel 1990). In his event history analysis, he includes effects of the physical setting (where contacts take place) but he does not yet include aspects of local network structure produced by previous contacts as is our goal here.

The first complication of event history models for network data has been covered in Section 2: the multilevel structure of observations nested in ordered pairs nested within vertices. Multilevel models offer a neat solution to this problem, so I would like to stick to this type of design. Furthermore, I will point out another advantage of a multilevel structure for event history data later in this section.

The second complication arises from our interest in the effects of previous local network structure on the action. As has been pointed out, local network context is very likely to change from one moment to the next due to the fact that new actions add arcs to it and progression of the retrospective window removes (or decreases the weights of) older arcs. In the terminology of event history models, local network structure is a time-varying covariate, which complicates event history analyses.

In all event history models, time-varying covariates that are not merely functions of time define duration intervals during which they are constant (Blossfeld et al. 1989: 199-201). The hazard or survivor function is estimated for each duration interval (Box-Steffensmeier and Jones 2004: 97). To this end, the data are transformed into counting process format, which is also the data set up for discrete-time event history models (Table 1). Basically, the counting process data format contains one case or record for each observation that is at risk of experiencing an event at a particular moment. There is at least one variable scored for each case, namely the censoring indicator showing whether (1) or not (0) the case experienced the event at this time or a variable indicating which event happened to the case in a multiple states model. The design of counting process data with time-varying covariates is different for, on the one hand, discrete-time and semi-parametric continuous time (Cox) models and, on the other hand, parametric event history models, such as the Weibull model.

t	tail *i*	head *j*	censoring	conformity	reciprocity	…
…	…	…	…	…	…	…
13.7	1	2	0	0	0	…
13.7	1	3	0	0	1	…
13.7	1	4	1	0	1	…
13.7	1	5	0	0	0	…
…	…	…	…	…	…	…
14.9	1	2	0	0	0	…
14.9	1	3	0	0	1	…
14.9	1	4	1	1	1	…
14.9	1	5	0	0	0	…
14.9	2	1	0	0	0	…
14.9	2	3	0	0	0	…
14.9	2	4	0	0	1	…
14.9	2	5	0	0	0	…
14.9	3	1	0	1	0	…
14.9	3	2	0	0	0	…
14.9	3	4	0	1	0	…
14.9	3	5	0	0	0	…
14.9	4	1	0	0	0	…
14.9	4	2	0	1	0	…
14.9	4	3	0	0	1	…
14.9	4	5	0	0	0	…
14.9	5	1	0	0	0	…
14.9	5	2	0	0	0	…
14.9	5	3	0	0	0	…
14.9	5	4	0	0	0	…

Table 1 - Counting process data for (part of) the example in Figure 2.

In parametric models, each episode between two events is split into intervals containing the time from a change in any of the time-varying covariates up to the next change in any time-varying covariate or the next event. The likelihood function is evaluated for each interval. This is clearly a prospective approach: how long does it take for an event to happen once any time-varying covariate changes? In the case of local network context as a time-varying covariate using a retrospective window, this means that any movement of the window involving dropping an arc or adding an arc changes the values of structural time-varying covariates, so it defines the start of a new duration interval and the new values of the covariates must be calculated. Although this is conceptually simple, it implies a lot of work in data preparation. Moreover, many rather short duration intervals will emerge, so it is quite likely that estimating the likelihood function for each of them separately produce computational problems. Therefore, I will not extensively discuss parametric event history models here.

Both the discrete-time event history model and the semi-parametric Cox model only take into account the ordering of events, not the exact span of time between them. As a consequence, it is not important to know whether a covariate change took place a long time ago or right before the event happens. The downside of this approach is that we cannot estimate the length of durations or the time at which an event is predicted to happen as in parametric models. The upside, however, is very important: we only have to analyze the situation at the times of observed events and we can still estimate the effects of covariates such as local network structure. For that reason, the counting process data matrix in Table 1 only contains observations for the time points at which new action takes place in the network depicted in Figure 2.

It is interesting to note that the counting process data matrix is exactly the same as in the endogenous time model discussed in the previous section. But now we look at the data in a slightly different way. The length of time, or rather the time relative to other individuals, that an individual does not experience a risk is represented by the number of records it contributes to the counting process data set. Assembling the records of one individual, the sheer number of records in combination with the censoring indicator expresses the hazard of experiencing an event in comparison to other individuals: the more records, the lower the risk.

In event history models for discrete time data, we can apply the logit (or exponential) model to counting process data, estimating the effects of a constant and a vector of covariates on the censoring variable with logistic regression analysis. The log-odds of an event can then be interpreted as the hazard probability. Duration dependency may be modelled by including duration time as a covariate at the lowest level of the multilevel model using dummy variables, a mathematical transformation of duration time, or duration time smoothed with spline or LOWESS functions estimated from the data. Time dependency of risks is approached as a nuisance which is controlled for rather than estimated. If the event happens seldom, which is to be expected in networks that are not very small, the underlying process may be seen as a very slow Poisson process, so a Poisson regression model is more appropriate. In a multilevel design, it has been proposed to use a complementary log log model with a binomial distribution (Goldstein 1995: 132).

When we are using a regression approach, it is straightforward to apply the multilevel model specified in Section 2. The additional advantage of a multilevel model with levels for ordered pairs and vertices is that the random factors absorb unmeasured heterogeneity among pairs or vertices with respect to the hazard of experiencing an event as in frailty models. Heterogeneity at this point is a problem because the more resistant observations will remain in the risk set longer. In addition, the constraints with respect to who can act towards whom are the same as discontinuous risks in event history models, that is, intervals in which an individual is not at risk of experiencing the event. They are easily incorporated by changing the counting process data set as described in the previous section: simply omit records in the process count data set for pairs that cannot interact. This includes restrictions that tell that we may never expect action within a particular ordered pair (compare the split population approach (Box-Steffensmeier and Jones 2004: 148-154)).

It may appear inappropriate to use a discrete-time event history model for data measured in continuous time. If available software permits a multilevel semi-parametric Cox event history model using partial likelihood estimation, this is surely to be preferred. As an alternative, Cox models with robust variance estimators can be used (Box-Steffensmeier and Jones 2004: 158ff). However, if this software is not available, taking a discrete-time approach is defendable. As some authors have argued, discrete-time and continuous-time Cox models have gradual rather than absolute differences (Beck et al. 1998; Box-Steffensmeier and Jones 2004: 83). Measurement in continuous time is just more detailed than measurement in discrete time and the choice between the two usually results from available data sources rather than from theoretical arguments. In addition, the preferred way of coping with tied observations in the Cox model, that is, events happening at the same moment, is the exact discrete method, which follows the logic of counting process data and regression of the censoring indicator albeit that it prefers to predict the pattern of events among the members of the risk set rather than each individual event.

Finally, continuous-time network data usually contain repeated events as in multiple episodes time event history models, so we need an extra level for episode. There are several possibilities here. If we expect different processes for different episodes, e.g., the pattern of invitations is different for a person's very first birthday than for later birthdays, we should group the counting process data for the first episode for each tail and use the episode's sequential number as a covariate at the level of the episode, which is situated above the tail vertex level. In a Cox model, this will yield proportionality of the marginal effects of covariates on the hazard probability (Goldstein 1995: 128). Of course, to be meaningful the data must start at the very first episode in the 'careers' of the vertices.

If we do have clearly distinguishable episodes but we do not expect systematic differences between particular episodes, as for instance in the case of New Year's visits or greetings, which come back every year in the same way, we can add the episode level between the repeated measurements for the time-varying covariates (level i in Equations 1-7) and the level of the ordered pairs (j). Thus, we obtain overall proportionality of covariates' effects but the marginal effects of covariates on the hazard probability are not proportional in a Cox model (Goldstein 1995: 128).

Finally, if we do not have clearly demarcated episodes, e.g., in a group discussion in which people can address another person repeatedly in a short span of time whereas other pairs don't interact within that time span, it is probably best not to specify a separate level for episodes at all. Just collect all repeated measurements at the lowest level of the model (level i).

5. *Conclusion and discussion*

A multilevel logit or Poisson model with counting process data is not the fanciest solution to analyzing continuous-time network data because it applies a discrete-time model. However, it is extremely flexible for modelling continuous-time network data, both if timing of actions is exogenously given (fixed occasions) and if it is to be explained endogenously within the model. Exogenous restrictions and time varying covariates can be used while correcting for multilevel dependencies created by repeated observations.

To users it is a big advantage that they can use fairly general statistical software to model dynamics of social networks. Instead of spending time on understanding new statistical models such as the parametric event history models and learning to handle new software packages – provided that they are available – they may concentrate on organizing their data set, incorporating substantive constraints on who can act upon whom as well as how to deal with time. This is not to say that special parametric models should not be fitted, on the contrary. The logit model with counting process data offers a good starting point for the analysis. It does not allow making inferences about the precise timing of events but most questions regarding social networks focus primarily on the effects of covariates on tie formation.

The exploration of statistical designs for actions and events highlights some assumptions about the nature of social ties in panel studies. What are the epistemological differences between actions or events on the one hand and enduring relations on the other hand? Should we conceptualize ties as something enduring or rather not? Or should we do both and investigate the links between the two?

Formal social relations, that is, relations that are formalized by official registration, contracts, etcetera, obviously can be regarded as enduring, having clearly separated beginnings and end points in time. Think, for instance, of having a job in an organization, being married to someone, etcetera. It is easy to think of these types of relations as a rather stable structure. With informal social relations, e.g., friendship, this is not so obvious on second thoughts. Friendship can be thought of in two ways, viz., as a series of friendly acts towards another person or between two persons, or as a perception of a tie by one or both persons involved in it. Continuous-time designs focus on the former while panel designs address the latter.

Are the two equally valid operationalizations of the concept of friendship? It is well known that perceptions of friendship usually vary within pairs; while one person may say the other is its friend, the other person may not report being its friend. In addition, if a person reports being a friend of somebody else but no friendly actions are found among them or when unfriendly actions occur, what does the statement about friendship mean? It could represent an aspiration or wishful thinking rather than actually being friends. However, it is probably equally wrong to reduce friendship to objectively observable interaction among people. The interpretation of recent or more distant interaction by the actors themselves is probably just as ‘true in its consequences’ as words spoken. Perceptions matter. In sum, it is advisable to understand action and perceptions of social ties as mutually dependent, perceptions being constructed from interaction and affecting interaction in a cyclic process. Enduring ties are most probably social constructs but the social constructs are consequential to action.

Designs for continuous-time data can incorporate data on perceptions of social ties whereas panel designs cannot analyze continuous-time data without aggregating them. In this respect, a continuous-time design is more flexible. If data are collected using an Event Oriented Observation Plan (Blossfeld et al. 1989: 24) or the Life History Calendar method (Freedman et al. 1988), questions about perceptions of social ties can be asked although one should not expect very reliable information on perceptions as they were long time ago. Ideally, one would have documents of perceptions of social ties or use a panel designs for knowing perceptions of social ties at different moments during the period under investigation. Krempel’s study (Krempel 1990) is one of the rare examples of such a design.

References

Adams, A. M., Madhavan, S., & Simon, D. (2002). Women's social networks and child survival in Mali. *Social Science & Medicine*, 54, 165-178.

Beck, N., Katz, J. N., & Tucker, R. (1998). Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable. *American Journal of Political Science*, 42, 1260-1288.

Bernasco, W., de Graaf, P. M., & Ultee, W. C. (1998). Coupled careers - Effects of spouse's resources on occupational attainment in the Netherlands. *European Sociological Review*, 14, 15-31.

Blossfeld, H.-P., Hamerle, A., & Mayer, K. U. (1989). *Event History Analysis. Statistical Theory and Application in the Social Sciences*. Hillsdale (NJ): Lawrence Erlbaum.

Bogart, D. (2007). Neighbors, networks, and the development of transport systems: Explaining the diffusion of turnpike trusts in eighteenth-century England. *Journal of Urban Economics*, 61, 238-262.

Bohman, L. (2006). Economic action and interfirm relations: Diffusion of stock repurchases on the Stockholm Stock Exchange 2000-2003. *European Sociological Review*, 22, 383-396.

Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event History Modeling. A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Brandt, M. (2006). Finding a job via social contacts. *Kolner Zeitschrift Fur Soziologie Und Sozialpsychologie*, 58, 468-+.

Burt, R. S., & Lin, N. (1977). Network time series from archival records. In R. S. Heise (Ed.), *Sociological Methodology, 1977.* (pp. 224-254). San Francisco: Jossey-Bass.

Butts, C. T., & Pixley, J. E. (2004). A structural approach to the representation of life history data. *Journal of Mathematical Sociology*, 28, 81-124.

Bygren, L. O., Konlaan, B. B., & Johansson, S. E. (1996). Attendance at cultural events, reading books or periodicals, and making music or singing in a choir as determinants for survival: Swedish interview survey of living conditions. *British Medical Journal*, 313, 1577-1580.

Cartwright, D., & Harary, F. (1956). Structural balance: A generalisation of Heider's Theory. *Psychological Review*, 63, 277-293.

Chaves, M. (1996). Ordaining women: The diffusion of an organizational innovation. *American Journal of Sociology*, 101, 840-873.

Clarysse, B., Debackere, K., & Rappa, M. A. (1996). Modeling the persistence of organizations in an emerging field: The case of hepatitis C. *Research Policy*, 25, 671-687.

Davis, G. F., & Stout, S. K. (1992). Organization Theory and the Market for Corporate-Control - a Dynamic Analysis of the Characteristics of Large Takeover Targets, 1980-1990. *Administrative Science Quarterly*, 37, 605-633.

Doreian, P. (1990). Mapping networks through time. In J. Weesie & H. Flap (Eds.), *Social Networks Through Time* (pp. 245-264). Utrecht: ISOR / University of Utrecht.

Edling, C., & Sandell, R. (2001). Social influence and corporate behaviour - A case study of interdependent decision-making in Sweden's publicly traded firms. *European Sociological Review*, 17, 389-399.

Freedman, D., Thomton, A., Cambum, D., Alwin, D., & Young-DeMarco, L. (1988). The Life History Calendar: A Technique for Collecting Retrospective Data. In C. C. Clogg (Ed.), *Sociological Methodology 1988* (pp. 3748). San Francisco: Jossey-Bass.

Freeman, L. C. (2004). *The Development of Social Network Analysis. A Study in the Sociology of Science*. Vancouver BC, Canada: Empirical Press.

Friedkin, N. E., & Johnsen, E. C. (1990). Social Influence and Opinions. *Journal of Mathematical Sociology*, 15, 193-205.

Goldstein, H. (1995). *Multilevel Statistical Models*. London: Arnold.

Gustafsson, T. M., Isacson, D. G. L., & Thorslund, M. (1998). Mortality in elderly men and women in a Swedish municipality. *Age and Ageing*, 27, 585-593.

Holland, P. W., & Leinhardt, S. (1977). A dynamic model for social networks. *Journal of Mathematical Sociology*, 5, 5-20.

Hummon, N. P., & Carley, K. (1993). Social Networks as Normal Science. *Social Networks*, 15, 71-106.

Hummon, N. P., & Doreian, P. (1989). Connectivity in a Citation Network: The Development of DNA Theory. *Social Networks*, 11, 39-63.

Hummon, N. P., Doreian, P., & Freeman, L. C. (1990). Analyzing the Structure of the Centrality-Productivity Literature Created between 1948 and 1979. *Knowledge-Creation Diffusion Utilization*, 11, 459-480.
Kang, S. H., & Bloom, J. R. (1993). Social Support and Cancer Screening among Older Black-Americans. *Journal of the National Cancer Institute*, 85, 737-742.
Katz, L., & Proctor, C. H. (1959). The configuration of interpersonal relations in a group as a time-dependent stochastic process. *Psychometrika*, 24, 317-327.
Kim, J. W., & Higgins, M. C. (2007). Where do alliances come from? The effects of upper echelons on alliance formation. *Research Policy*, 36, 499-514.
Krempel, L. (1990). lnterpersonal Structure and Contact. Empirical Evidence from the Analysis of a Series of Social Networks in Time and Contact as a Process. An Analysis with Tools of 'Event History' Analysis. In J. Weesie & H. Flap (Eds.), *Social Networks Through Time* (pp. 65-90). Utrecht: ISOR.
Lipp, C., & Krempel, L. (2001). Petitions and the social context of political mobilization in the revolution of 1848/49: A microhistorical actor-centred network analysis. *International Review of Social History*, 46, 151-169.
Litwin, H., & Shiovitz-Ezra, S. (2006). Network type and mortality risk in later life. *Gerontologist*, 46, 735-743.
Marsden, P. V., & Podolny, J. M. (1990). Dynamic analysis of network diffusion processes. In J. Weesie & H. Flap (Eds.), *Social Networks Through Time* (pp. 197-214). Utrecht: ISOR.
McKeown, T. J. (1994). The Epidemiology of Corporate Pac Formation, 1975-84. *Journal of Economic Behavior & Organization*, 24, 153-168.
McPherson, J. M., Popielarz, P. A., & Drobnic, S. (1992). Social Networks and Organizational Dynamics. *American Sociological Review*, 57, 153-170.
Mintrom, M., & Vergari, S. (1998). Policy networks and innovation diffusion: The case of state education reforms. *Journal of Politics*, 60, 126-148.
Mossholder, K. W., Settoon, R. P., & Henagan, S. C. (2005). A relational perspective on turnover: Examining structural, attitudinal, and behavioral predictors. *Academy of Management Journal*, 48, 607-618.
Patterson, T. L., Shaw, W. S., Semple, S. J., Cherner, M., McCutchan, J. A., Atkinson, J. H., Grant, I., & Nannis, E. (1996). Relationship of psychosocial factors to HIV disease progression. *Annals of Behavioral Medicine*, 18, 30-39.
Payette, H., Coulombe, C., Boutier, V., & Gray-Donald, K. (2000). Nutrition risk factors for institutionalization in a free-living functionally dependent elderly population. *Journal of Clinical Epidemiology*, 53, 579-587.
Pearson, M., Steglich, C., & Snijders, T. (2006). Homophily and assimilation among sport-active adolescent substance users. *Connections*, 27, 47-63.
Robins, G., Elliott, P., & Pattison, P. (2001a). Network models for social selection processes. *Social Networks*, 23, 1-30.
Robins, G., & Pattison, P. (2001). Random graph models for temporal processes in social networks. *Journal of Mathematical Sociology*, 25, 5-41.
Robins, G., Pattison, P., & Elliott, P. (2001b). Network models for social influence processes. *Psychometrika*, 66, 161-189.
Robinson, D. T., & Smith-Lovin, L. (2001). Getting a laugh: Gender, status, and humor in task discussions. *Social Forces*, 80, 123-158.
Runger, G., & Wasserman, S. (1979). Longitudinal analysis of friendship networks. *Social Networks*, 2, 143-154.
Samuelsson, G., & Dehlin, O. (1993). Family Network and Mortality - Survival Chances through the Life-Span of an Entire Age Cohort. *International Journal of Aging & Human Development*, 37, 277-295.
Snijders, T. A. B. (1995). Methods for longitudinal social network data: Review and Markov process models. In E.-M. Tiit, T. Kollo & H. Niemi (Eds.), *Multivariate statistics and matrices in statistics : proceedings of the 5th Tartu Conference, Tartu-Pühajärve, Estonia, 23-28 May 1994* (pp. 211-227). Vilnius, Utrecht [etc.]: TEV, VSP.
— (2005). Models for longitudinal network data. In P. J. Carrington, J. Scott & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis* (pp. 215-247). Cambridge: Cambridge University Press.
Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. London: SAGE.

Snijders, T. A. B., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. *Personal Relationships*, 6, 471-486.
Snijders, T. A. B., Steglich, C. E. G., & Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. In K. v. Montfort, H. Oud & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 41-71). Mahwah, NJ: Lawrence Erlbaum.
Soule, S. A. (1997). The student divestment movement in the United States and tactical diffusion: The shantytown protest. *Social Forces*, 75, 855-882.
Strang, D., & Tuma, N. B. (1993). Spatial and Temporal Heterogeneity in-Diffusion. *American Journal of Sociology*, 99, 614-639.
Sutton, J. R., & Chaves, M. (2004). Explaining schism in American protestant denominations, 1890-1990. *Journal for the Scientific Study of Religion*, 43, 171-190.
Taris, T. W. (2000). *A Primer in Longitudinal Data Analysis*. London: SAGE.
Trovato, F. (1998). Nativity, marital status and mortality in Canada. *Canadian Review of Sociology and Anthropology-Revue Canadienne De Sociologie Et D Anthropologie*, 35, 65-91.
Tsai, W. P. (2000). Social capital, strategic relatedness and the formation of intraorganizational linkages. *Strategic Management Journal*, 21, 925-939.
Van den Bulte, C., & Lillien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology*, 106, 1409-1435.
Villingshoj, M., Ross, L., Thomsen, B. L., & Johansen, C. (2006). Does marital status and altered contact with the social network predict colorectal cancer survival? *European Journal of Cancer*, 42, 3022-3027.
Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks .1. An introduction to Markov graphs and p*. *Psychometrika*, 61, 401-425.
Wasserman, S., & Robins, G. (2005). An introduction to random graphs, dependence graphs, and p*. In P. J. Carrington, J. Scott & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis* (pp. 148-161). Cambridge: Cambridge University Press.
Weesie, J., & Flap, H. (1990). Social Networks Through Time (pp. vii-xi). Utrecht: ISOR.
White, H. C. (1970). *Chains of opportunity : System models of mobility in organizations*. Cambridge, Massachusetts: Harvard University Press.
Yakubovich, V., & Kozina, I. (2000). The changing significance of ties - An exploration of the hiring channels in the Russian transitional labor market. *International Sociology*, 15, 479-500.