Statistical Methods for Social Network Dynamics

C: Networks and Behavior

Tom A.B. Snijders

University of Groningen
University of Oxford

March 2023

1. Networks as dependent and independent variables

Co-evolution

Simultaneous endogenous dynamics of networks and behaviour: e.g.,

- individual humans & friendship relations:
  attitudes, behaviour (lifestyle, health, etc.)
- individual humans & cooperation relations:
  work performance
- companies / organisations & alliances, cooperation:
  performance, organisational success.

Two-way influence between networks and behaviour

Relational embeddedness is important
for well-being, opportunities, etc.

Actors are influenced in their behaviour, attitudes, performance
by other actors to whom they are tied
e.g., network resources (social capital), social control.

(N. Friedkin, *A Structural Theory of Social Influence*, C.U.P., 1998).

In return, many types of tie (friendship, cooperation, liking, etc.)
are influenced positively by
similarity on relevant attributes: *homophily*
(e.g., McPherson, Smith-Lovin, & Cook, *Ann. Rev. Soc.*, 2001.)

More generally, actors choose relation partners
on the basis of their behaviour and other characteristics
(similarity, opportunities for future rewards, etc.).

Influence, network & behaviour effects on *behaviour*;
Selection, network & behaviour effects on *relations*.

Terminology

relation = network = pattern of ties in group of actors;
behaviour = any individual-bound changeable attribute
(including attitudes, performance, etc.).

Relations and behaviours are endogenous variables
that develop in a simultaneous dynamics.

Thus, there is a feedback relation in the dynamics
of relational networks and actor behaviour / performance:
macro $\Rightarrow$ micro $\Rightarrow$ macro $\cdot \ \cdot \ \cdot\cdot$

(although network perhaps is meso rather than macro)

The investigation of such social feedback processes is difficult:

- Both the *network* $\Rightarrow$ *behaviour*
  and the *behaviour* $\Rightarrow$ *network* effects
  lead ‘network autocorrelation’:
  “friends of smokers are smokers”
  “high-reputation firms don’t collaborate
  with low-reputation firms”.
  It is hard to ascertain the strengths
  of the causal relations in the two directions.
- For many phenomena
  quasi-continuous longitudinal observation is infeasible.
  Instead, it may be possible to observe
  networks and behaviours at a few discrete time points.

Data

We consider again panel data:
network panel data, in which at two or more waves
for all actors in the network we observe

- $\Rightarrow$ network: who is tied to whom
- $\Rightarrow$ behavior,

where the behavior variable is assumed to be *ordinal discrete*
with integer values; simplest case: dichotomous.

(continuous behavior variables: Niezink & Snijders, *Soc. Methodology*, 2019)

Aim: disentangle effects *networks* $\Rightarrow$ *behavior*
from effects *behavior* $\Rightarrow$ *networks*.

Statistical Methodology for the evolution of networks and behavior

Integrate the *influence* (dep. var. = behavior)
and *selection* (dep. var. = network) processes.

Again the model assumes an evolution in continuous time;
the 'state' of the process now is
the combination of the network and the behavior of all actors;

each dependent variable (network, behavior)
has its own rate and evaluation function,
depending on <u>both</u> dependent variables,
which leads to their mutual dependence /
entwinement in a joint feedback process.

Outline of the co-evolution model: micro-step

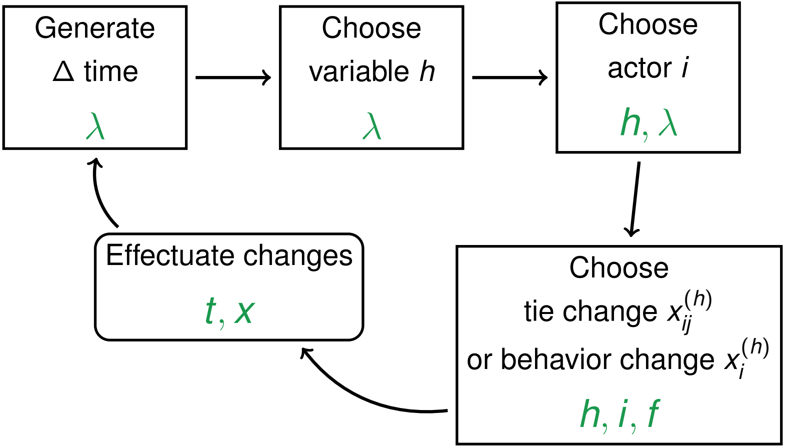
The co-evolution of a network X and a behavior variable Z proceeds in the following *smallest* steps:

1. at a random 'next' moment, an actor i is chosen,
   and a variable V is chosen which can be X or Z;
2. $\Rightarrow$ if $V = X$ then actor i chooses an actor j for creating or dropping the tie $i \rightarrow j$, or leaves everything unchanged;
   $\Rightarrow$ if $V = Z$ then actor i chooses an increment $-1,\ 0$, or $+1$ as the change for Z (restricted by its range).
3. the change (if any) is put into effect, and the process restarts.

(If there are several networks and/or behaviors,
there will be more V's.)

Flow chart for the micro-step

The co-evolution Markov chain is a succession of micro-steps; variables can be networks or actor-level variables.



Specification for the network-and-behavior model

The network and behavior both have their own evaluation function, with a basis constituted by what drives the variable itself, and added to this a dependence on each other.

For the network, the basis is as above;
for the behavior, the basis is a feedback model for Z (including regression to the mean) based on available variables;

Dependence on each other, e.g.:

selection : network ties $i \rightarrow j$ more likely when Z_i and Z_j are similar;

influence : when i's 'friends' on average are higher w.r.t. Z, Z_i will have a stronger upward tendency.

Actor-driven models – elaboration

Each actor "controls" not only his outgoing ties,
collected in the row vector $\big(X_{i1}(t), ..., X_{in}(t)\big)$,
but also his behaviour $Z_i(t) = \big(\ Z_{i1}(t), ..., Z_{iH}(t)\ \big)$
(H is the number of dependent behaviour variables).

Network change process and behaviour change process
run simultaneously, and influence each other
being each other's changing constraints.

At stochastic times
(*rate functions* λ^X for changes in network,
λ^{Z_h} for changes in behaviour h),
the actors may change a tie or a behaviour.

Probabilities of change are increasing functions of
evaluation functions of the new state,
defined specifically for network, f^X,
and for each behaviour, f^{Z_h} .

Again, only the smallest possible steps are allowed:
change one tie variable,
or move one step up or down on a behaviour variable.

For network change, change probabilities are as before.

For the behaviours, the formula of the change probabilities is

$$p_{ihv}(\beta, z) = \frac{\exp\big(f(i,h,v)\big)}{\sum_u \exp\big(f(i,h,u)\big)}$$

where $f(i,h,v)$ is the evaluation function calculated
for the potential new situation after a behaviour change,

$$f(i,h,v) = f_i^z\big(\beta, z(i, h \rightsquigarrow v)\big)\ .$$

Again, multinomial logit form.
The summation in the denominator extends
over the 2 or 3 options of permitted changes in $\{-1, 0, +1\}$.

Again, a 'myopic stochastic optimizing' interpretation is possible.

Micro-step for change in network:

Remember: at random moments occurring at a rate λ_i^X,
actor i is designated
to make a change in one tie variable:
the *micro-step* (on $\Rightarrow$ off, or off $\Rightarrow$ on.)

Micro-step for change in network:

Remember: at random moments occurring at a rate λ_i^X,
actor i is designated
to make a change in one tie variable:
the *micro-step* (on $\Rightarrow$ off, or off $\Rightarrow$ on.)

micro-step for change in behaviour:

At random moments occurring at a rate $\lambda_i^{Z_h}$,
actor i is designated to make a change in behaviour h
(one component of Z_i, assumed to be ordinal):
the *micro-step* is a change to an adjacent category; or stay the same.

Again, many micro-steps can *accumulate* to big differences.

Optimizing interpretation:

When actor i 'may' change an outgoing tie variable to some actor j,
he/she chooses the 'best' j by maximizing
the evaluation function $f_i^X(\beta, X, z)$ of the situation obtained
after the coming network change
plus a random component representing unexplained influences;

and when this actor 'may' change behaviour h,
he/she chooses the "best" change (up, down, nothing)
by maximizing the evaluation function $f_i^{Z_h}(\beta, x, Z)$ of the situation
obtained after the coming behaviour change
plus a random component representing unexplained influences.

There is no comparison network — behaviour.

Optimal network change:

The new network is denoted by $x^{(\pm ij)}$.
The attractiveness of the new situation
(evaluation function plus random term)
is expressed by the formula

$$f_i^X(\beta, x^{(\pm ij)}, z) \quad + \quad \underset{\substack{\Uparrow \\ \text{random component}}}{U_i^X(t, x, j)}\,.$$

(Note that the network is also permitted to stay the same.)

Optimal behaviour change:

Whenever actor i may make a change in variable h of Z,
he changes only one behaviour, say z_{ih}, to the new value v.
The new vector is denoted by $z(i, h \rightsquigarrow v)$.
Actor i chooses the "best" h, v by maximizing the evaluation function of the situation obtained after the coming behaviour change plus a random component:

$$f_i^{Z_h}(\beta, x, z(i, h \rightsquigarrow v)) \quad + \quad U_i^{Z_h}(t, z, h, v)\,.$$

$$\Uparrow$$

random component

(behaviour is permitted to stay the same.)

Specification of the behaviour model

Many different reasons why networks are important for behaviour; e.g.

1. *imitation* :
   individuals imitate others
   (basic drive; uncertainty reduction).
2. *social capital* :
   individuals may use resources of others;
3. *coordination* :
   individuals can achieve some goals
   only by concerted behaviour;

Theoretical elaboration helpful for a good data analysis.

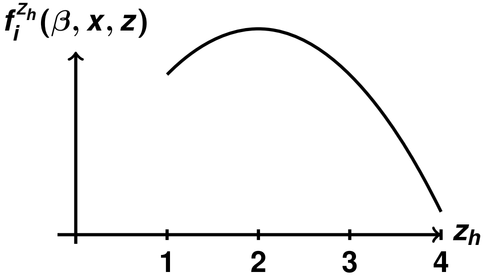
Evaluation function for dynamics of behaviour f_i^Z is again a linear combination

$$f_i^Z(\beta, x, z) = \sum_{k=1}^{L} \beta_k \, s_{ik}(x, z) \, .$$

Basic effects:

1. *linear shape* ,
   $s_{i1}^Z(x, z) = z_{ih}$
2. *quadratic shape*, ‘effect behaviour on itself’,
   $s_{i2}^Z(x, z) = z_{ih}^2$
   Quadratic shape effect important for model fit.

For a negative quadratic shape parameter,
the model for behaviour is a unimodal preference model.



For positive quadratic shape parameters ,
the behaviour evaluation function can be bimodal
('positive feedback').

3. *behaviour-related average similarity*,
   average of behaviour similarities between i and friends
   $s_{i3}(x) = \frac{1}{x_{i+}} \sum_j x_{ij} \, \mathrm{sim}(z_{ih}, z_{jh})$
   where $\mathrm{sim}(z_{ih}, z_{jh})$ is the similarity between v_i and v_j,

$$\mathrm{sim}(z_{ih}, z_{jh}) \, = \, 1 - \frac{|z_{ih} - z_{jh}|}{R_{Z^h}} \; ,$$

   R_{Z^h} being the range of Z^h;

3. *behaviour-related average similarity*,
   average of behaviour similarities between i and friends
   $s_{i3}(x) = \frac{1}{x_{i+}} \sum_j x_{ij} \, \mathrm{sim}(z_{ih}, z_{jh})$
   where $\mathrm{sim}(z_{ih}, z_{jh})$ is the similarity between v_i and v_j,
   $$\mathrm{sim}(z_{ih}, z_{jh}) \, = \, 1 - \frac{|z_{ih} - z_{jh}|}{R_{Z^h}} \, ,$$
   R_{Z^h} being the range of Z^h;
4. *average behaviour alter* — an alternative to similarity:
   $s_{i4}(x, z) = z_{ih} \frac{1}{x_{i+}} \sum_j x_{ij} z_{jh}$

Effects 3 and 4 are alternatives for each other:

they express the same theoretical idea of influence

in mathematically different ways.

The data, and/or theory, will have to differentiate between them.

Network position can also have influence on behaviour dynamics
e.g. through degrees rather than through behaviour
of those to whom one is tied:

5. *popularity-related tendency*, (in-degree)
   $s_{i5}(x, z) = z_{ih}\, x_{+i}$

Network position can also have influence on behaviour dynamics
e.g. through degrees rather than through behaviour
of those to whom one is tied:

7. *popularity-related tendency*, (in-degree)
   $s_{i7}(x, z) = z_{ih}\, x_{+i}$
8. *activity-related tendency*, (out-degree)
   $s_{i8}(x, z) = z_{ih}\, x_{i+}$

7. *dependence on other behaviours* $(h \neq \ell)$,
   $s_{i7}(x, z) = z_{ih}\, z_{i\ell}$
8. *influence from other characteristics V*
   $s_{i8}(x, z) = z_{ih}\, \frac{1}{x_{i+}} \sum_j x_{ij} v_j$,
   analogous to average alter for behaviour.

For both the network and the behaviour dynamics,
extensions are possible depending on the network position.

Now focus on the *similarity effect* in evaluation function :

sum of absolute behaviour differences between i and his friends
$s_{i2}(x, z) = \sum_j x_{ij} \, \mathrm{sim}(z_{ih}, z_{jh})$.

This is fundamental both
to network selection based on behaviour,
and to behaviour change based on network position.

A positive coefficient for this effect means that the actors prefer friends with similar Z_h values (*network autocorrelation*).

A positive coefficient for this effect means that the actors prefer friends with similar Z_h values (*network autocorrelation*).

Actors can attempt to attain this by changing their own Z_h value to the average value of their friends (*network influence*, *contagion*),

A positive coefficient for this effect means that the actors prefer friends with similar Z_h values (*network autocorrelation*).

Actors can attempt to attain this by changing their own Z_h value to the average value of their friends (*network influence*, *contagion*),

or by becoming friends with those with similar Z_h values (*selection on similarity*).

Statistical estimation: networks & behaviour

Procedures for estimating parameters in this model are similar to estimation procedures for network-only dynamics: Methods of Moments & Stochastic Approximation, conditioning on the first observation $X(t_1), Z(t_1)$.

The two different effects,
networks $\Rightarrow$ behaviour and behaviour $\Rightarrow$ networks,
both lead to network autocorrelation of behaviour;

but they can be (in principle)
distinguished empirically by the time order: respectively
association between ties at t_m and behaviour at t_{m+1};
and association between behaviour at t_m and ties at t_{m+1}.

Statistics for use in method of moments:

for estimating parameters in network dynamics:

$$\sum_{m=1}^{M-1} \sum_{i=1}^{n} s_{ik}(X(t_{m+1}), Z(t_m)) \ ,$$

and for the behaviour dynamics:

$$\sum_{m=1}^{M-1} \sum_{i=1}^{n} s_{ik}(X(t_m), Z(t_{m+1})) \ .$$

'cross-lagged statistics'.

The data requirements for these models are strong:
few missing data; enough change on the behavioural variable.

Currently, work still is going on about good ways
for estimating parameters in these models.

Maximum likelihood estimation procedures
(currently even more time-consuming; under construction...)
are preferable for small data sets.

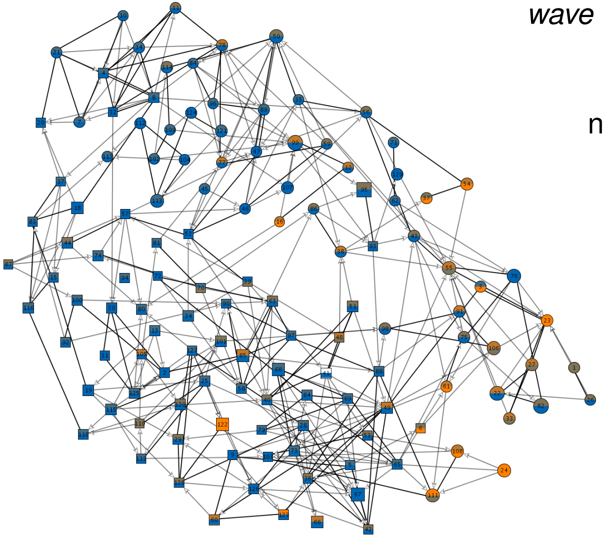
Example :

Study of smoking initiation and friendship
(following up on earlier work by P. West, M. Pearson & others)
(Steglich, Snijders & Pearson, *Sociological Methodology*, 2010).

One school year group from a Scottish secondary school
starting at age 12-13 years, was monitored over 3 years;
total of 160 pupils, of which 129 pupils present at all 3 observations;
with sociometric & behaviour questionnaires at three moments, at
appr. 1 year intervals.

Smoking: values 1–3;
drinking: values 1–5;

covariates:
gender, smoking of parents and siblings (binary), pocket money.



wave 1

girls: circles

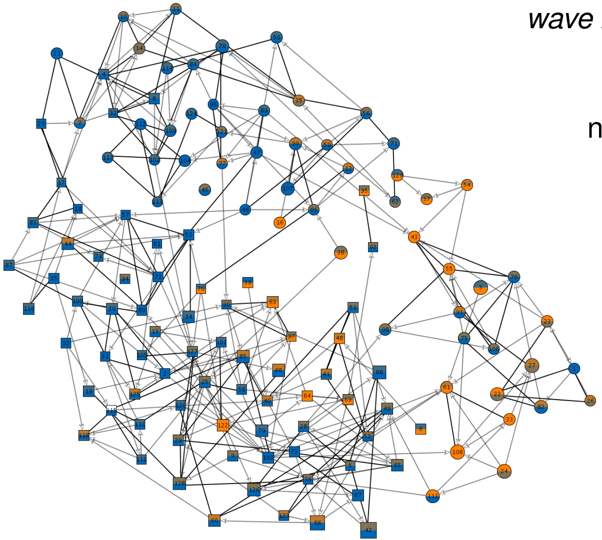
boys: squares

node size: pocket money

color: top = drinking

bottom = smoking

(orange = high)



wave 2

girls: circles

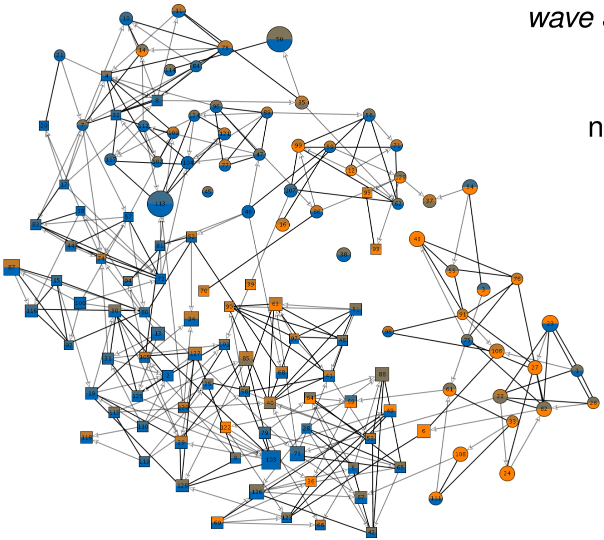
boys: squares

node size: pocket money

color: top = drinking

bottom = smoking

(orange = high)



wave 3

girls: circles

boys: squares

node size: pocket money

color: top = drinking

bottom = smoking

(orange = high)

Descriptives of covariate change: drinking

Observed changes in alcohol use in the Glasgow data, pooled over periods.

		t_{end}				
		1	2	3	4	5
	1: I don't drink (alcohol)	3	3	5	1	0
	2: once or twice a year	0	35	27	14	3
t_{begin}	3: about once a month	1	13	31	20	3
	4: about once a week	0	4	10	25	8
	5: more than once a week	0	0	2	4	11

The idea of an underlying process of micro-steps seems reasonable.

Descriptives of covariates: smoking

Observed changes in tobacco use, pooled over periods.

		t_{end}		
		1	2	3
t_{begin}	1: non-smoker	193	9	18
	2: occasional smoker	6	3	9
	3: regular smoker	3	3	27

Not so much variation.

Results

The table of results is distributed over 4 pages:

- structural effects and effect of sex
- friendship: effects of smoking, drinking, pocket money
- drinking
- smoking.

Effect	par.	(s.e.)
Network Dynamics		
constant friendship rate (period 1)	11.403	(1.147)
constant friendship rate (period 2)	9.237	(0.943)
outdegree (density)	–2.693***	(0.312)
reciprocity	3.388***	(0.290)
GWESP-FF ($\alpha = 0.30$)	2.430***	(0.131)
indegree - popularity	–0.053*	(0.024)
outdegree - activity	0.030	(0.044)
reciprocal degree - activity	–0.143*	(0.068)
indegree - activity	–0.120**	(0.046)
sex alter	–0.084	(0.101)
sex ego	0.017	(0.111)
same sex	0.558***	(0.087)
reciprocity $\times$ GWESP-FF	–0.913***	(0.256)

Network Dynamics		
Effect	par.	(s.e.)
Network Dynamics		
drinking alter	–0.016	(0.093)
drinking squared alter	–0.107	(0.096)
drinking ego	0.183†	(0.108)
drinking e–a difference squared	–0.090	(0.058)
smoking alter	0.132	(0.098)
smoking ego	–0.177	(0.116)
smoking similarity	0.437*	(0.179)
money/10 alter	0.105	(0.075)
money/10 squared alter	0.063	(0.040)
money/10 ego	–0.103	(0.076)
money/10 e–a difference squared	–0.067**	(0.025)

Effect	par.	(s.e.)
Behaviour Dynamics: drinking		
rate drinking (period 1)	1.634	(0.336)
rate drinking (period 2)	2.454	(0.534)
drinking linear shape	0.436**	(0.141)
drinking quadratic shape	–0.605**	(0.192)
drinking average alter	1.226*	(0.545)
drinking: effect from sex	0.068	(0.212)
drinking: effect from smoking	–0.096	(0.202)
drinking: effect from moneys	0.021	(0.015)

Effect	par.	(s.e.)
Behaviour Dynamics: smoking		
rate smoking (period 1)	4.389	(1.686)
rate smoking (period 2)	4.162	(1.345)
smoking linear shape	–3.375***	(0.356)
smoking quadratic shape	2.595***	(0.332)
smoking average alter	1.562**	(0.600)
smoking: effect from sex	–0.002	(0.270)
smoking: effect from smoking at home	–0.114	(0.264)
smoking: effect from drinking	–0.113	(0.245)
smoking: effect from moneys	0.016	(0.019)

† $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$;

convergence t ratios all < 0.03. Overall maximum convergence ratio 0.11.

The results for the structural network effects and for the effect of sex and money are almost the same as for the network-only analysis;
the effects of smoking and drinking on friendship are somewhat different, and have smaller standard errors;
their joint effect tests are less strongly significant.

Joint effect of drinking: $\chi^2_4 = 6.2, p = 0.19$.
Joint effect of smoking: $\chi^2_3 = 8.9, p = 0.03$.
Joint effect of pocket money: $\chi^2_4 = 15.3, p < 0.005$.

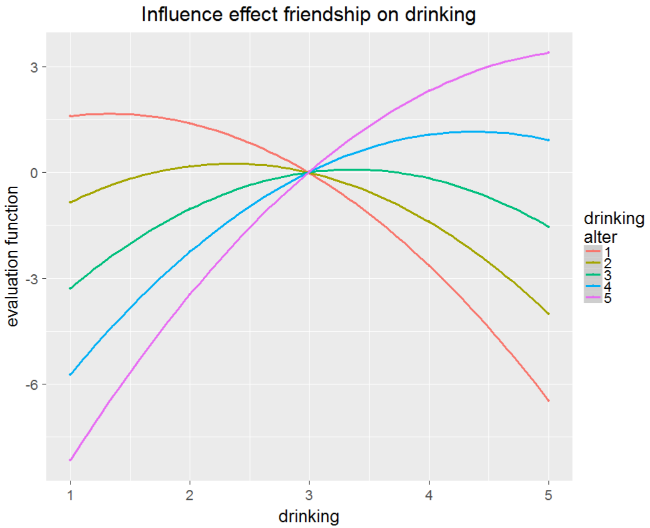
The influence effects for smoking and drinking are significant.

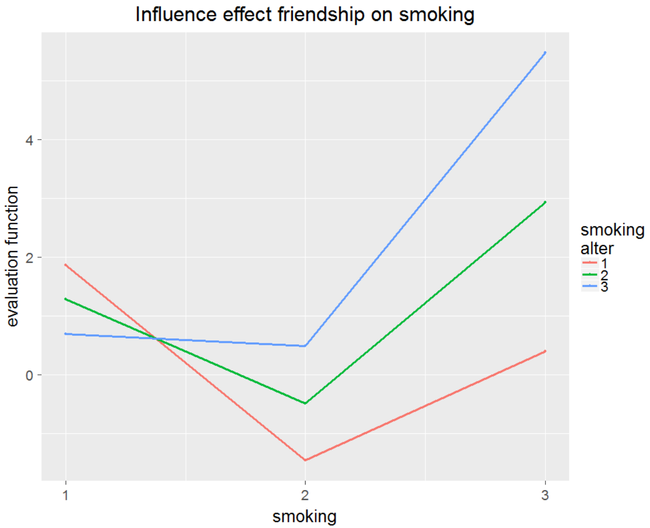
By the way, if for drinking the model is specified as ego, alter, and similarity, then similarity is marginally significant ($t = 1.62,\ p = 0.06$);
this illustrates the importance of choosing the model
before looking at results in case of a strict testing approach.

Parameter interpretation for behaviour change

The evaluation function for behaviour can be plotted as a function of Z, the behavior itself, for various different values of the average behaviour of the friends ('average alter').

This is treated in the manual as the *Ego-alter Influence Table*, and the website contains a script `InfluenceTables.r`.


Influence effect friendship on drinking
3
0
-3
-6
evaluation function
1
2
3
4
5
drinking
drinking alter
1
2
3
4
5


Influence effect friendship on smoking
evaluation function
4
2
0
1
2
3
smoking
smoking alter
1
2
3

Mind the different shapes of the functions for smoking and drinking:

For drinking, the influence function is concave,
and it is convex for smoking.
This is expressed by
the sign of the coefficient of the quadratic shape effect,
which is the quadratic term in the evaluation function.

Co-evolution, more generally

The idea of '*network-behaviour co-evolution*':

network is considered as one complex variable $X(t)$;

behaviour is considered as one complex variable $Z(t)$;

these are evolving over time in mutual dependence $X(t) \leftrightarrow Z(t)$,
changes occurring in many little steps,
where changes in X are a function of the current values of $\big(X(t), Z(t)\big)$,
and the same holds for changes in Z.

Co-evolution, more generally

The idea of '*network-behaviour co-evolution*':

network is considered as one complex variable $X(t)$;

behaviour is considered as one complex variable $Z(t)$;

these are evolving over time in mutual dependence $X(t) \leftrightarrow Z(t)$,
changes occurring in many little steps,
where changes in X are a function of the current values of $\big(X(t), Z(t)\big)$,
and the same holds for changes in Z.

This may be regarded as a 'systems approach',
and is also applicable to more than one network
and more than one behaviour.

4. Other co-evolution models

The co-evolution approach can be applied also to network-network co-evolution.

One or both of the networks could also be an affiliation network, i.e., a two-mode network where the first mode is the actor set and the second mode a set of binary non-exclusive attributes.

For example: individuals and clubs; firms and activities; etc.

Multiple networks require multilevel thinking

Interdependencies between networks can play on various levels; e.g., for friendship and advice:

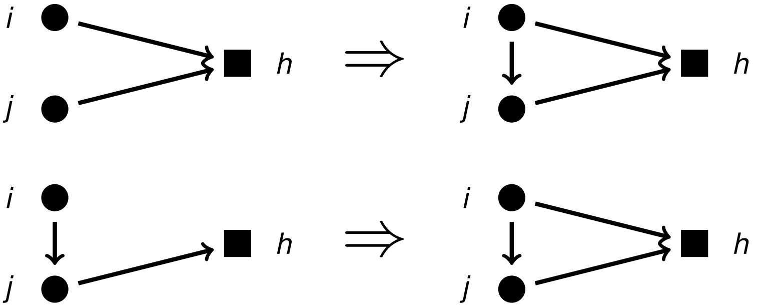
1. dyadic entrainment: friends become advisors;
2. dyadic exchange:
   I ask advice from those who say I am their friend;
3. actor level: those who have many friends get many advisors
   (not necessarily the same persons)
   (4 combinations in/outdegrees);
4. mixed closure 1: friends of friends become advisors;
5. mixed closure 2: advisors of friends become advisors;
6. and other mixed closures.

(See Snijders, Lomi, Torlò 2013; Snijders, 2016)

One-mode – two-mode co-evolution

For one-mode – two-mode co-evolution,
influence and selection can be modified to the comparison of affiliation-based focal closure and association-based affiliation closure:
(Cf. Easley & Kleinberg, 2010; Lomi & Stadtfeld, 2014)

Do we associate with those who have the same activities,
or do we choose the same activities as those with whom we associate?



Mixed closure in a combined one- and two-mode network.

Circles (left) are mode-1, squares (right) are mode-2 nodes.

Top: affiliation-based network closure;
bottom: network-based affiliation closure.

5. Miscellanea

Finally, a number of topics
that play around the background of this type of modeling.

Change and the Stochastic Actor-oriented Model

Parameters in the actor-oriented model determine how change occurs, but are not directly reflected by changes in network features.

Note that even though the conditional probabilities
as determined by the evaluation function are constant
(unless it contains time-dependent covariates),
the network itself may and usually will be changing
in the direction of some dynamic equilibrium
(like all Markov processes).

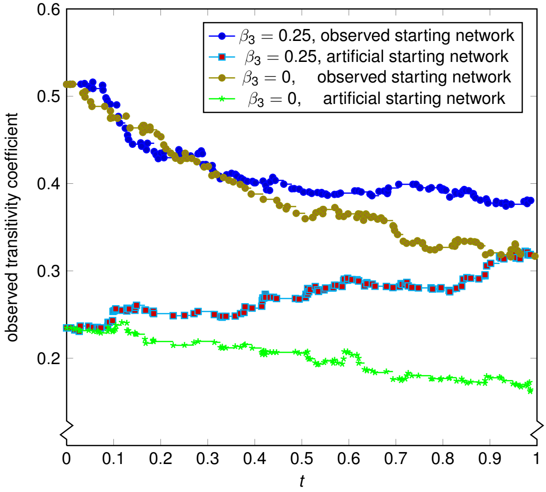
‘Constant transition distribution, changing marginal distribution’

Change and the Stochastic Actor-oriented Model (2)

Example : a positive transitivity parameter means that
there is a systematic tendency favoring transitivity;
but it does not mean that on average transitivity is increasing,
because there also are random tendencies away from transitivity.

For a network that starts with little transitive closure
a positive transitivity parameter will imply increasing transitivity;
but for a network that starts highly transitive,
a positive transitivity parameter may go together
with decreasing transitivity.

Next page shows a simulation example, combining
two different parameters and two different starting networks.



Artificial initial network:
reduced transitivity;
(light colors)

β_3 = transitivity parameter
in simulations.
(blue: 0.3; green: 0)

Blue curves have same parameters but different starting networks;

green curves likewise.

Model specification

For a good model specification, we need to start with reflection about what might influence the creation and disappearance of network ties, balancing between what is theoretically likely or possible and what is empirically discernible.

But we still know little about network dynamics.

- outdegree effect: balances between creation-termination of ties;
- reciprocity: 'always' there;
- transitivity: also 'always' there,
  but has several possible representations;
- degree effects:
  outdegrees vary because of (e.g.) response tendencies or resource differences, indegrees vary because of (e.g.) popularity or status differences, should be included by default.

Model specification: continued

For larger networks, the structure of the environment and the associated meeting opportunities must be represented; e.g., ‘same classroom’, distance, ‘same sector’.

Interactions are possible, also between covariates and structure.

Some checks for the model specification can be obtained by studying goodness of fit for distributions of indegree / outdegrees, triad census, distribution of geodesic distances.

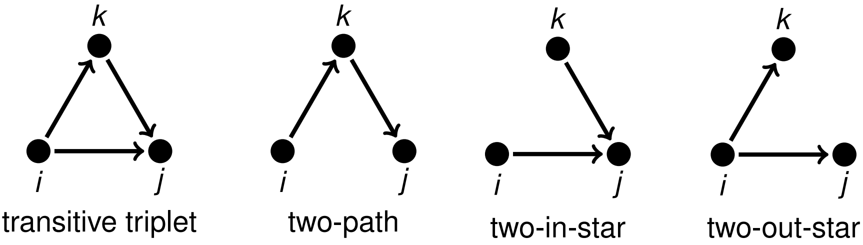
It is currently unknown how robust results are for misspecification.

Model specification: hierarchy requirements

There are hierarchy principles somewhat like in regression analysis: simpler configurations should be used as controls for complicated configurations.

This leads to heavy controls for multiple network co-evolution and complicated multi-node effects.

Hierarchy



The transitive triplet (left) includes three subgraphs (right);
actor *i* can create a transitive triplet by closing $i \to j$ or $i \to k$;
therefore, to properly test transitivity, the two-path
and two-in-star configurations should be included in the model.
These correspond to the
outdegree-popularity and indegree-popularity effects.

Causality?

Network data are often observational, and relations are crucial for how social actors try to attain their goals.
Therefore, networks in real life are highly endogenous.

Attaining causal conclusions about network effects from non-experimental studies is hard, because if ties are changed, actors will try something else that is similarly helpful for what they try to attain.

Causality in observational research, certainly in network research, is a Holy Grail: a lofty and important aim,
which we should not expect to attain;
cf. Shalizi & Thomas (2011):
selection and influence are generically confounded.

D.R. Cox / R.A. Fisher about causality: *Make your theories elaborate*, construct explanations at a deeper level.

P. Hedström & P. Ylikoski: *causal mechanisms*.

Network approaches themselves are a deeper level
than traditional quantitative social science approaches,
representing interaction processes,
and in this sense may help in coming closer to causal insights.

The approach of Stochastic Actor-oriented Modeling does not
lead to causal conclusions in the Holland-Rubin counterfactual sense;
it leads to conclusions about time sequentiality.

Network delineation

For a good network analysis, *network delineation* is important:
the analysis proceeds as if the delineated set is the whole world –
anathema to the basic tenets of the network approach.

Linked to this is the property that missing data, even randomly missing, can severely bias results of network analysis.

However, much network research is not ideal in this respect.

My impression is that, if the sampled network contains, for the actors included, the main parts of their relevant personal network, the general conclusions will tend to be correct; even if parameter estimates are biased.

This is supported by some very limited simulations.

6. Conclusion

- These models represent network structure
  as well as attributes / behaviour.
- Theoretically: they combine agency and structure.
- Available in
  package RSiena in the statistical system R.

6. Conclusion

- These models represent network structure
  as well as attributes / behaviour.
- Theoretically: they combine agency and structure.
- Available in
  package RSiena in the statistical system R.
- What was treated here is just the basic structure.
  Further possibilities, e.g.: multivariate,
  valued (only for few values!), two-mode, non-directed,
  continuous behaviour variables.
- Important: model choice, goodness-of-fit.
- The method is in a stage of continuous development:
  networks are very complicated data structures,
  we are only starting to understand them.

Discussion (2)

- This approach attempts to tackle peer effects questions by process modeling: data-intensive and potentially assumption-intensive.

Discussion (2)

- This approach attempts to tackle peer effects questions
  by process modeling: data-intensive
  and potentially assumption-intensive.
  Cox / Fisher: *Make your theories elaborate.*

Discussion (2)

- This approach attempts to tackle peer effects questions by process modeling: data-intensive and potentially assumption-intensive.
  Cox / Fisher: *Make your theories elaborate.*
- This type of analysis offers a very restricted take an causality: only *time sequentiality*; but network approaches can get closer to 'mechanisms' than approaches with atomic actors.

Discussion (2)

- This approach attempts to tackle peer effects questions by process modeling: data-intensive and potentially assumption-intensive.
  Cox / Fisher: *Make your theories elaborate.*
- This type of analysis offers a very restricted take an causality: only *time sequentiality*; but network approaches can get closer to ‘mechanisms’ than approaches with atomic actors.
- Assessing network effects is full of confounders.
  Careful theory development, good data are important.
  Asses goodness of fit of estimated model.

What distinguishes a statistical modeling approach from other kinds of network analysis?

What distinguishes a statistical modeling approach from other kinds of network analysis?

⇒ Direct combination of networks and attributes
and: combination of structure and agency.

What distinguishes a statistical modeling approach from other kinds of network analysis?

- $\Rightarrow$ Direct combination of networks and attributes
  and: combination of structure and agency.
- $\Rightarrow$ Distinction dependent $\Leftrightarrow$ explanatory variables

What distinguishes a statistical modeling approach from other kinds of network analysis?

$\Rightarrow$ Direct combination of networks and attributes
and: combination of structure and agency.

$\Rightarrow$ Distinction dependent $\Leftrightarrow$ explanatory variables

$\Rightarrow$ Hypothesis testing,
clearer support of theory development.

$\Rightarrow$ Combination of multiple mechanisms: test theories
while controlling for alternative explanations.

What distinguishes a statistical modeling approach from other kinds of network analysis?

$\Rightarrow$ Direct combination of networks and attributes
and: combination of structure and agency.

$\Rightarrow$ Distinction dependent $\Leftrightarrow$ explanatory variables

$\Rightarrow$ Hypothesis testing,
clearer support of theory development.

$\Rightarrow$ Combination of multiple mechanisms: test theories
while controlling for alternative explanations.

$\Rightarrow$ Assessment of uncertainties in inference.

$\cdots$ . . . but the classical network studies are also important
(positions, equivalence, centrality, blockmodeling, ...) !

Other work (recent, current, near future)

1. Changing composition of node set (Huisman & Snijders, *SMR* 2003).
2. Score-type tests (Schweinberger, *BJMSP* 2011).
3. Time heterogeneity (Lospinoso et al., *ADAC* 2011), function <sienaTimeTest>.
4. Goodness of fit (Lospinoso & Snijders, *Meth. Innovations* 2019), function <sienaGOF>.
5. Bayesian estimation; Maximum Likelihood estimation (Koskinen & Snijders, *J.Stat.Plann.Inf.* 2007); (Snijders, Koskinen, & Schweinberger, *Ann.Appl.Statist.* 2010).
6. Treatment of missing data (Krause et al., *Ital. J. Stat.*, 2018); script on website.
7. Explained variation ('R^2') (Snijders, *Math.Soc.Sci.* 2004; Indlekofer, 2014); function *sienaRI*.

Model extensions

1. Non-directed relations. (Snijders & Pickup, 2017)
2. Multivariate relations. (Snijders, Lomi, & Torlò, *SoN* 2013)
3. Valued relations (example in Elmer, Boda, Stadtfeld, *Network Sci.* 2017).
4. Two-mode networks. (Koskinen & Edling, *SoN* 2011;
   and Snijders, Lomi, & Torlò, *SoN* 2013)

Model extensions

1. Non-directed relations. (Snijders & Pickup, 2017)
2. Multivariate relations. (Snijders, Lomi, & Torlò, *SoN* 2013)
3. Valued relations (example in Elmer, Boda, Stadtfeld, *Network Sci.* 2017).
4. Two-mode networks. (Koskinen & Edling, *SoN* 2011; and Snijders, Lomi, & Torlò, *SoN* 2013)
5. Diffusion of innovations (effects `avExposure`, etc.; Greenan 2015).
6. Multilevel network analysis (meta analysis approach) (function <siena08>; Snijders & Baerveldt, *J.Math.Soc.* 2003).
7. Random effects multilevel network models (function <sienaBayes>; Koskinen, Snijders).
8. Continuous dependent behaviour variables (Niezink).
9. Larger networks, dropping assumption of complete information (settings model; Preciado/Snijders).

Further study – keeping updated

1. The version of RSiena at CRAN is not so frequently updated; check website - News whether the R-Forge version is preferable.
2. Basic tutorial: Tom A.B. Snijders, Gerhard G. van de Bunt, Christian E.G. Steglich (2010), Introduction to actor-based models for network dynamics. *Social Networks*, 32, 44–60.
3. The manual (available from website) has a lot of material.
4. Go through the website to see what's there:
   http://www.stats.ox.ac.uk/siena/
   For example, many useful scripts!
5. There is also a user's group:
   https://groups.io/g/RSiena

7. Some references about longitudinal models

- Tom A.B. Snijders (2001), The Statistical Evaluation of Social Network Dynamics, *Sociological Methodology*, 31, 361–395;
- Tom A.B. Snijders (2005), Models for Longitudinal Network Data, Ch. 11 in P. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis*. New York: Cambridge University Press.
- Tom A.B. Snijders (2017), Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and Its Application*, 4, 343–363.
- Tom A.B. Snijders, Johan Koskinen, and Michael Schweinberger (2010). Maximum Likelihood Estimation for Social Network Dynamics. *Annals of Applied Statistics*, 4, 567–588.
- Tom A.B. Snijders and Alessandro Lomi (2019). Beyond homophily: incorporating actor variables in statistical network models. *Network Science*, 7, 1–19.
- Tom A.B. Snijders and Mark Pickup (2017). Stochastic Actor-Oriented Models for Network Dynamics. *Oxford Handbook of Political Networks*, edited by J.N. Victor, A.H. Montgomery, and M. Lubell. Oxford: Oxford University Press.
- Tom A.B. Snijders, Gerhard G. van de Bunt, Christian E.G. Steglich (2010), Introduction to actor-based models for network dynamics. *Social Networks*, 32, 44–60.

Some references about network-behaviour co-evolution

- Yuval Kalish (2020). Stochastic Actor-Oriented Models for the co-evolution of networks and behavior: an introduction and tutorial. *Organizational Research Methods*, 23, 511-534.
- Christian E.G. Steglich, Tom A.B. Snijders, and Mike Pearson (2010). Dynamic Networks and Behavior: Separating Selection from Influence. *Sociological Methodology*, 40, 329–392.
- Tom A.B. Snijders and Christian E.G. Steglich (2015), Representing Micro-Macro Linkages by Actor-Based Dynamic Network Models. *Sociological Methods & Research*, 44, 222–271.
- René Veenstra, Jan Kornelis Dijkstra, Christian Steglich and Maarten H.W. Van Zalk (2013). Network-Behaviour Dynamics. *Journal of Research on Adolescence*, 23, 399–412.

Some references about multilevel networks and agent-based models.

- Viviana Amati, Alessandro Lomi, Daniele Mascia, and Francesca Pallotti (2021). The co-evolution of organizational and network structure: the role of multilevel mixing and closure mechanisms. *Organizational Research Methods*, 24, 285–318.
- Emmanuel Lazega and Tom A.B. Snijders, eds. (2016). *Multilevel Network Analysis for the Social Sciences; Theory, Methods and Applications*. Cham: Springer.
- Tom A.B. Snijders (2016). The Multiple Flavours of Multilevel Issues for Networks. Pp. 15–46 in Lazega & Snijders (2016).
- Tom A.B. Snijders, Alessandro Lomi, and Vanina Torlò (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks*, 35, 265–276.
- Tom A.B. Snijders and Christian E.G. Steglich (2015), Representing Micro-Macro Linkages by Actor-Based Dynamic Network Models. *Sociological Methods & Research*, 44, 222–271.

See `SIENA` manual and homepage for further references.

Some references in various languages

- Ainhoa de Federico de la Rúa, L'Analyse Longitudinale de Réseaux sociaux totaux avec SIENA – Méthode, discussion et application.
  BMS, Bulletin de Méthodologie Sociologique, 84, October 2004, 5–39.
- Ainhoa de Federico de la Rúa, El análisis dinámico de redes sociales con SIENA. Método, Discusión y Aplicación.
  Empiria, 10, 151–181 (2005).
- Mark Huisman and Tom A.B. Snijders, Een stochastisch model voor netwerkevolutie.
  Nederlands Tijdschrift voor de Psychologie, 58 (2003), 182-194.
- Laura Savoia (2007), L'analisi della dinamica del network con SIENA.
  In: A. Salvini (a cura di), *Analisi delle reti sociali. Teorie, metodi, applicazioni*, Milano: FrancoAngeli.
- Christian Steglich and Andrea Knecht (2009), Die statistische Analyse dynamischer Netzwerkdaten.
  In: Christian Stegbauer and Roger Häußling (Hsrg.), *Handbuch der Netzwerkforschung*, Wiesbaden (Verlag für Sozialwissenschaften).
- Tom A.B. Snijders 网络分析的动力学进路 Chinese translation (2018) for Chongqing University Press of 'Network Dynamics', Chapter 33 in John Scott and Peter J. Carrington (eds.), *The SAGE Handbook of Social Network Analysis* (London: Sage, 2011).