SIMULATION-BASED STATISTICAL INFERENCE FOR EVOLUTION OF SOCIAL NETWORKS

Tom A.B. Snijders

ICS

University of Groningen,

The Netherlands

January 2006

Social networks:

structures of relations between individuals.

Older studies of social support and influence
considered networks as *independent variables*
for explaining well-being (etc.);
this later led to studies of
network resources, social capital, solidarity,
in which the network is also a *dependent variable*.

Networks are dependent as well as independent variables:
intermediate structures in macro–micro–macro phenomena.

In this presentation:
focus first on networks as dependent variables,
then on mutual dependence networks and behavior
('behavior' stands here also for other individual attributes).

Single observations of networks are snapshots,
the results of untraceable history.
Therefore, explaining them has limited importance.

Longitudinal modeling offers more promise
for understanding of network structure.

The more descriptively oriented type of statistical modeling
of linear regression analysis etc.
cannot be transplanted to network analysis,
where the focus has to be on *modeling dependencies*.

Instead, longitudinal statistical modeling of networks
relies heavily on *modest process modeling*:
purposeful actors who optimize myopically
according to random utility models
subject to weak & limited rationality postulates.

Longitudinal data collection and modeling of social networks
has an *important advantage* over use of one-moment observations:

for modeling a single observation of a network,
"everything depends on everything else",
which leads to big problems in modeling and statistical inference;

for longitudinal modeling of social networks,
the first observation may be taken as given rather than modeled,
and then the remaining dependence is unidirectional in time
and less difficult to model.

1. Networks as dependent variables

Repeated measurements on social networks:
at least 2 measurements (preferably more).

Data requirements:
The repeated measurements must be close enough together,
but the total change between first and last observation
must be large enough
in order to give information about rules of network dynamics.

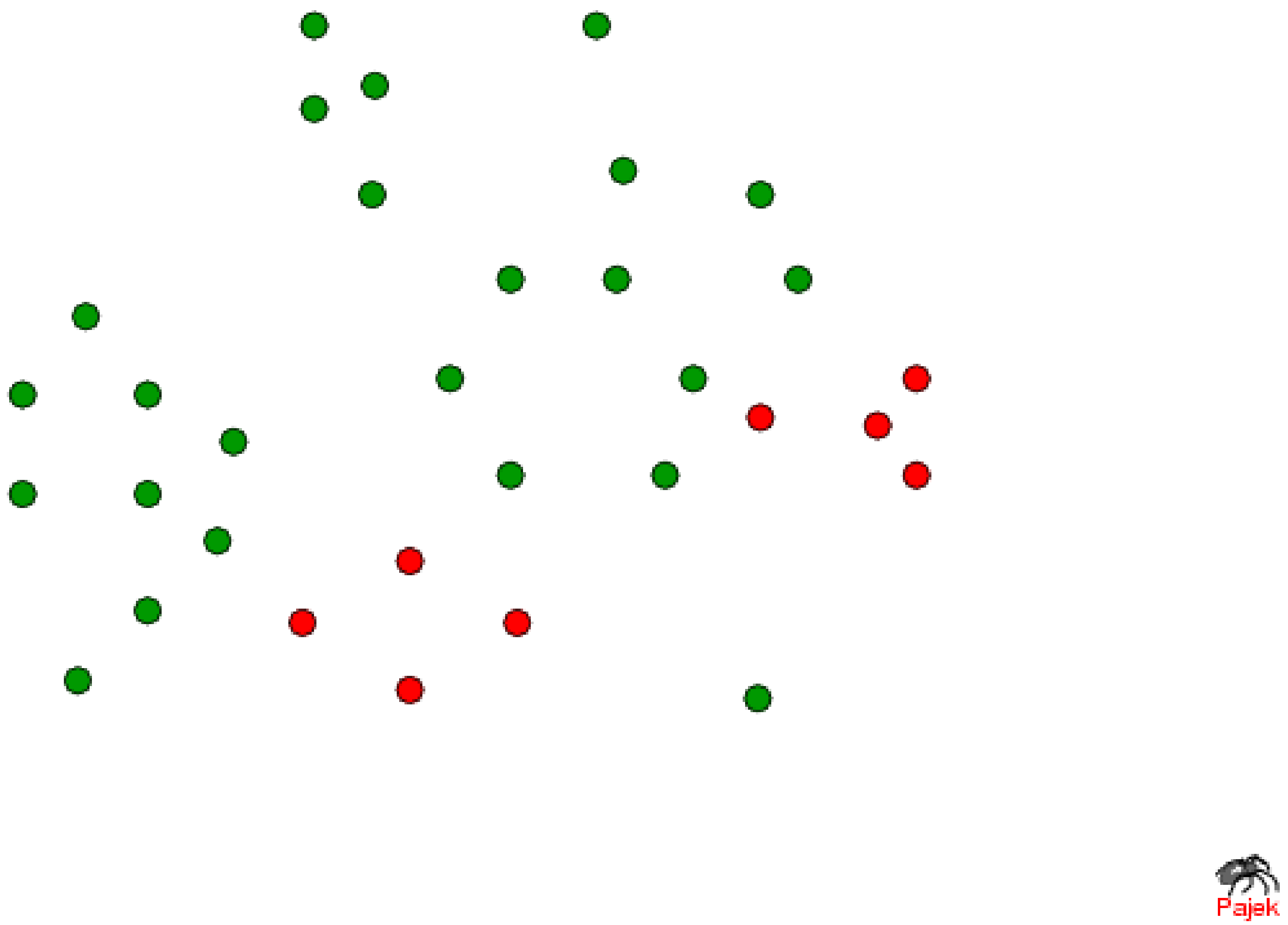
Go to modeling principles

Example:
Studies Gerhard van de Bunt

1. Study of 32 freshman university students,
   7 waves in 1 year.
   See van de Bunt, van Duijn, & Snijders,
   Computational & Mathematical Organization Theory,
   5 (1999), 167 – 192.

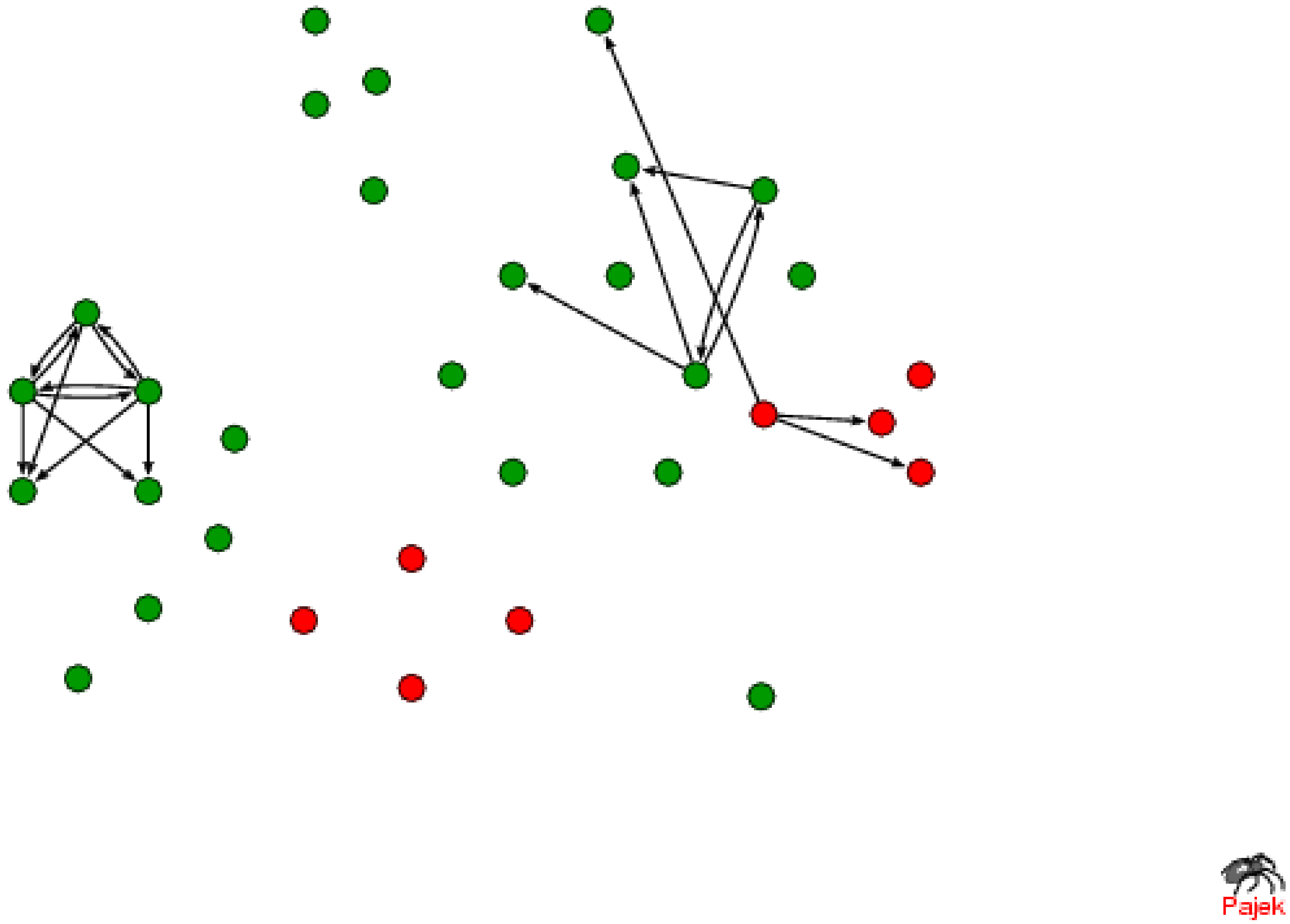
2. Study of hospital employees,
   2 departments (49 and 30 actors), 4 waves.

This presentation concentrates on the first data set,
which can be pictured by the following graphs
(arrow stands for ‘best friends’).



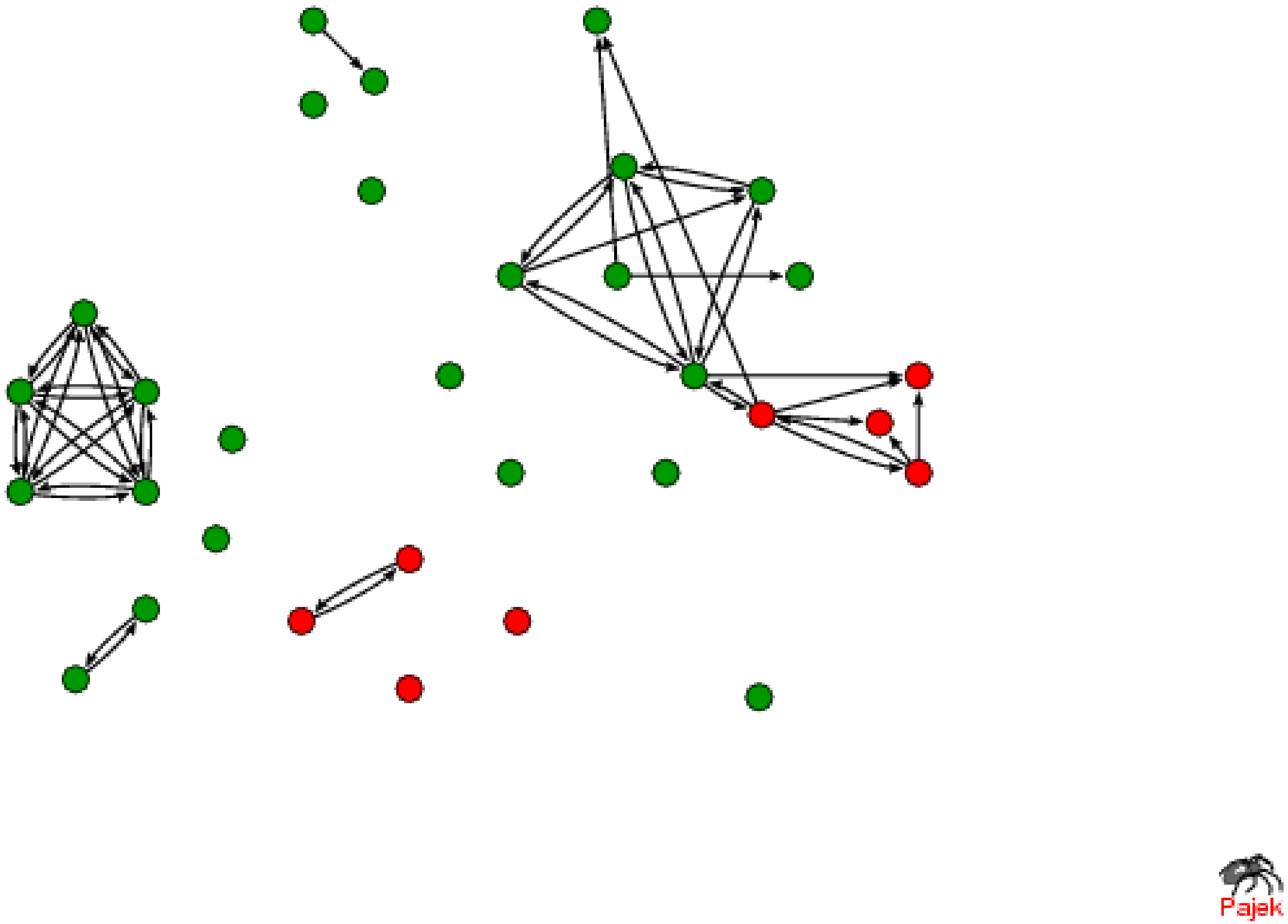
Friendship network time 1.

Average degree 0.0; missing fraction 0.0.



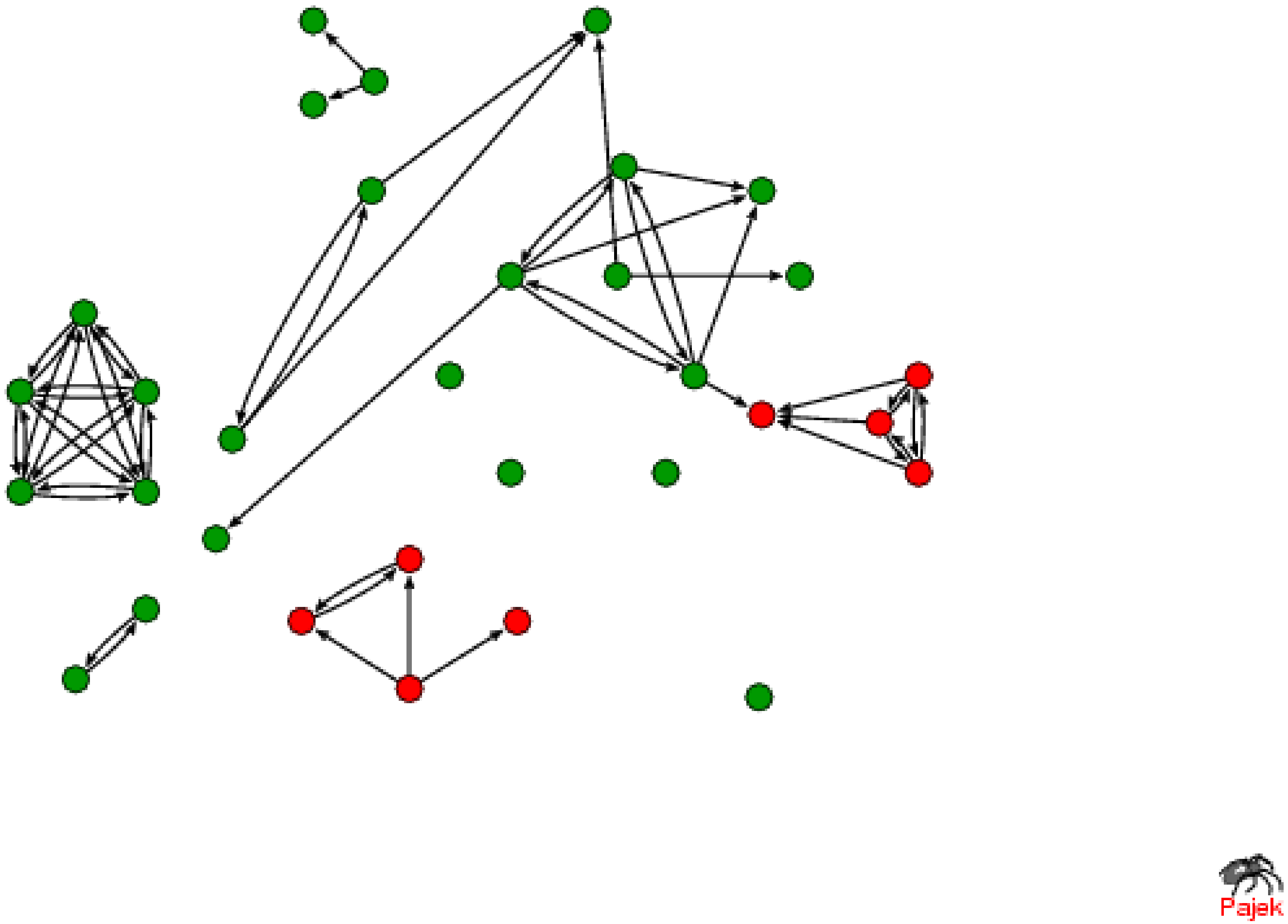
Friendship network time 2.

Average degree 0.7; missing fraction 0.06.



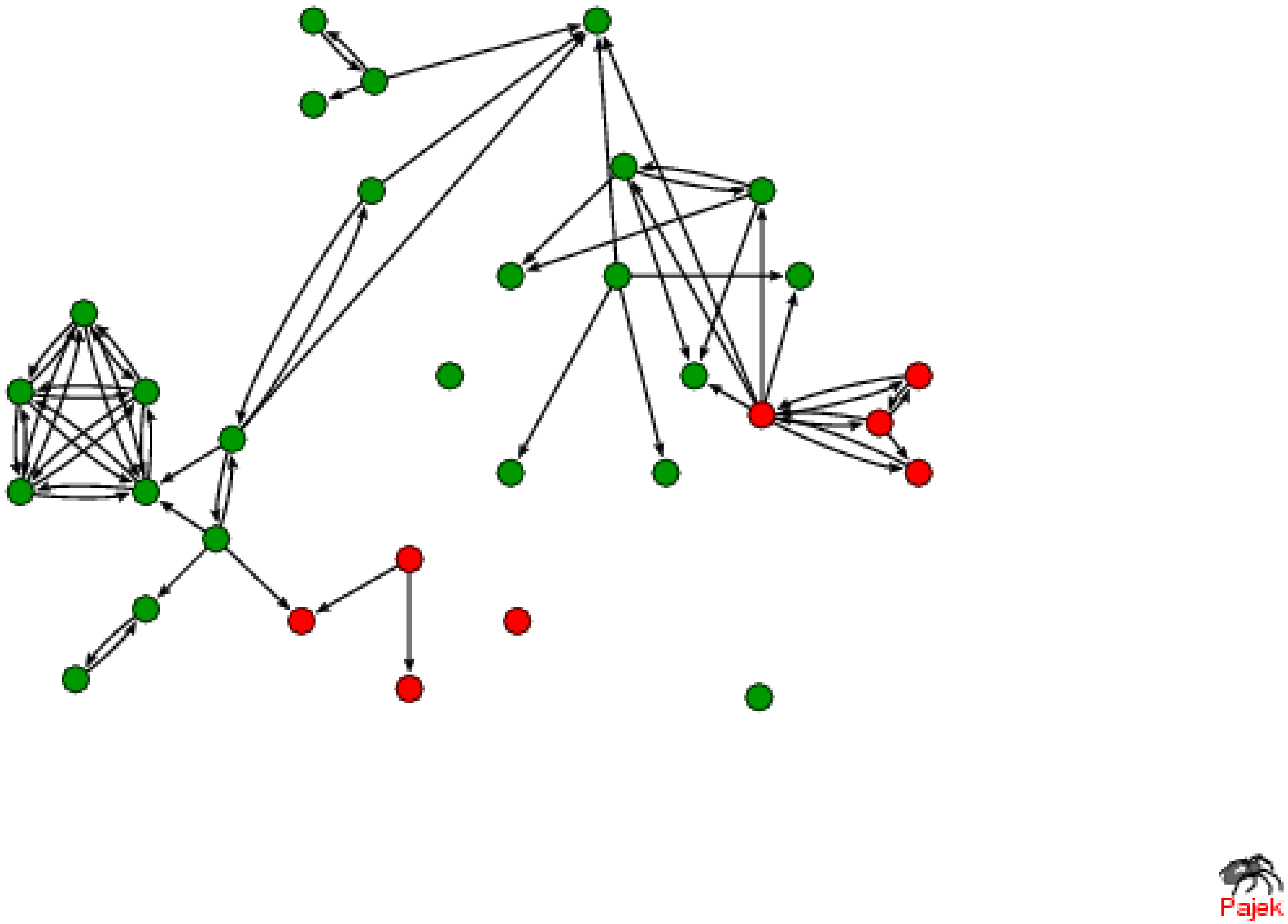
Friendship network time 3.

Average degree 1.7; missing fraction 0.09.



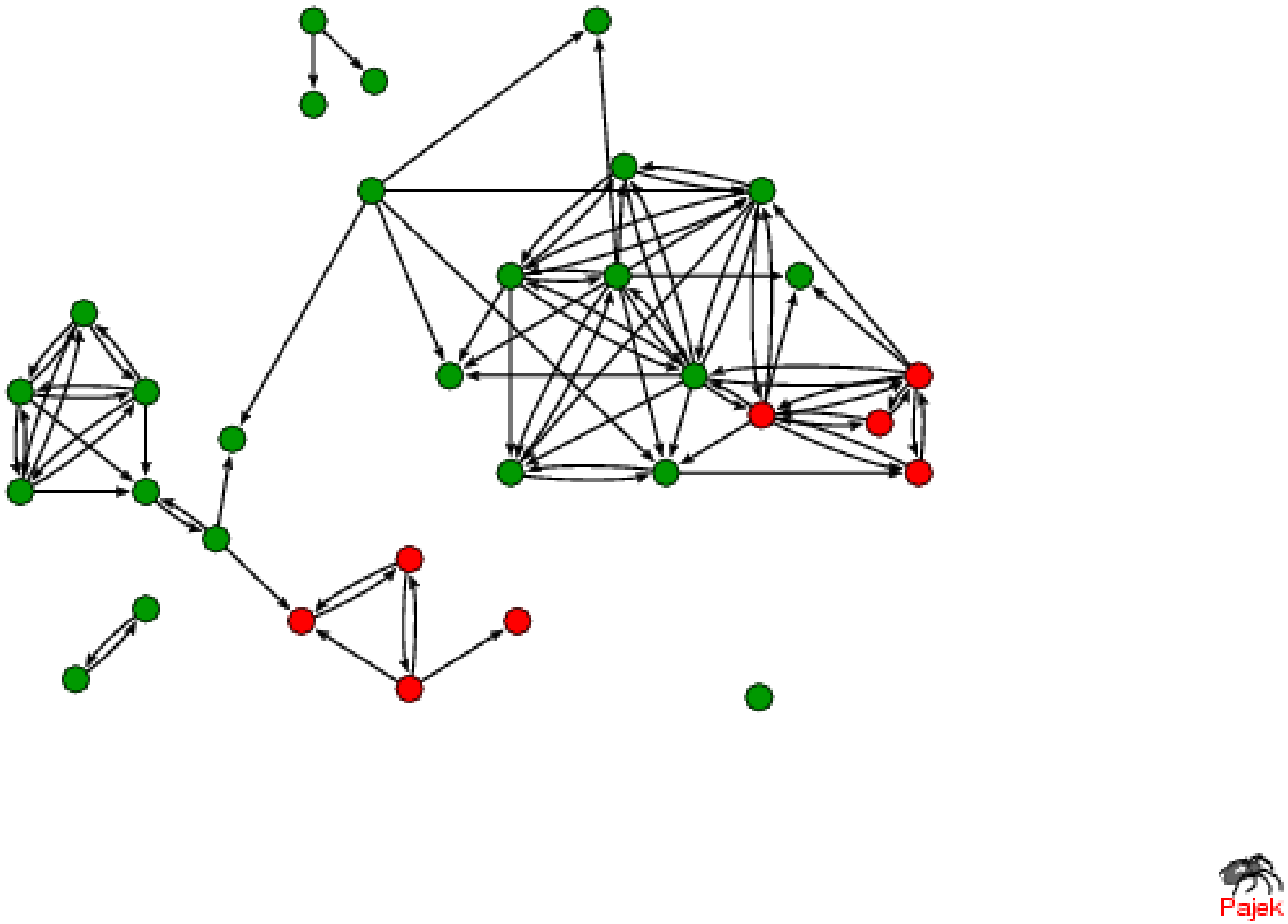
Friendship network time 4.

Average degree 2.1; missing fraction 0.16.



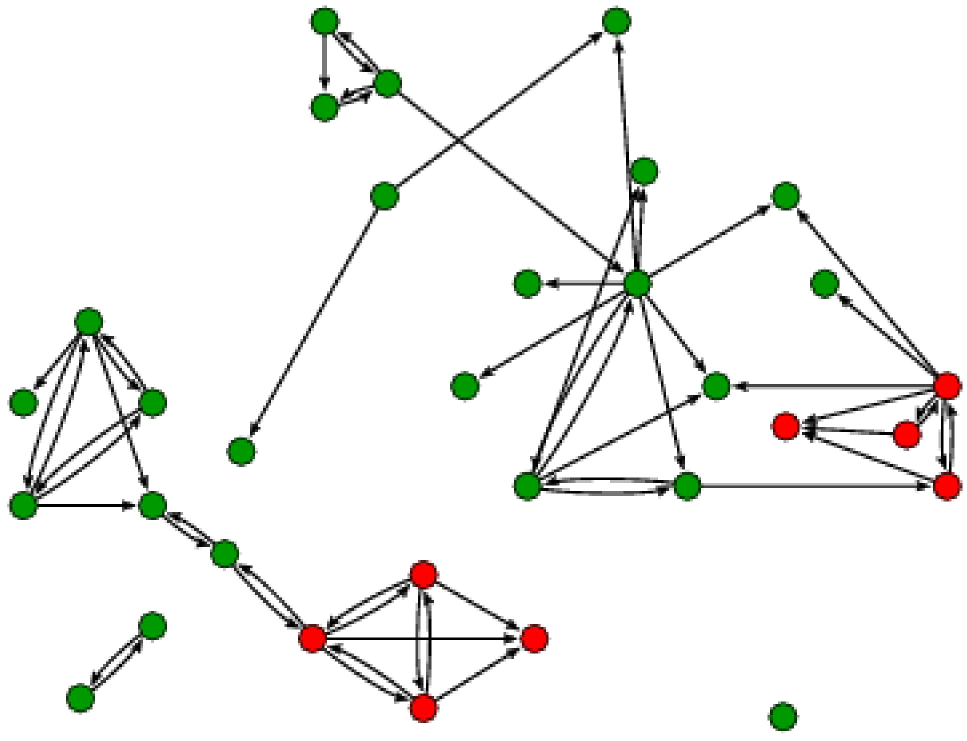
Friendship network time 5.

Average degree 2.5; missing fraction 0.19.



Friendship network time 6.

Average degree 2.9; missing fraction 0.04.



Friendship network time 7.

Average degree 2.3; missing fraction 0.22.

Which conclusions can be drawn from such a data set?

Dynamics of social networks are complicated
because "network effects" are endogenous feedback effects:
e.g., reciprocity, transitivity, popularity, subgroup formation.

In much other work:

- ⇝ computer simulation used in network analysis for investigation of theoretical properties:
  rich models.

For statistical modeling, one goes no further
than what can be estimated from data:
parsimonious modeling.

However, the statistical models must do justice
to the complexities of network structure:
complicated enough to be realistic,
not more complicated than necessary.

For a correct interpretation of empirical observations,
it is crucial to consider a model with *latent change* going on more or less continuously between the observation moments.

E.g., groups may be regarded as the result of the coalescence of relational dyads helped by a process of transitivity ("friends of my friends are my friends").
Which groups form may be contingent on unimportant details;
that groups will form is a sociological regularity.

Therefore: use dynamic models with *continuous time parameter*:
time runs on between observation moments.

An advantage of using continuous-time
models, even if observations are made at a few discrete time points,
is that a more natural and simple representation may be found,
especially in view of the endogenous dynamics.
(cf. Coleman, 1964).
No problem with irregularly spaced data.

For *discrete data*:
Kalbfleisch & Lawless, JASA, 1985;
for *continuous data*:
mixed state space modelling well-known in engineering,
in economics e.g. Bergstrom (1976, 1988),
in social science Tuma & Hannan (1984),
work by H. Singer in the 1990s.

Here: *discrete data with complicated structure*.

Purpose of statistical inference:
investigate network evolution as function of

1. structural effects (reciprocity, transitivity, etc.)
2. explanatory actor variables
3. explanatory dyadic variables

all controlling for each other.

By controlling adequately for structural effects, it is possible to test hypothesized effects of variables on network dynamics (without such control these tests would be unreliable).

The structural effects imply that the presence of arcs is highly dependent on the presence of other arcs.

Principles for this approach to analysis of network dynamics:

* use simulation models as *models for data*

* comprise a random influence in the simulation model
  to account for 'unexplained variability'

* develop statistical inference
  for probability models implemented as simulation models

* employ a continuous-time model
  (even if observations are at discrete time points)
  to represent endogenous network evolution

* condition on the first observation
  and refrain from modeling it:
  no stationarity assumption.

Notation and assumptions

(some of which may and will be relaxed):

1. *Actors* $i = 1, \ldots, n$ (individuals in the network),
   pattern X of *ties* between them :
   one, binary, network X;
   $X_{ij} = 0$, or 1 if there is no tie, or a tie, from i to j.
2. Exogenously determined independent variables:
   actor-dependent covariates v_h, dyadic covariates w_h.
   These can be constant or changing over time.
3. Although the data collection may follow a panel design,
   in the underlying model, time (t) is a continuous parameter.
4. The change process is stochastic.
5. The current state of the network $\big(X(t)\big)$
   acts as a dynamic constraint for its own process of change:
   Markov process.

6. The actors control their outgoing ties.
7. The ties have inertia: they change in small steps.
   At any single moment in time,
   only one variable $X_{ij}(t)$ may change.

8. Changes are made by the actors to optimize their situation,
   as it will obtain immediately after this change.
9. Assessment by actors of their situation comprises
   random element, expressing aspects not modeled explicitly.

(8) and (9): goal-directed behavior,
in the weak sense of myopic stochastic optimization.

Assessment of the situation is represented by
objective function, interpreted as
'that which the actors seem to strive after in the short run'.

Further elaboration:

At randomly determined moments t,
actors i have opportunity to change a tie variable X_{ij}: *micro step*.
(It is also allowed not to change anything.)
The frequency of such micro steps is determined
by *rate functions*.

If the rate of change of the network by actor i is λ_i ,
this means that, for a certain short time interval $(t, t+\epsilon)$,
the probability that this actor randomly gets an opportunity
to change one of his/her outgoing ties, is given by $\epsilon \times \lambda_i$.

When a micro step is taken,
the actor optimizes an *objective function*
which is the sum of a deterministic and a random part.
The random part reflects unexplained variation.

Specification: rate function

'how fast is change / opportunity for change ?'

Rate functions can depend on the observation period $(t_{m-1},\ t_m)$, actor covariates, actor behavior, and network position, through an exponential link function.

In a simple specification, the rate functions are constant within periods.

Network rate functions could also e.g. increase with out-degrees: more activity $\sim$ more change, especially for strongly heterogeneous actors.

Specification: objective function

'what is the direction of change?'

Objective functions for the network will be defined as the sum of:

1. *evaluation function*, expressing satisfaction
   with the network;
2. *endowment function*, expressing aspects of satisfaction
   with the network
   that are obtained 'free' but are lost at a value
   (to allow asymmetry between creation and deletion of ties);
3. *random variable* with a Gumbel distribution
   leading to probabilities as in multinomial logit modeling.

The objective function does not reflect the eventual 'utility'
of the situation to the actor, but short-time goals
following from preferences, constraints, opportunities.

The evaluation and endowment functions express how the dynamics of the network process depends on its current state.

Evaluation function and endowment function modeled
as linear combinations of theoretically argued components
of actors' assessment of the network.
The weights in the linear combination are the statistical parameters
(cf. regression coefficients).

The focus of modeling is first on the evaluation function;
then on the rate and endowment functions.

Example: SIENA applet.

Mini-step:

At random moments (frequency determined by rate function), a random actor gets the opportunity to make a change in one tie variable: the *mini-step* (on ⇒ off, or off ⇒ on).

It is allowed to change nothing.

This actor tries to improve his/her objective function and looks only on its value immediately after this mini-step (*myopia*) .

This absence of strategy or farsightedness in the model implies the *definition* of effects as "what the actors try to achieve in the short run".

All mini-steps are sequential, *no coordination* between actors (such as swapping partners).

Many mini-steps can *accumulate* to big differences between consecutively observed networks.

Simple model specification:

* The actors all change their relationships at random moments, at the same rate ρ.

* Each actor tries to optimize an *objective function* with respect to the network configuration,

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

  which indicates the preference of actor i for the relational situation represented by x; objective function depends on *parameter* β.

Whenever actor i may make a change,
he changes only one relation, say x_{ij}.
The new network is denoted by $x(i \rightsquigarrow j)$.
(Formally denote by $x(i \rightsquigarrow i)$
the network where nothing has changed: $x(i \rightsquigarrow i) = x$.)
Actor i chooses the "best" j by maximizing

$$f_i\Big(\beta, x(i \rightsquigarrow j)\Big) \;+\; U_i(t, x, j)\,.$$

⇑
random component

For a convenient choice of the distribution
of the random component,
(type 1 extreme value = Gumbel distribution)
given that i is allowed to make a change,
the probability that i changes his relation with j is

$$p_{ij}(\beta, x) = \frac{\exp\big(f(i,j)\big)}{\sum_{h=1, h\neq i}^{n} \exp\big(f(i,h)\big)} \quad (j \neq i).$$

where $f(i,j) = f_i\Big(\beta, x(i \rightsquigarrow j)\Big)$;
the probability of no change is $p_{ii}(\beta, x)$.

This is the multinomial logit form of a *random utility* model.

The Gumbel distribution has variance $\pi^2/6 = 1.645$ and s.d. 1.28.

Skip details.

Intensity matrix

This specification implies that X follows a *continuous-time Markov chain* with intensity matrix

$$q_{ij}(x) = \lim_{dt \downarrow 0} \frac{\mathsf{P}\left\{X(t+dt) = x(i \rightsquigarrow j) \mid X(t) = x\right\}}{dt} \qquad (i \neq j)$$

given by

$$q_{ij}(x) = \lambda_i(\alpha, \rho, x)\, p_{ij}(\beta, x)\,.$$

Computer simulation algorithm

for arbitrary rate function $\lambda_i(\alpha, \rho, x)$

1. Set $t = 0$ and $x = X(0)$.
2. Generate S according to the negative exponential distribution with mean $1/\lambda_+(\alpha, \rho, x)$ where

$$\lambda_+(\alpha, \rho, x) = \sum_i \lambda_i(\alpha, \rho, x) \ .$$

3. Select randomly $i \in \{1, ..., n\}$ using probabilities

$$\frac{\lambda_i(\alpha, \rho, x)}{\lambda_+(\alpha, \rho, x)} \ .$$

4. Select randomly $j \in \{1, ..., n\}$, using probabilities $p_{ij}(\beta, x)$.
5. Set $t = t + S$ and $x = x(i \rightsquigarrow j)$.
6. Go to step 2
   (unless stopping criterion is satisfied).

Model specification :

The objective functions f_i reflect
network effects (endogenous) and covariate effects (exogenous).

Covariates can be actor-dependent: v_i
or dyad-dependent: w_{ij} .

Convenient definition of objective function f_i
is a weighted sum

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

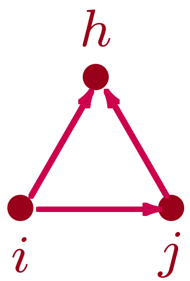
where weights β_k are statistical parameters
indicating strength of effect $s_{ik}(x)$.

Choose possible network effects for actor i, e.g.:
(others to whom actor i is tied are called here i's 'friends')

1. *density effect*, out-degree
   $s_{i1}(x) = x_{i+} = \sum_j x_{ij}$
2. *reciprocity effect*, number of reciprocated relations
   $s_{i2}(x) = \sum_j x_{ij}\, x_{ji}$
3. *popularity effect*, sum of in-degrees of i's friends
   $s_{i3}(x) = \sum_j x_{ij}\, x_{+j} = \sum_j x_{ij} \sum_h x_{hj}$
4. *activity effect*, sum of the out-degrees of i's friends
   $s_{i4}(x) = \sum_j x_{ij}\, x_{j+} = \sum_j x_{ij} \sum_h x_{jh}$

Three effects related to network closure:

5. *transitivity effect*,
   number of transitive patterns in i's relations
   $(i \to j,\ j \to h,\ i \to h)$
   $s_{i5}(x) = \sum_{j,h} x_{ij}\, x_{jh}\, x_{ih}$



transitive triplet

6. *indirect relations effect*,
   number of actors j to whom i is indirectly related
   (through at least one intermediary: $x_{ih} = x_{hj} = 1$)
   but not directly ($x_{ij} = 0$),
   = number of geodesic distances equal to 2,
   $s_{i6}(x) = \#\{j \mid x_{ij} = 0, \max_h(x_{ih}\, x_{hj}) > 0\}$

7. *balance* or structural equivalence,
similarity between out-relations of i
with out-relations of his friends,

$$s_{i7}(x) = \sum_{j=1}^{n} x_{ij} \sum_{\substack{h=1 \\ h \neq i,j}}^{n} \Big(1 - |x_{ih} - x_{jh}|\Big) ,$$

[note that $\Big(1 - |x_{ih} - x_{jh}|\Big) = 1$ if $x_{ih} = x_{jh}$,
and 0 if $x_{ih} \neq x_{jh}$, so that

$$\sum_{\substack{h=1 \\ h \neq i,j}}^{n} \Big(1 - |x_{ih} - x_{jh}|\Big)$$

measures agreement between i and j .]

Differences between these three network closure effects:

$\Rightarrow$ transitive triplets effect:
i more attracted to j if there are
more indirect ties $i \rightarrow h \rightarrow j$;

$\Rightarrow$ negative indirect connections effect:
i more attracted to j if there is
at least one such indirect connection ;

$\Rightarrow$ balance effect:
i prefers others j who make same choices as i.

Non-formalized theories usually do not distinguish
between these different closure effects.
It is possible to 'let the data speak for themselves'
and see what is the best formal representation of closure effects.

Three kinds of objective function effect associated with actor covariate v_i :

8. *covariate-related popularity*,
   sum of covariate over all of i 's friends
   $s_{i8}(x) = \sum_j x_{ij}\, v_j$;
9. *covariate-related activity*,
   i's out-degree weighted by covariate
   $s_{i9}(x) = v_i\, x_{i+}$;
10. *covariate-related similarity*,
    sum of measure of covariate similarity between i and his friends, e.g.
    $s_{i10}(x) = \sum_j x_{ij}\, \big(1 - |\, v_i - v_j \,|\big)$
    if V has values between 0 and 1.

Objective function effect for dyadic covariate w_{ij} :

11. *covariate-related preference*,
sum of covariate over all of i's friends,
i.e., values of w_{ij} summed over all others to whom i is related,
$s_{i11}(x) = \sum_j x_{ij}\, w_{ij}$.

If this has a positive effect, then the value of a tie $i \rightarrow j$ becomes higher when w_{ij} becomes higher.

Example

Data collected by Gerhard van de Bunt:
group of 32 university freshmen,
24 female and 8 male students.

Three observations used here $(t_1,\ t_2,\ t_3)$:
at 6, 9, and 12 weeks after the start of the university year.
The relation is defined as a 'friendly relation'.

Missing entries $x_{ij}(t_m)$ set to 0
and not used in calculations of statistics.

Densities increase from 0.15 at t_1 via 0.18 to 0.22 at t_3 .

Very simple model: only density and reciprocity effects

	Model 1	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	3.26	(0.48)
Rate $t_2 - t_3$	2.83	(0.42)
Density	−1.03	(0.19)
Reciprocity	1.76	(0.27)

Overly simplistic model:

rate parameters :
actors make on average about 3 changes between observations;

density parameter negative:
on average, cost of friendly ties higher than their benefits;

reciprocity effect strong and highly significant ($t = 1.76/0.27 = 6.5$).

Objective function is

$$f_i(x) = \sum_j \Big(-1.03\, x_{ij} + 1.76\, x_{ij}\, x_{ji} \Big).$$

This expresses how much actor i likes the network.

Adding a reciprocated tie (i.e., for which $x_{ji} = 1$) gives

$$-1.03 + 1.76 = 0.73.$$

Adding a non-reciprocated tie (i.e., for which $x_{ji} = 0$) gives

$$-1.03,$$

i.e., this has negative benefits.

Conclusion: reciprocated ties are valued positively,
unreciprocated ties negatively;
actors will be reluctant to form unreciprocated ties;
by ‘chance’ (the random term),
such ties will be formed nevertheless
and these are the stuff on the basis of which
reciprocation by others can start.

(Incoming unreciprocated ties, $x_{ji} = 1, x_{ij} = 0$ do not play a role
because for the objective function
only those parts of the network are relevant
that are under control of the actor,
so terms not depending on the outgoing relations of the actor
are irrelevant.)

More adequate structural model: also two network closure effects

	Model 2	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	3.87	(0.57)
Rate $t_2 - t_3$	3.12	(0.48)
Density	−1.45	(0.26)
Reciprocity	1.90	(0.33)
Transitive triplets	0.22	(0.12)
Indirect relations	−0.32	(0.07)

Both network closure effects are significant
(controlling for each other!),
but negative indirect relations effect is stronger.

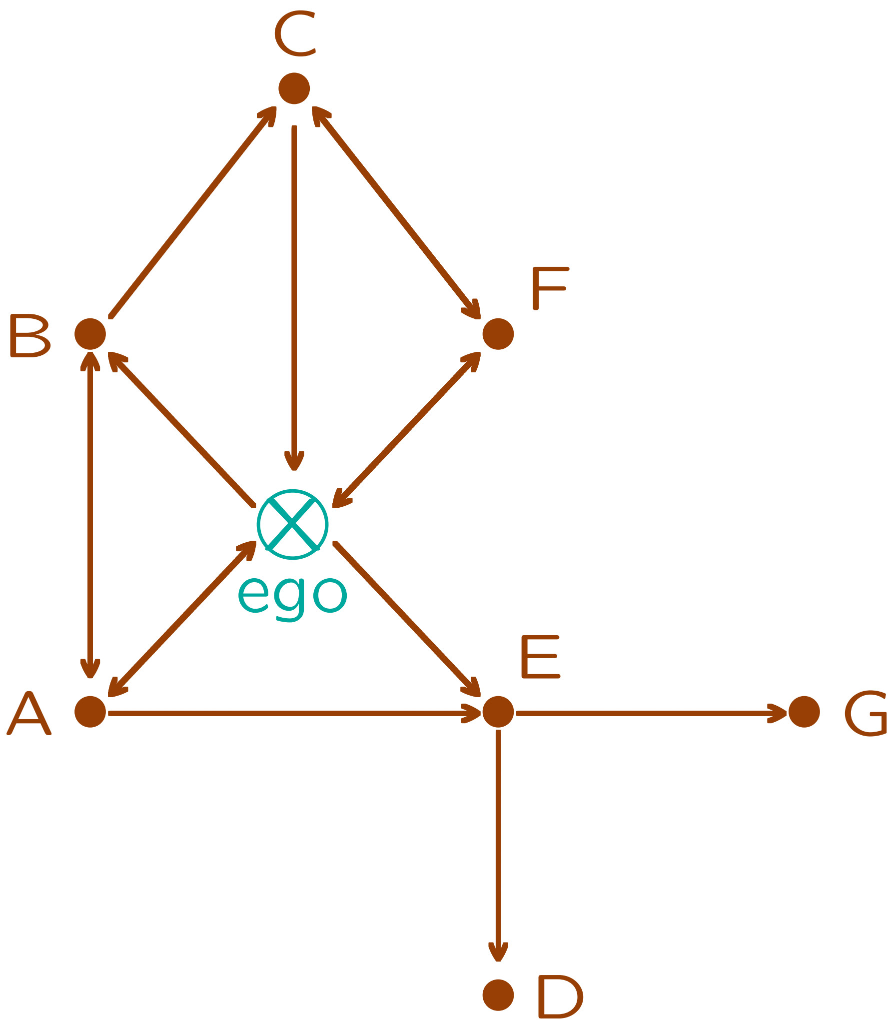
Skip details interpretation

For an interpretation, consider the simpler model without the transitive triplets effect. The estimates are:

Structural model with one network closure effect

	Model 3	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	3.82	(0.60)
Rate $t_2 - t_3$	3.17	(0.52)
Density	−1.05	(0.19)
Reciprocity	2.43	(0.40)
Indirect relations	−0.55	(0.08)

Example: Personal network of ego.



out-degree x_{i+} = 4 ,

#{recipr. ties} = 2,

#{at distance 2} = 3.

Objective function is

$$f_i(x) = \sum_j \Big(-1.05\, x_{ij} + 2.43\, x_{ij}\, x_{ji} - 0.55\, (1 - x_{ij}) \max_h \big(x_{ih}\, x_{hj}\big)\Big)$$

$\Big($ note that $\sum_j (1 - x_{ij}) \max_h \big(x_{ih}\, x_{hj}\big)$

is #{indiv. at distance 2} $\Big)$

so its current value for this actor is

$$f_i(x) = -1.05 \times 4 + 2.43 \times 2 - 0.55 \times 3 = -0.99.$$

Some options for actor i when this actor can make a ministep:

	out-degree	recipr.	distance 2	gain
current	4	2	3	
new tie to C	5	3	2	+1.93
new tie to D	5	2	2	-0.50
new tie to G	5	2	2	-0.50
drop tie to A	3	1	4	-1.93
drop tie to F	3	1	3	-1.38
drop tie to E	3	2	1	+2.15

The actor adds random influences to this (with s.d. 1.28), and chooses the change with the best total value.

Effects of sex and program, smoking similarity

	Model 3	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	4.05	(0.66)
Rate $t_2 - t_3$	3.10	(0.47)
Density	−1.55	(0.24)
Reciprocity	1.83	(0.32)
Transitive triplets	0.21	(0.11)
Indirect relations	−0.30	(0.07)
Sex (M) popularity	0.57	(0.25)
Sex (M) activity	−0.39	(0.27)
Sex similarity	0.35	(0.25)
Program similarity	0.38	(0.13)
Smoking similarity	0.37	(0.20)

Conclusion:
Women more active;
men more popular;
no significant
sex similarity effect.

To interpret the three effects of actor covariate *gender*,
it is more instructive to consider them simultaneously.
Gender was coded originally by with 0 for *F* and 1 for *M*
but this dummy variable was centered (the mean was subtracted)
which led to scores $z_i = -0.25$ for *F* and $+0.75$ for *M*.
The joint effect of the gender-related effects
for the tie variable x_{ij} from i to j is

$$-0.39\, z_i + 0.57\, z_j + 0.35\, I\{z_i = z_j\} \ .$$

$i \ \backslash \ j$	F	M
F	0.03	−0.02
M	−0.50	0.15

Conclusion:
the gender effect is mainly, that men seem not to like female friends.

Skip extended model specification

Extended model specification

1. Gratification function / endowment effect $g_i(\gamma, x, j)$

This represents the "gratification" experienced by the actor when he *makes* a particular *change* in his relations, rather than when he *has* a particular configuration of relations.

Is used to represent models where certain effects work differently for *creation* of ties ($0 \rightarrow 1$) than for *termination* of ties ($1 \rightarrow 0$).

Function $g_i(\gamma, x, j)$ expresses the gratification for i, when starting from the present network structure x, as a consequence of changing his relation with j.

With this extension, actor i chooses to make the relational change that maximizes

$$f_i\Big(\beta, x(i \rightsquigarrow j)\Big) + g_i(\gamma, x, j) + U_i(t, x, j) \ .$$

The gratification function again can be a weighted sum

$$g_i(\gamma, x, j) = \sum_{h=1}^{H} \gamma_h \, r_{ijh}(x) \ .$$

Examples of components of gratification function:

1. $\gamma_1 \, x_{ij} \, x_{ji}$
   γ_1 reflects benefits of breaking off a reciprocated tie (expected to be negative).

2. $\gamma_2 \, (1 - x_{ij}) \sum_h x_{ih} \, x_{hj}$
   number of actors through whom i is indirectly tied to j

3. $\gamma_3 \, x_{ij} \, w_{ij}$
   effect of dyadic covariate w_{ij}
   different for creating than for breaking a tie.

Continuation example :

Reciprocity of a relation can have different effects
for creating than for breaking a relation.

Table next page:
total effect due to reciprocity conducive to *creating* a tie is 1.43;
total effect due to reciprocity
conducive to *breaking* a tie is -1.43 - 1.21 = -2.64.

Conclusion: reciprocity works stronger for terminating
than for creating ties.

Add gratification effect of breaking reciprocated tie

	Model 4	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	4.34	(0.71)
Rate $t_2 - t_3$	3.20	(0.51)
Density	−1.60	(0.24)
Reciprocity	1.43	(0.39)
Transitive triplets	0.21	(0.11)
Indirect relations	−0.29	(0.07)
Sex (M) popularity	0.53	(0.24)
Sex (M) activity	−0.48	(0.27)
Sex similarity	0.34	(0.26)
Program similarity	0.39	(0.14)
Smoking similarity	0.41	(0.19)
Breaking reciprocated tie	−1.21	(0.46)

Skip non-constant rate function

Extended model specification

2. Non-constant rate function $\lambda_i(\alpha, x)$.

This means that some actors change their relations more quickly than others, depending on covariates or network position.

Dependence on covariates:

$$\lambda_i(\alpha, x) = \exp(\sum_h \alpha_h v_{hi}) .$$

(Note that rate function must be positive; ⇒ exponential function.)

Dependence on network position:
e.g., dependence on in- and out-degrees:

$$\lambda_i(\alpha, x) = x_{i+} \exp(\alpha_1) + (n - 1 - x_{i+}) \exp(-\alpha_1) \; .$$

Also, $\sharp$ reciprocated relations of actor i may be used.

Now the parameter is $\theta = (\rho, \alpha, \beta, \gamma)$.

Continuation example

Rate function depends on out-degree:
those with higher out-degrees
also change their tie patterns more quickly.

Gratification function depends on tie recipration
and gender dissimilarity:

$$g_i(\gamma, x, j) = \gamma_1 \, x_{ij} \, x_{ji} + \gamma_2 \, x_{ij} \, | \, w_i - w_j \, |$$

Reciprocity and gender dissimilarity operate differently
for tie initiation than for tie withdrawal.

Parameter estimates model with rate and gratification effects

	Model 4	
Effect	par.	(s.e.)
Rate (period 1)	5.05	
Rate (period 2)	3.95	
Out-degree effect on rate	0.90	(0.47)
Density	-0.99	(0.20)
Reciprocity	2.82	(0.56)
Indirect relations	-0.508	(0.091)
Gender activity	-0.52	(0.31)
Gender popularity	0.55	(0.30)
Gender similarity	-0.08	(0.37)
Breaking recipr. relation	-0.58	(1.06)
Breaking relation with different-gender other	1.64	(0.62)

Conclusion:

weak evidence that actors with higher out-degrees tend to change their relations more often ($t = 0.90/0.47 = 1.91$),

ties with others of the other sex terminated more quickly than ties with others of the same sex ($t = 1.64/0.62 = 2.65$).

no evidence that effect of reciprocation on tie creation differs from the negtiave of its effect on tie formation.

Further continuation example

Add effect out-degrees on rate of change
and also gratification effect of sex dissimilarity
(representing that sex similarity may have different effects for creating than for terminating ties)
(other dissimilarity variables had no significant effects).

	Model 5	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	5.40	(0.82)
Rate $t_2 - t_3$	4.06	(0.63)
Out-degrees effect on rate	0.82	(0.52)
Density	−1.44	(0.25)
Reciprocity	1.76	(0.43)
Transitive triplets	0.18	(0.12)
Indirect relations	−0.30	(0.07)
Sex (M) popularity	0.51	(0.28)
Sex (M) activity	−0.37	(0.28)
Sex similarity	0.10	(0.38)
Program similarity	0.41	(0.13)
Smoking similarity	0.39	(0.19)
Breaking reciprocated tie	−0.49	(0.47)
Breaking tie with other sex	1.39	(0.66)

Conclusion:
gratification effect breaking reciprocated tie becomes non-significant, is taken over for gratification effect dissimilar sex tie.

Deleting (most) non-significant effects one by one leads to following model.

Parsimonious model

	Model 6	
Effect	par.	(s.e.)
Rate $t_1 - t_2$	4.17	(0.69)
Rate $t_2 - t_3$	3.18	(0.53)
Density	−1.52	(0.26)
Reciprocity	1.70	(0.36)
Transitive triplets	0.22	(0.11)
Indirect relations	−0.28	(0.07)
Sex (M) popularity	0.54	(0.23)
Sex similarity	0.01	(0.35)
Program similarity	0.40	(0.13)
Smoking similarity	0.35	(0.19)
Breaking tie with other sex	1.35	(0.61)

Final conclusion:
men are more popular; different-sex ties are less stable.

Non-directed networks

The actor-driven modeling is less straightforward
for non-directed relations,
because two actors are involved in deciding about a tie.

Various modeling options are possible:

1. Forcing model:
   one actor takes the initiative and unilaterally imposes
   that a tie is created or dissolved.

2. Unilateral initiative with reciprocal confirmation:
one actor takes the initiative and proposes a new tie
or dissolves an existing tie;
if the actor proposes a new tie, the other has to confirm,
otherwise the tie is not created.

3. Pairwise conjunctive model:
a pair of actors is chosen and reconsider whether a tie
will exist between them; a new tie is formed if both agree.

4. Pairwise disjunctive (forcing) model:
a pair of actors is chosen and reconsider whether a tie
will exist between them;
a new tie is formed if at least one wishes this.

5. Pairwise compensatory (additive) model:
a pair of actors is chosen and reconsider whether a tie will exist between them; this is based on the sum of their utilities for the existence of this tie.

Option 1 is close to the actor-driven model for directed relations.

In options 3–5, the pair of actors (i, j) is chosen depending on the product of the rate functions $\lambda_i \lambda_j$ (under the constraint that $i \neq j$).

The numerical interpretation of the ratio function differs between options 1–2 compared to 3–5.

The decision about the tie is taken on the basis of the objective functions $f_i \, f_j$ of both actors.

Statistical estimation

Suppose that at least 2 observations on $X(t)$ are available, for observation moments $t_1,\ t_2$.

How to estimate θ?

Condition on $X(t_1)$ *:*
the first observation is accepted as given, contains in itself no observation about θ.

No assumption of a stationary network distribution.

Method of moments :

choose a suitable statistic $Z = (Z_1, \ldots, Z_K)$,
i.e., K variables which can be calculated from the network;
the statistic Z must be *sensitive* to the parameter θ
in the sense that higher values of θ_k
lead to higher values of the expected value $\mathsf{E}_{\widehat{\theta}}(Z_k)$;

determine value of $\theta = (\rho, \beta)$ for which
observed and expected values of suitable statistic are equal.

Questions:

* What is a suitable (K-dimensional) statistic?
  Corresponds to objective function.

* How to find this value of θ?
  By stochastic approximation (Robbins-Monro process)
  based on repeated simulations of the dynamic process,
  with parameter values
  getting closer and closer to the moment estimates.

Skip details.

Suitable statistics for method of moments

Assume first that $\lambda_i(x) = \rho = \theta_1$,
and 2 observation moments.

This parameter determines the expected "amount of change".

A sensitive statistic for $\theta_1 = \rho$ is

$$C = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \mid X_{ij}(t_2) - X_{ij}(t_1) \mid \ ,$$

the "observed total amount of change".

For the weights β_k in the objective function

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

a higher value of β_k means that all actors strive more strongly after a high value of $s_{ik}(x)$, so $s_{ik}(x)$ will tend to be higher for all i, k.

This leads to the statistic

$$S_k = \sum_{i=1}^{n} s_{ik}(X(t_2)) \ .$$

This statistic will be sensitive to β_k :
a high β_k will to lead to high values of S_k.

Moment estimation will be based on the vector of statistics

$$Z = (C, S_1, ..., S_{K-1}) \ .$$

Denote by z the observed value for Z.

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

$$\mathsf{E}_{\widehat{\theta}}\{Z\} = z \ .$$

Robbins-Monro algorithm

The moment equation cannot be solved by analytical or the usual numerical procedures, because

$$\mathsf{E}_\theta\{Z\}$$

cannot be calculated explicitly.

However, the solution can be approximated by the Robbins-Monro (1951) method for stochastic approximation.

Iteration step:

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

where D is a suitable matrix,
and z_N is a simulation of Z with parameter $\hat{\theta}_N$,
and a_N is a sequence $a_N \to 0$.

How to choose matrix D?

Ruppert (1988) and Polyak (1990) proposed
that (under certain conditions)
D can be a positive diagonal matrix
if θ is estimated not by the last value $\hat{\theta}_{N(\max)}$
but by the average of the sequence $\hat{\theta}_N \quad \big(N = 1, \ldots, N(\max)\big)$.

Pflug (1990) proposed
to keep a_N constant during the iterations
until a certain convergence criterion is satisfied,
and decrease it only then.

Covariance matrix

The method of moments yields the covariance matrix

$$\text{cov}(\hat{\theta}) \approx D_\theta^{-1} \, \Sigma_\theta \, D_\theta'^{\,-1}$$

where

$$\begin{aligned} \Sigma_\theta &= \text{cov}\{Z \,|\, X(t_1) = x(t_1)\} \\ D_\theta &= \frac{\partial}{\partial \theta} \text{E}\{Z \,|\, X(t_1) = x(t_1)\} \, . \end{aligned}$$

(Note: Z is function of $X(t_1)$ and $X(t_2)$).

After the presumed convergence of the algorithm
for approximately solving the moment equation,
extra simulations are carried out

(*a*) to check that indeed $\mathsf{E}_{\hat{\theta}}\{Z\} \approx z$,

(*b*) to estimate Σ_θ,

(*c*) and to estimate D_θ
using a score function algorithm
(earlier algorithm used
difference quotients and common random numbers).

Skip conditional estimation

Modified estimation method:

conditional estimation .

Condition on the observed numbers of differences between successive observations,

$$c_m = \sum_{i,j} |\, x_{ij}(t_{m+1}) - x_{ij}(t_m) \,| \quad .$$

For continuing the simulations do not mind the values of the time variable t,
but continue between t_m and t_{m+1} until the observed number of differences

$$\sum_{i,j} \mid X_{ij}(t) - x_{ij}(t_m) \mid$$

is equal to the observed c_m .
This is defined as time moment t_{m+1} .

This procedure is a bit more stable;
requires modified estimator of ρ_m .

Summary of estimation algorithm

3 phases:

1. brief phase for preliminary estimation of $\partial \mathrm{E}_{\hat{\theta}}\{Z\}/\partial\theta$ for defining D;
2. estimation phase with Robbins-Monro updates, where a_N remains constant in *subphases* and decreases between subphases;
3. final phase where θ remains constant at its estimated value; this phase is for checking that

$$\mathrm{E}_{\hat{\theta}}\{Z\} \approx z \ ,$$

and for estimating D_θ and Σ_θ to calculate standard errors.

The procedure is implemented in the program

S imulation
I nvestigation for
E mpirical
N etwork
A nalysis

(new beta version 2.4) which can be downloaded from
http://stat.gamma.rug.nl/snijders/siena.html
(programmed by Tom Snijders, Christian Steglich, Michael Schweinberger, Mark Huisman).

A Windows shell is contained in the StOCNET package
(new version 1.7)
developed by Peter Boer
(contributions by Bert Straatman, Evelien Zeggelink, Mark Huisman, Christian Steglich)

http://stat.gamma.rug.nl/stocnet/

Further explanation :

Tom A.B. Snijders,
"The Statistical Evaluation of Social Network Dynamics".
Sociological Methodology 2001, 361–395;

—, "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 (2005).

Mark Huisman and Tom A.B. Snijders,
"Statistical analysis of longitudinal network data with changing composition".
Sociological Methods & Research, 32 (2003), 253-287.

—, "Een stochastisch model voor netwerkevolutie".
Nederlands Tijdschrift voor de Psychologie, 58 (2003), 182-194.

SIENA manual.

2. Networks as dependent & independent variables

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

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

Two-way influence between networks and behavior

Relational embeddedness is important
for well-being, behavior, opportunities, etc.;
cf. studies of *social capital*.

Also, actors are influenced in their behavior and attitudes
by other actors to whom they are tied

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

In addition, 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 behavior and other characteristics
(similarity, opportunities for future rewards, etc.).

Influence, network effects on behavior;
Selection, behavior effects on relations.

Skip extra discussion

Terminology:

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

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

Examples:

⇒ Risky social behaviors (like smoking, taking alcohol or drugs) are ’contagious’ among friends but also operative in friendship formation.

⇒ How hard pupils and employees work often is subject to social control.

⇒ Firms choose partners for collaboration based on complementary expertise, reputation, trust, etc.

Thus, there is a feedback relation in the dynamics of relational networks and actor behavior / performance.

The investigation of such social feedback processes is difficult:

* Both the *network ⇒ behavior* and the *behavior ⇒ network* effects lead to an association between current behavior and network: "friends of smokers are smokers" (cf. work by Baumann, Kirke), "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 behaviors at a few discrete time points.
Such an observation design is the point of departure here.

Data:

One bounded set of actors
(e.g. school class, group of professionals, set of firms);

several observation moments (at least 3);

for each observation moment:
⇒ network: who is tied to whom
⇒ behavior of all actors

Aim: disentangle effects *networks* ⇒ *behavior*
from effects *behavior* ⇒ *networks*.

Statistical Methodology for the simultaneous evolution of networks $X(t)$ and behavior $Z(t)$.

Integrate the *influence* (network ⇒ characteristics)
and *selection* (characteristics ⇒ network) processes.

Notation:

In addition to the network X, associated to each actor i
there is a vector $Z_i(t)$ of actor characteristics
indexed by $h = 1, \ldots, H$.
For the moment: ordered discrete
(simplest case: one dichotomous variable).

Actor-driven models :

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

Network change process and behavior 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_h^Z for changes in behavior h),
the actors may change a tie or a behavior.

The actors try to attain a rewarding configuration
of the network and the behaviors
expressed in *objective functions* f^X and f^Z .

Again, conditional independence,
given the current network structure.

Functions λ^X, λ^Z, f^X, and f^Z depend on
K-dimensional statistical parameter $\theta \in \Theta \subset \mathbb{R}^K$.

Skip details

Mini-step for change in network:

At random moments occurring at a rate λ^X,
a random actor is designated to make a change in one tie variable:
the *mini-step* (on ⇒ off, or off ⇒ on.)

Mini-step for change in behavior:

At random moments occurring at a rate λ_h^Z,
a random actor is designated to make a change in behavior h
(one component of Z_i, assumed to be ordinal):
the *mini-step* is a change to an adjacent category.

Again, many mini-steps can *accumulate* to big differences
between consecutive observations.

When actor i ‘may’ change an outgoing tie to some other actor j,
he/she chooses the ’best’ j by maximizing
the objective 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 behavior h,
he/she chooses the “best” change (up, down, nothing)
by maximizing the objective function $f_i^Z(\beta, x, Z)$ of the situation
obtained after the coming behavior change
plus a random component representing unexplained influences.

The new network is denoted by $x(i \rightsquigarrow j)$.
The attractiveness of the new situation
(objective function plus random term)
is expressed by the formula

$$f_i^X(\beta, x(i \rightsquigarrow j), z) \;+\; U_i^X(t, x, j)\,.$$

$\Uparrow$

random component

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

Whenever actor i may make a change in variable h of Z,
he changes only one behavior, 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 objective function of the situation obtained
after the coming behavior change
plus a random component representing unexplained influences:

$$f_i^Z(\beta, x, z(i, h \rightsquigarrow v)) \;+\; U_i^Z(t, z, h, v)\,.$$

⇑

random component

(behavior is permitted to stay the same.)

For the behaviors, the formula of the change probabilities is

$$p_{ihv}(\beta, z) = \frac{\text{exp}(f(i,h,v))}{\sum_{k,u} \text{exp}(f(i,k,u))}$$

where

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

Again, multinomial logit form.

Specification of the behavior model

There are many different reasons
why networks are important for behavior:

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

In this presentation, only imitation is considered,
but the other two reasons are also of eminent importance.

Basic effects for dynamics of behavior f_i^Z:

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

1. *tendency* ,
   $s_{i1}(x, z) = z_{ih}$
2. *covariate-related similarity*,
   sum of covariate similarities
   between i and his friends
   $s_{i2}(x, z) = \sum_j x_{ij} \left(1 - | \, z_{ih} - z_{jh} \, | \right)$,
   if Z_h assumes values between 0 and 1.

3. *popularity-related tendency*,
   $s_{i3}(x, z) = z_{ih}\, x_{+i}$;
4. *activity-related tendency*,
   $s_{i4}(x, z) = z_{ih}\, x_{i+}$;
5. *dependence on other behaviors* $(h \neq h')$,
   $s_{i5}(x, z) = z_{ih}\, z_{ih'}$

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

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

sum of absolute covariate differences between i and his friends
$s_{i2}(x,z) = \sum_j x_{ij} \Big(1 - |\, z_{ih} - z_{jh} \,|\Big)$.

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

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*).

Skip details

A comparison of these two ways of change can give us some insight into the question whether the objective functions f^X and f^Z should be the same.

In the first place, note that a common purpose for the network and the individual behavior does not imply that the coefficients β_k^X and β_k^Z should be the same:
the random utility formulation compares all effects
with each other and with all unmeasured effects (random term),
and this random influence might be larger
for changes in networks than for changes in behaviors.

Therefore, the question is whether
the two sets of coefficients are proportional.

Consider two behaviors Z_1 and Z_2,
for both of which the actors prefer positive network autocorrelation.

It is very wel conceivable that actors attempt to reach this
by network changes for Z_1 and behavior changes for Z_2.

E.g., for a friendship relation:
let Z_1 be religion and Z_2 musical taste.

E.g., for cooperation between firms:
let Z_1 be location and Z_2 administrative organization.

Since the functions f^X and f^Z
express the trade-off between rewards and costs,
even when the rewards are the same,
the costs of change may well be different.

Statistical estimation

Procedures for estimating parameters in this model
are analogous 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 $\Rightarrow$ behavior and behavior $\Rightarrow$ networks,
both lead to
a contemporaneous network autocorrelation of behavior;
but they can be (in principle)
distinguished empirically by the time order: respectively
association between ties at t_m and behavior at t_{m+1};
and association between behavior 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 behavior dynamics:

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

Example :

Study of smoking initiation and friendship
in a Scottish secondary school
(following up on earlier work by P. West, M. Pearson & others,
see Pearson & Michell, *Drugs: educ., prev. and policy*, 2000.)

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

What does this data set tell us about the mutual effects
between friendship and smoking?

First, results for a model
for dynamics in networks and in smoking behavior
under the assumption that both are unrelated.

Smoking measured in three categories:
1 = no, 2 = occasionally, 3 = regularly.

There is more network change than behavior change;
⇒ more power for discovering effects acting on network dynamics.

In this group, girls smoke more than boys.

Parameter estimates for network dynamics:
assumed independent of smoking.

	Effect	Estimate	Standard error
Rate function			
λ_0^X	Rate parameter t_1–t_2	11.63	1.42
λ_1^X	Rate parameter t_2–t_3	9.27	0.96
Objective function			
β_1^X	Density / out-degree	−2.16	0.06
β_2^X	Reciprocity	2.03	0.08
β_3^X	Number of distances 2	−0.68	0.014
β_4^X	Transitive triplets	0.22	0.011
β_5^X	Gender (F) popularity	−0.23	0.08
β_6^X	Gender (F) activity	0.17	0.08
β_7^X	Gender similarity	0.82	0.11

Parameter estimates for smoking dynamics:
assumed independent of friendship.

	Effect	Estimate	Standard error
Rate function			
λ_0^Z	Rate parameter t_1–t_2	0.68	0.22
λ_1^Z	Rate parameter t_2–t_3	0.72	0.24
Objective function			
β_1^Z	Tendency	−0.26	0.24
β_2^Z	Gender (F)	2.05	1.16
β_3^Z	Parents' smoking	0.70	1.20

Now a similar analysis,
but with a model in which there is a mutual effect
between smoking and friendship formation.

Parameter estimates for network dynamics:
dependent on smoking.

	Effect	Estimate	Standard error
Rate function			
λ_0^X	Rate parameter t_1–t_2	11.74	1.25
λ_1^X	Rate parameter t_2–t_3	9.53	1.07
Objective function			
β_1^X	Density / out-degree	−2.17	0.05
β_2^X	Reciprocity	2.06	0.08
β_3^X	Number of distances 2	−0.80	0.013
β_4^X	Transitive triplets	0.17	0.009
β_5^X	Gender (F) popularity	−0.20	0.08
β_6^X	Gender (F) activity	0.18	0.08
β_7^X	Gender similarity	0.80	0.09
β_8^X	Smoking similarity	0.17	0.05

Parameter estimates for smoking dynamics: dependent on friendship.

	Effect	Estimate	Standard error
Rate function			
λ_0^Z	Rate parameter t_1–t_2	0.84	0.30
λ_1^Z	Rate parameter t_2–t_3	0.84	0.27
Objective function			
β_1^Z	Tendency	0.13	0.35
β_2^Z	Gender (F)	1.26	1.14
β_3^Z	Parents' smoking	1.00	1.28
β_4^Z	Friendship	0.33	0.37

Conclusions :

Evidence for effect of smoking on friendship development;
no evidence for effect of friendship on smoking initiation;
note that
taking the mutual effect of smoking and friendship into account
(even though the effect friendship $\Rightarrow$ smoking was not significant)
reduces strongly the estimated effect of gender:

in the first analysis
there seemed some evidence for a gender effect on smoking,
but this can equally be explained as an effect of friendship
(friendship is rather gender-homogeneous).

Discussion issues

⇒ the *fit* of these models
and the *robustness* of the conclusions;

⇒ to what extent is this *causal* modeling?
very modest claims: “as if” approach,
we are describing data & testing substantive theories
using models that express causality;
effects of explanatory variables are ‘maximally’ controlled
for structural network effects;

⇒ richer modeling of network effects is important:
e.g., of network positions of individual actors
(cf. Pearson & West, *Connections*, 2003.)

Further points (current & future work) :

1. Goodness of fit (paper Michael Schweinberger).
2. Maximum likelihood and Bayesian estimation.
3. Explained variation (‘R^2’) (paper Tom Snijders).
4. Non-directed relations.
5. Random actor effects.
6. Random effects multilevel network models.
7. Relations with more than 2 ordered values.
8. Multivariate relations.
9. Fit diagnostics.