Modelling Network Dynamics with SIENA

Christian Steglich and Mark Huisman

University of Groningen

Workshop at Sunbelt XXVI
April 26, 2006

Evolution of social networks:
structures of relations between individuals (actors)
that evolve over time.

Dynamics of social networks

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

Longitudinal modeling offers more promise
for understanding of network structure.
Evolving networks can show the rules of relation choices.

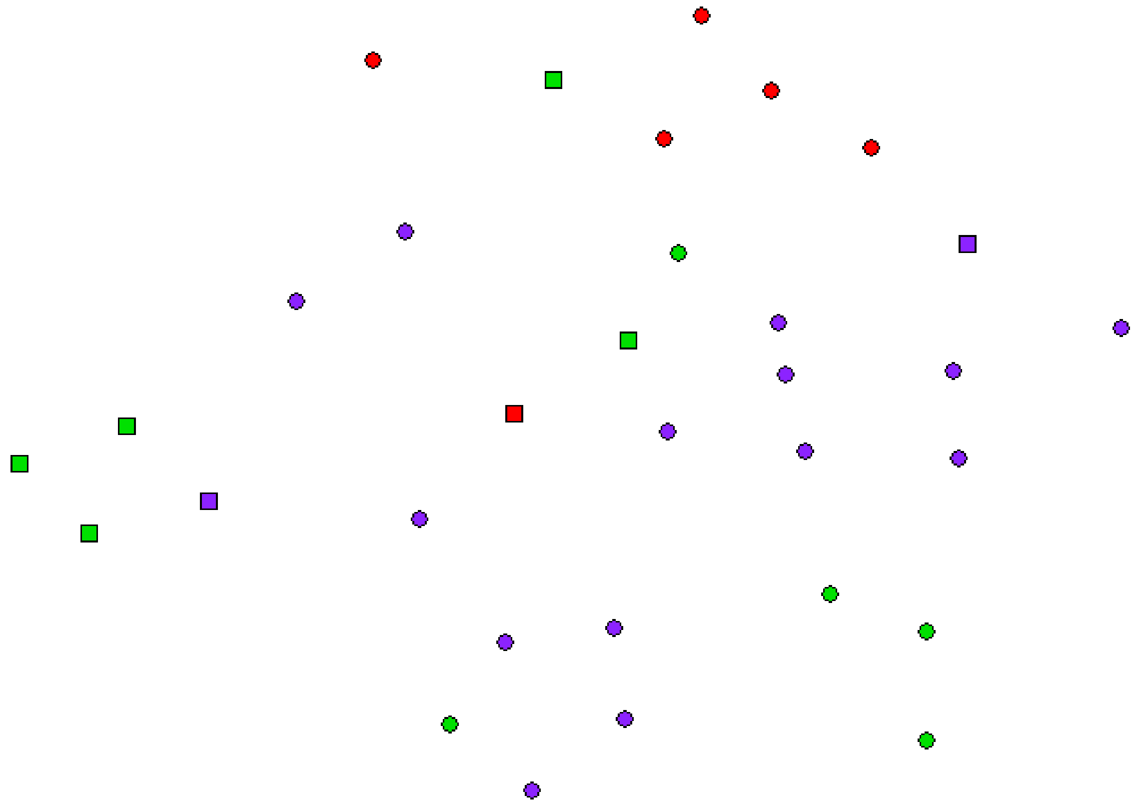
Example data

Van de Bunt's friendship network

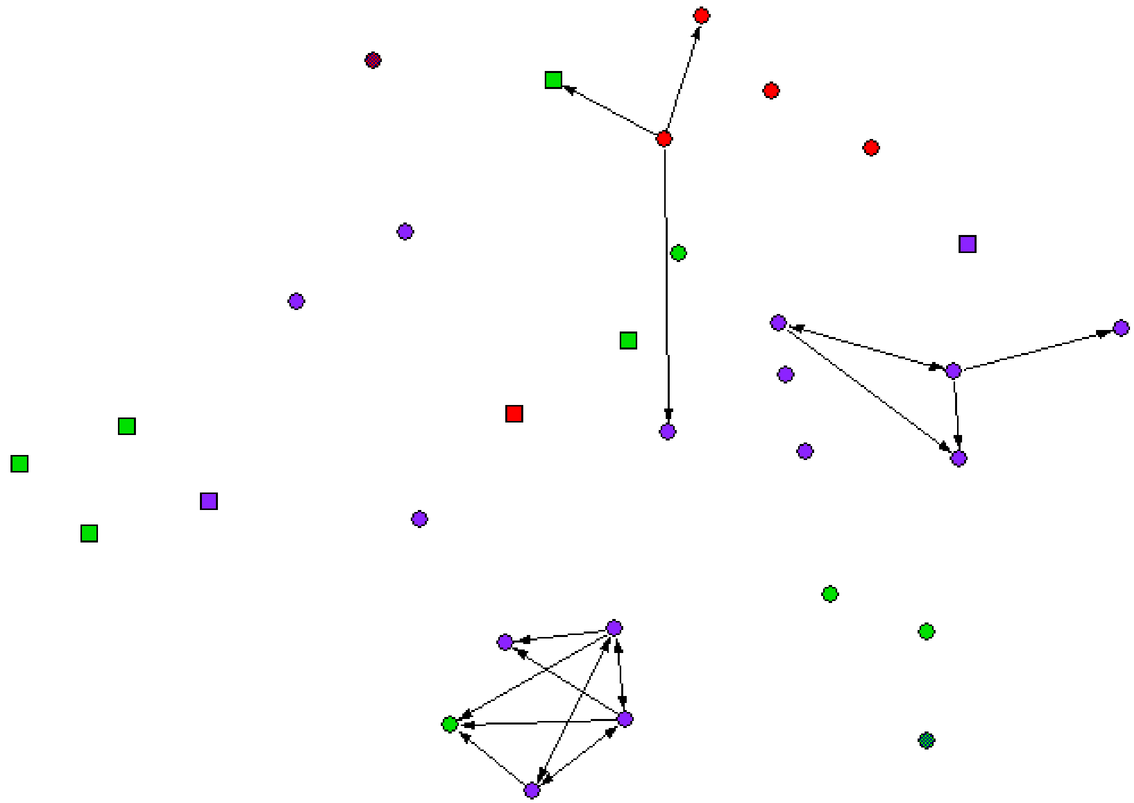
- 32 actors: freshman university students
- Tie = *friendship*:
  (best) friend vs. friendly relation/known/unknown
- Actor covariates: progran (color), gender (shape), smoking
- 7 measurements, several weeks/months apart

Van de Bunt, Van Duijn, & Snijders (1999), *Computational & Mathematical Organization Theory*, *5*, 167–192.

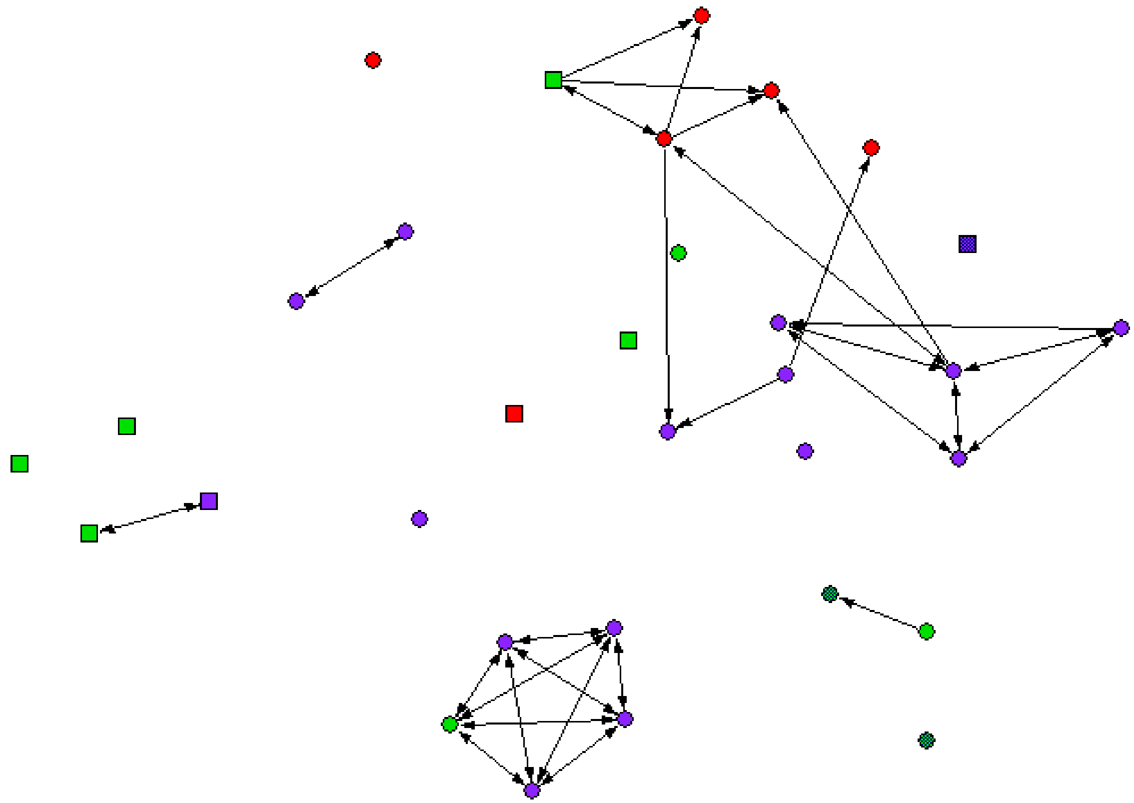
Exploration by visualization (with Pajek software)



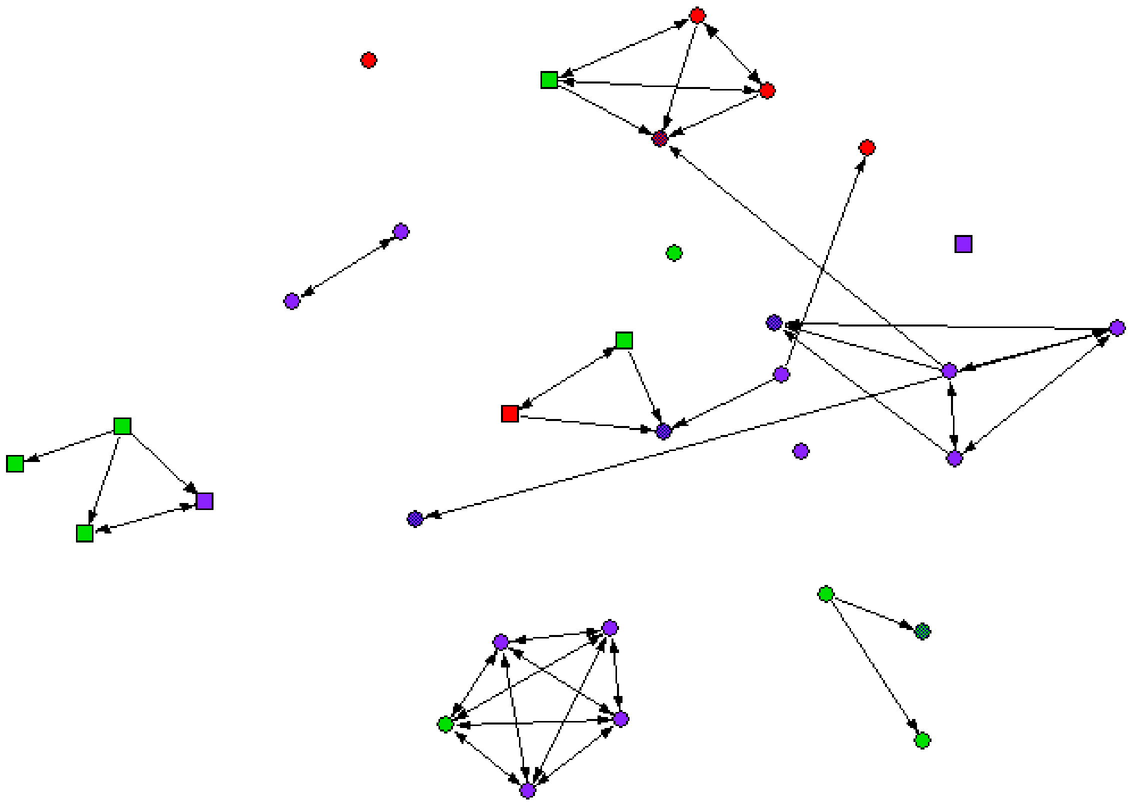
Friendship network time 0
Average degree 0.0, missing fraction 0.00



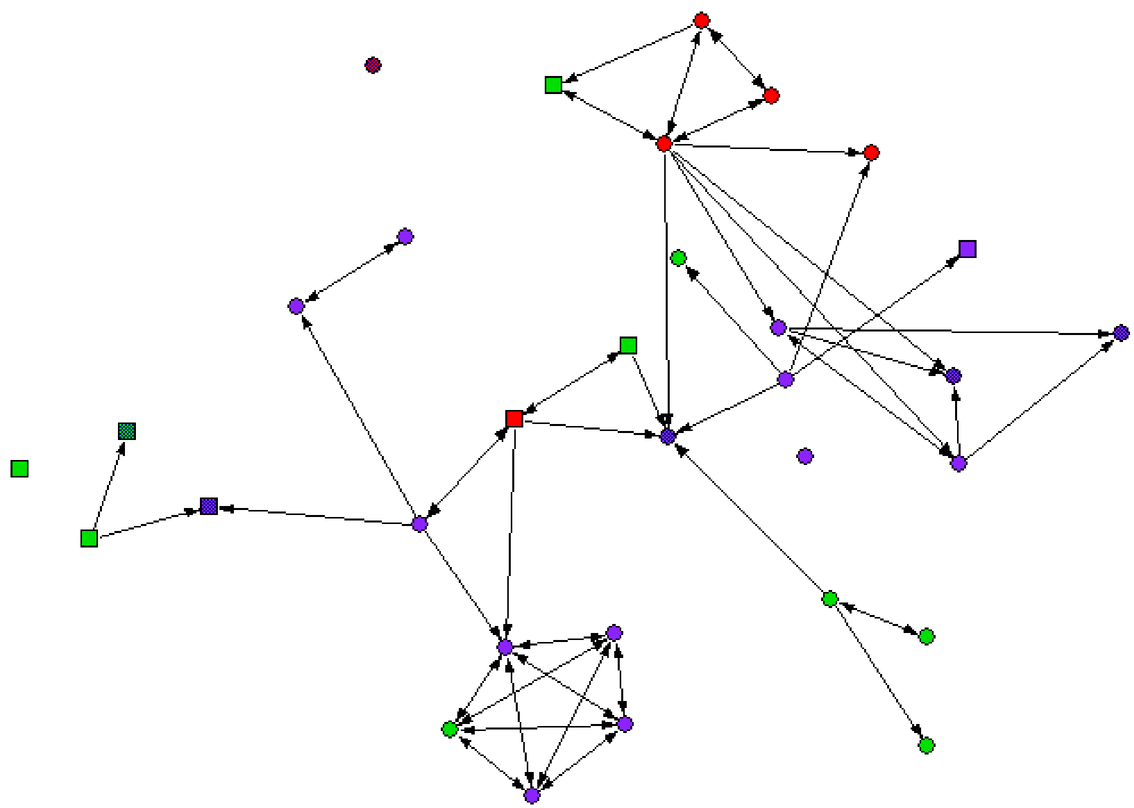
Friendship network time 1
Average degree 0.6, missing fraction 0.06



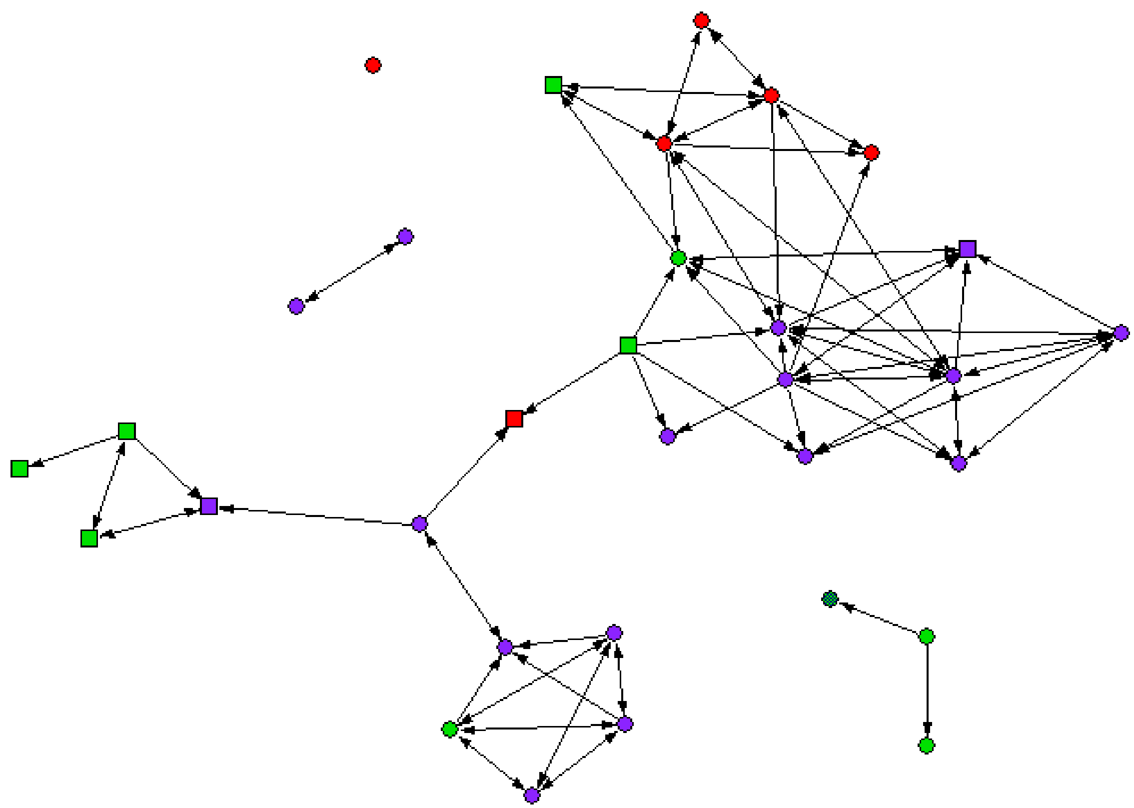
Friendship network time 2
Average degree 1.6, missing fraction 0.09



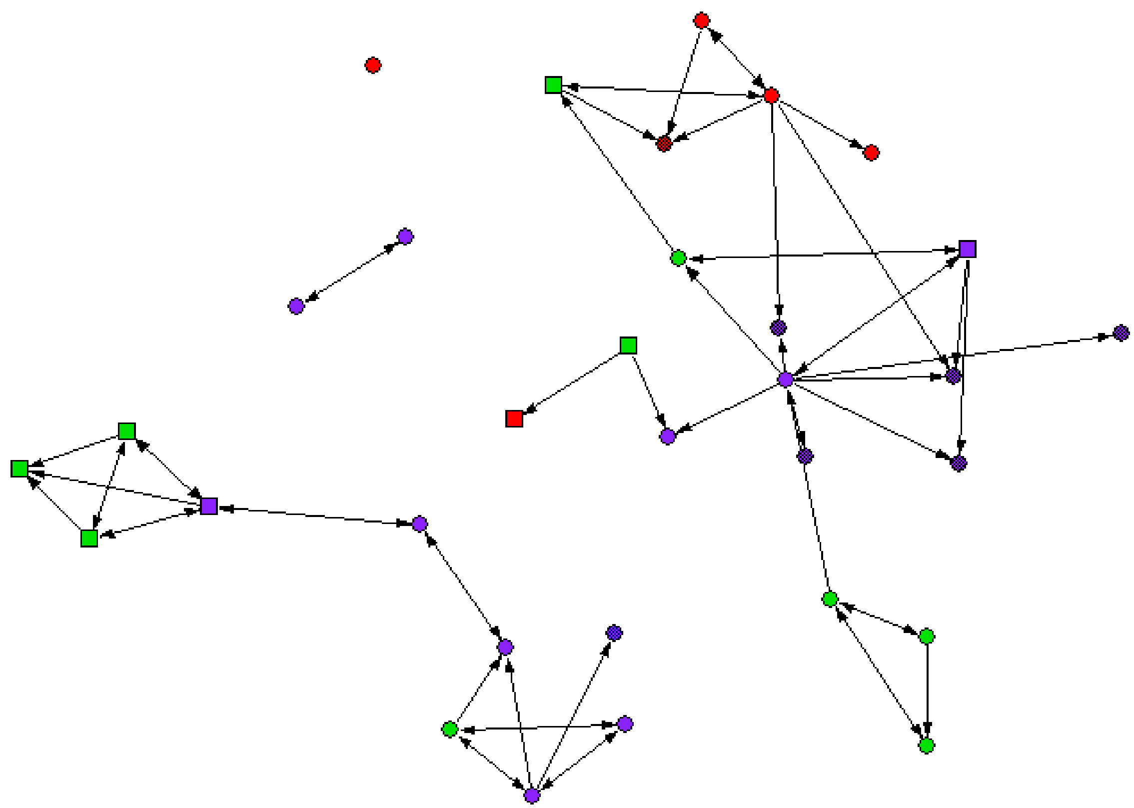
Friendship network time 3
Average degree 2.0, missing fraction 0.16



Friendship network time 4
Average degree 2.4, missing fraction 0.19



Friendship network time 5
Average degree 2.8, missing fraction 0.04



Friendship network time 6
Average degree 2.2, missing fraction 0.22

StOCNET - module Examine

Change statistics

```
--------------------------------------------------------
Numbers of changes between subsequent observations
obs. times   0->0  0->1  1->0  1->1  Distance  Missing
  1 to 2      911    19     0     0        19    62 ( 6%)
  2 to 3      820    29     0    19        29   124 (13%)
  3 to 4      690    14     1    39        15   248 (25%)
  4 to 5      633     6     3    40         9   310 (31%)
  5 to 6      685    22    13    46        35   226 (23%)
  6 to 7      683     8     8    45        16   248 (25%)
--------------------------------------------------------
```

Distance = total number of changes

Also changes in dyads (MAN) and triplets (transitivity)

StOCNET - module Examine

Descriptives per observation time

Dyads and triplets

Network	0	1	2	3	4	5	6
Complete dyads	496	435	406	351	325	456	300
% Mutuals	0.000	0.009	0.047	0.054	0.058	0.061	0.050
Triplets*	0	19	100	92	119	249	77
% Transitive	0.000	0.842	0.840	0.957	0.681	0.554	0.442

* (i->j), (j->h), and (i,h) non-missing

Purpose of statistical modeling:
investigate network evolution as function of

1. structural (network) effects like reciprocity or transitivity
2. explanatory actor variables like program, gender, smoking behavior
3. (possibly explanatory dyadic variables)

All effects control for each other effect

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)

Modeling network dynamics

Data: *repeated measurements of a social network*

Principles:

1. Condition on the first observation (refrain from modeling it)
2. Regard the observations as discrete observations of a process developing in continuous time, where actors can make unobserved changes between the observation moments, being each others' changing environment

$\Rightarrow$ The actors change the network *AND* the changing network acts as a dynamic constraint on each actors' behavior

$\Rightarrow$ Treat network dynamics as an *endogenous dynamic process*

Stochastic actor-oriented models

(Snijders, 1996, 2001, 2005)

Model the process as a series of *micro-steps*:
a sequence of small unobserved changes resulting in differences between two observed networks

At random moments, a random actor takes a micro-step:
make a change in one relation x_{ij} (dichotomous)
(on $\Rightarrow$ off, off $\Rightarrow$ on, or leave unchanged)

Many micro-steps *accumulate* to big differences between consecutive observations of the network

Model specification

To model the micro-steps two functions are needed:

1. The *rate function*
   indicating rate at which an actor may take a micro-step
   (*rate of change*)
2. The *objective function*
   indicating preference of an actor for the
   current state of the network

Simple specification:

- The *rate function* is *constant*: ρ_m
  in time period (t_m, t_{m+1})

- The *objective function* is a weighted sum of network and covariate effects

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

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

Actors try to attain a rewarding position in the network

Aim: achieve high value of the objective function $f_i(\beta, x)$

Actions are also propelled by a *random component* expressing *unexplained change* ('residual term'):
maximize $f_i(\beta, x(i \rightsquigarrow j)) + U_i(t, x, j)$

Objective function

Is the focus of the model construction

Is a weighted sum of effects reflecting

1. structural (network) effects (endogenous):
   density, reciprocity, transitivity, popularity, etc.
2. actor-dependent covariate effects (exogenous):
   covariate-related popularity, activity, or similarity
3. dyad-dependent covariate effects (exogenous)

Stochastic actor-oriented models

Each actor “controls” his/her outgoing relations

At any given moment, with a given current network structure, the actors act independently (i.e. take *micro-steps*), one-at-a-time

No strategy: objective function reflects short-term goals, opportunities, constraints

The stochastic moments are governed by the *rate function*
The changes (micro-steps) are governed by the *objective function*

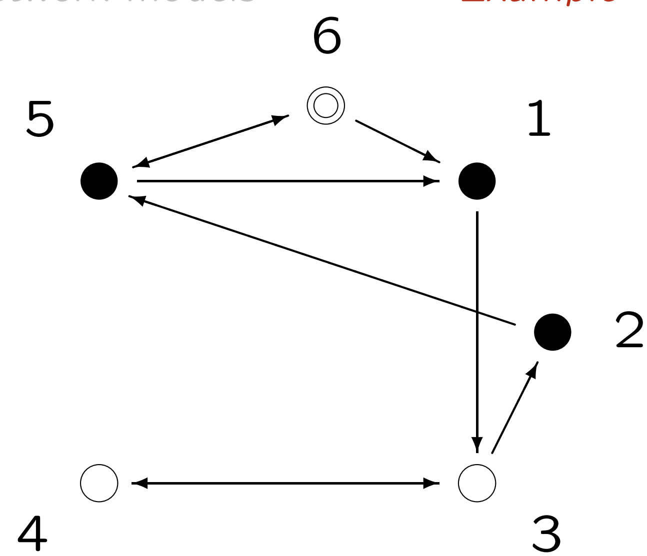
Simulation of networks

The specification of the rate and objective function implies that the sequence of micro-steps can be modeled as a *continuous-time Markov chain*

Starting from the first observation (considered given), and given some estimated value of the parameters, a *Markov-chain of networks* can be simulated using the specified distributions

This setup allows also for actors to join or leave the network at fixed time points (modeled as exogenous events)

T_1



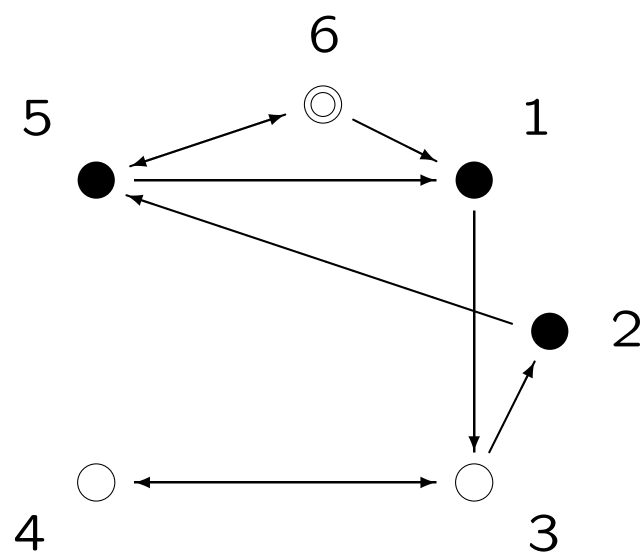
1. *Constant rate*: ρ.
2. *Objective function* includes *density effect* s_1, *reciprocity effect* s_2, and *covariate-related similarity* s_9:

$$f_i(\widehat{\beta}, x) = \widehat{\beta}_1 \, s_{i1}(x) + \widehat{\beta}_2 \, s_{i2}(x) + \widehat{\beta}_9 \, s_{i9}(x)$$

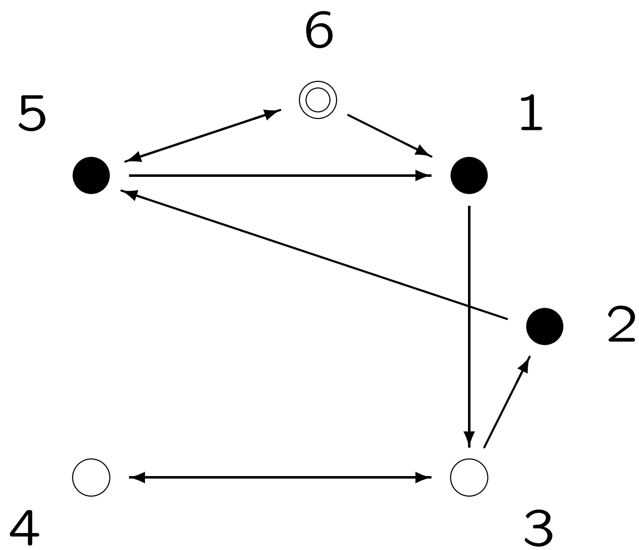
3. *Change in composition*: at time point 0.25 actor 6 leaves at time points 0.50, and 0.75 actors 7 and 8 join

$\Rightarrow$ given values of $\widehat{\beta}$: *Simulate evolution*

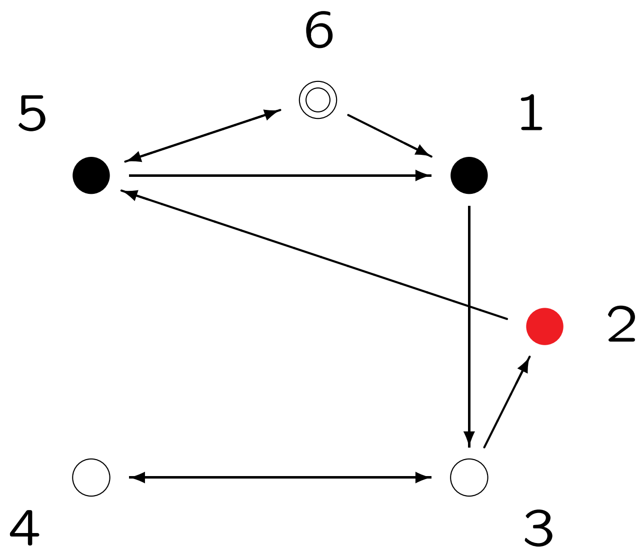
$t = 0$



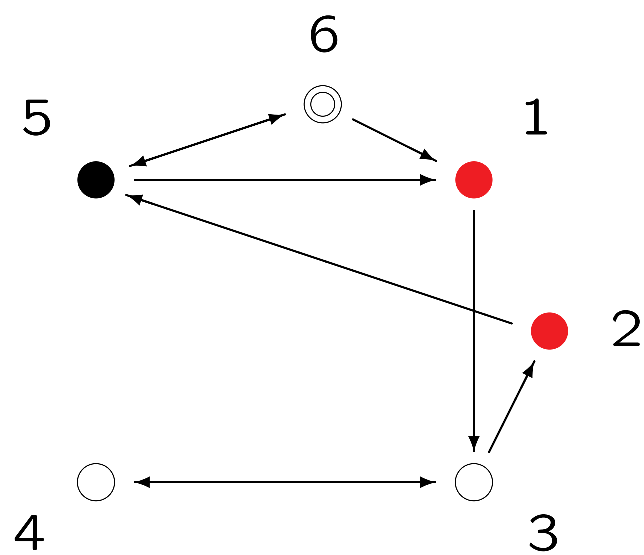
$t = 0.03$



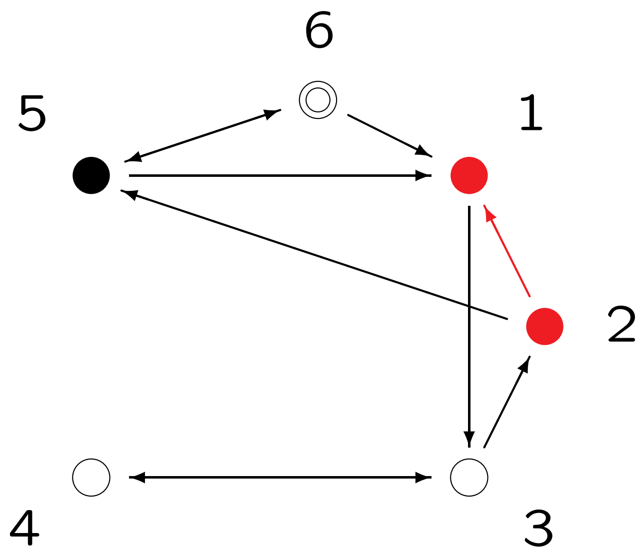
$t = 0.03$



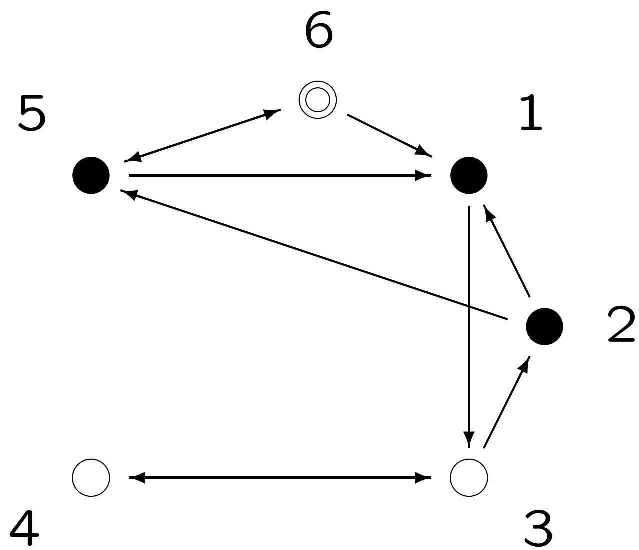
$t = 0.03$



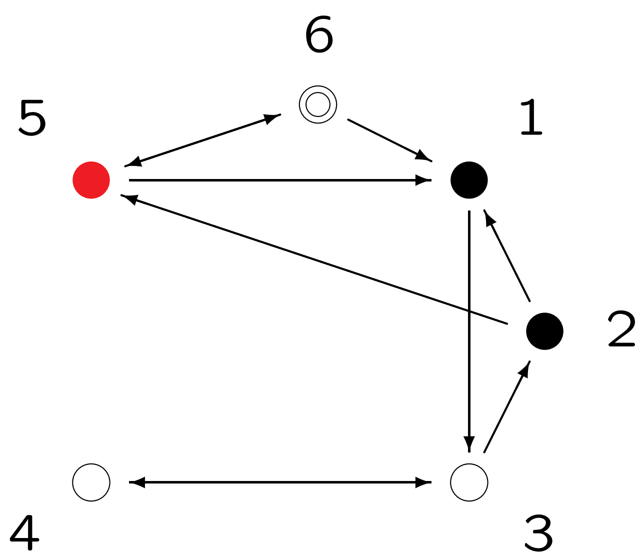
$t = 0.03$



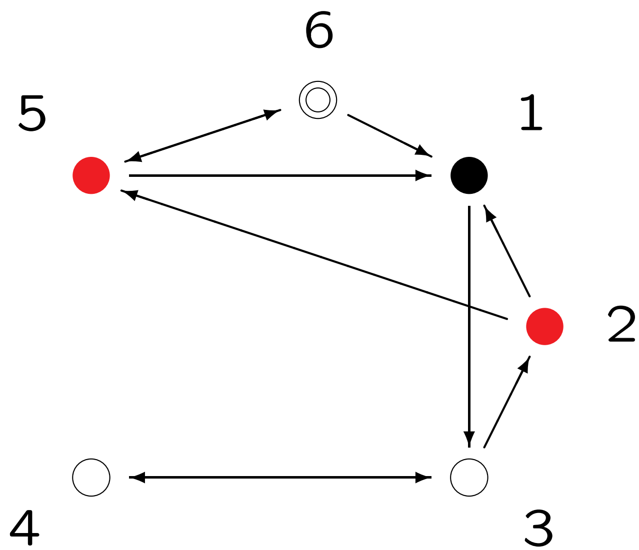
$t = 0.09$



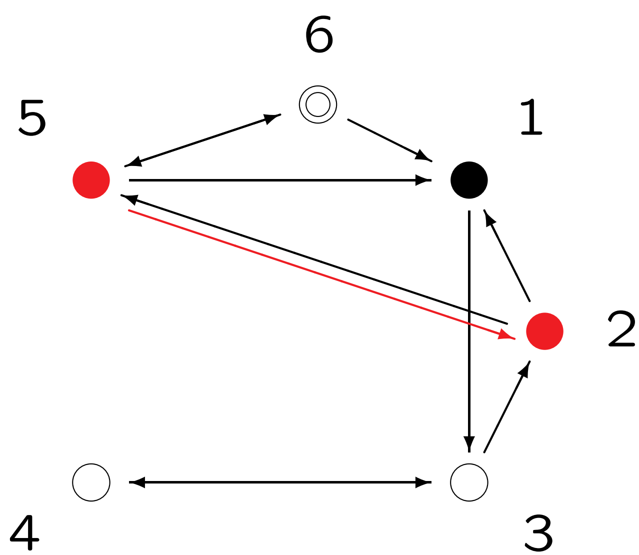
$t = 0.09$



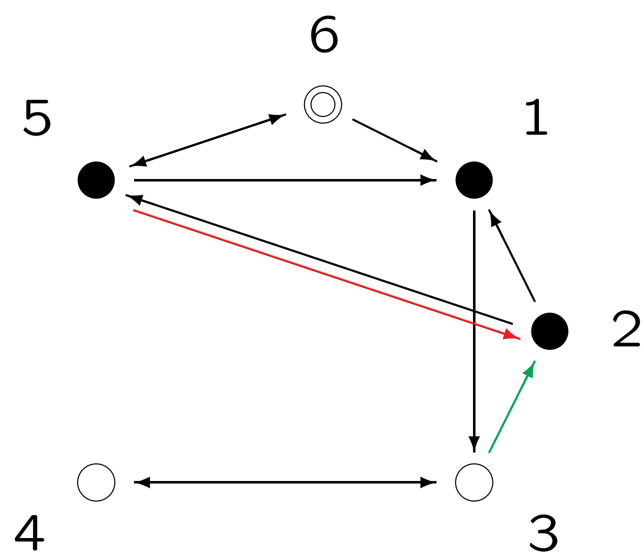
$t = 0.09$



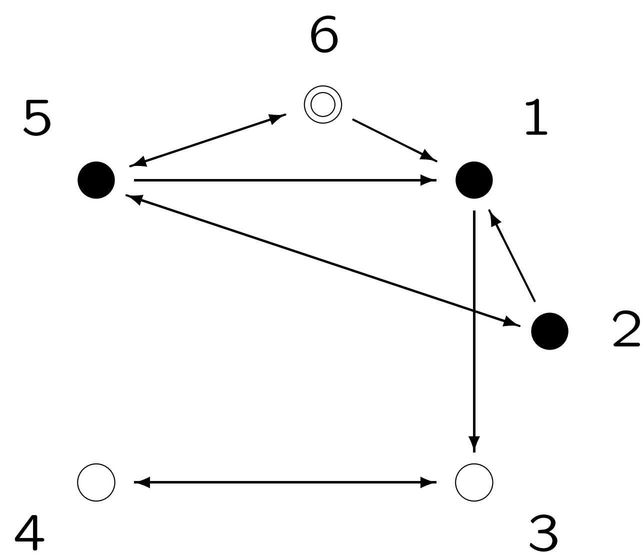
$t = 0.09$



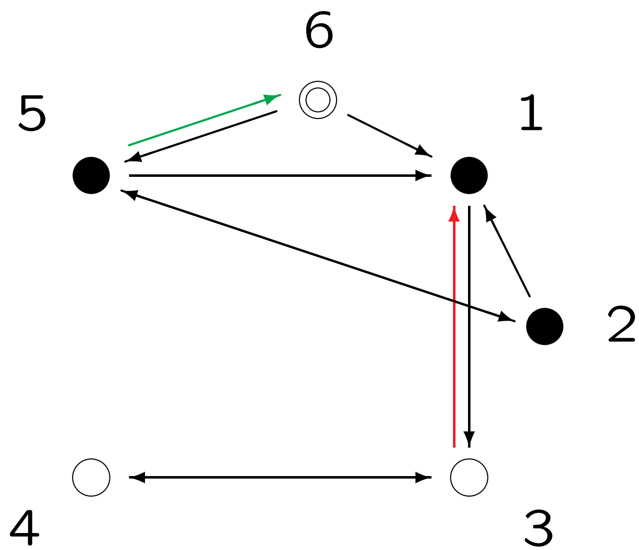
$t = 0.09$



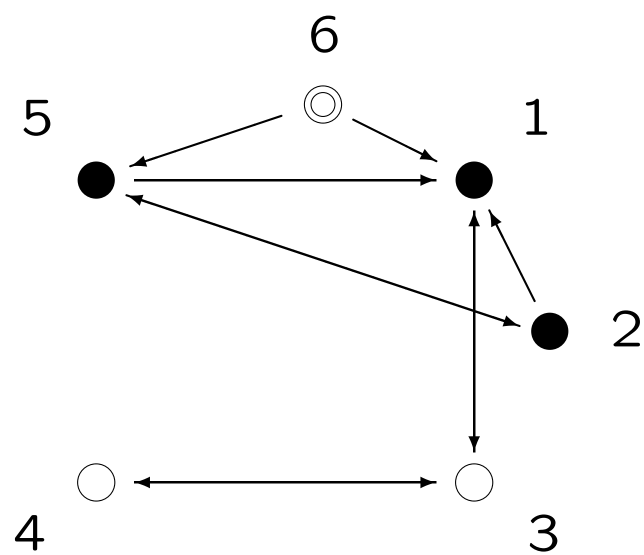
$t = 0.11$



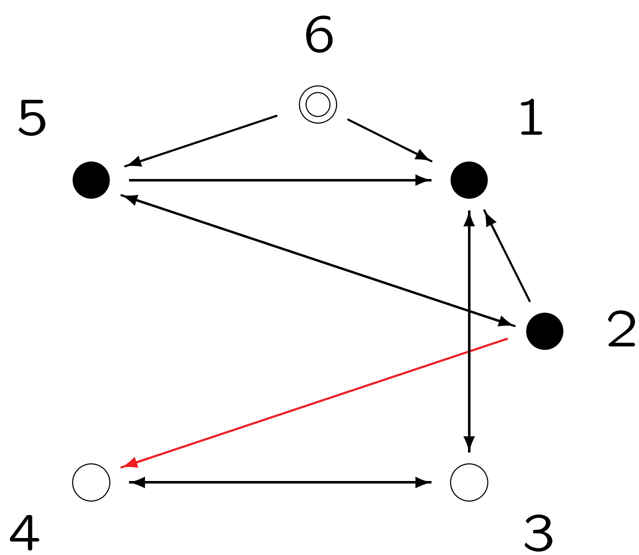
$t = 0.19$



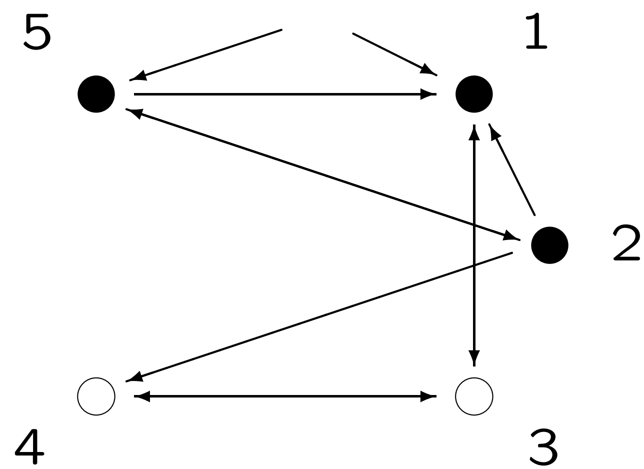
$t = 0.20$



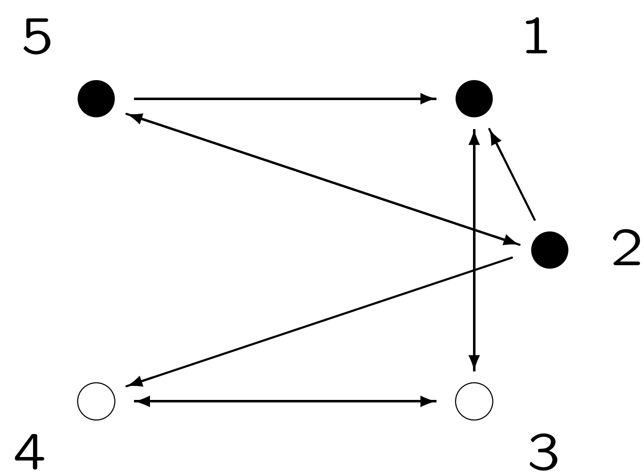
$t = 0.24$



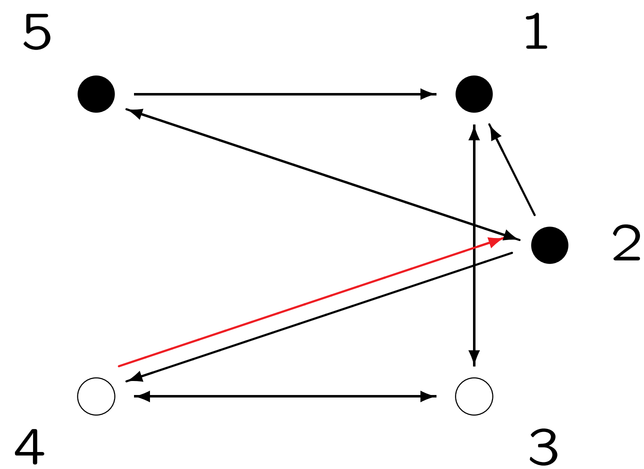
$t = 0.25$



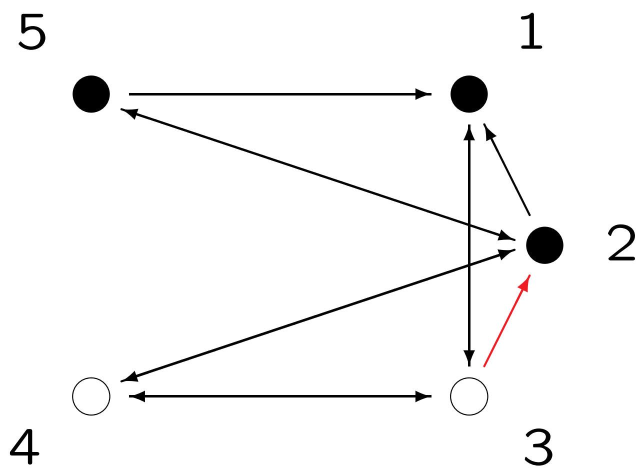
$t = 0.25$



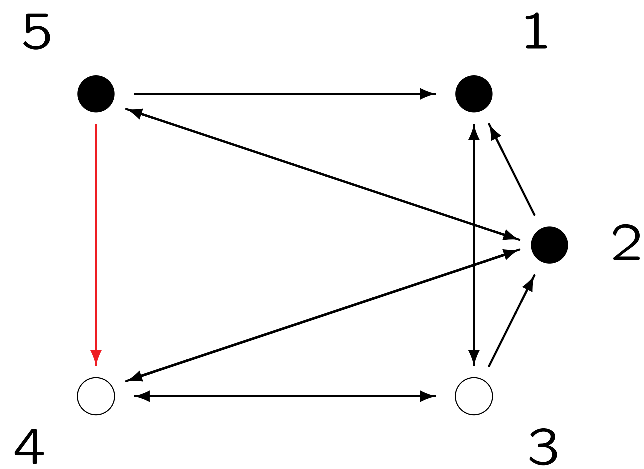
$$t = 0.28$$



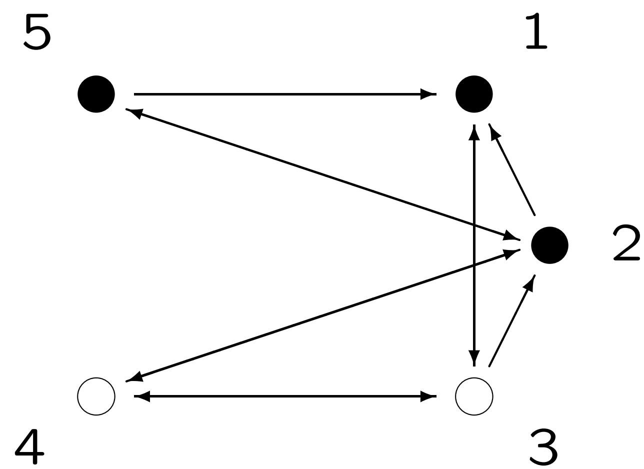
$t = 0.30$



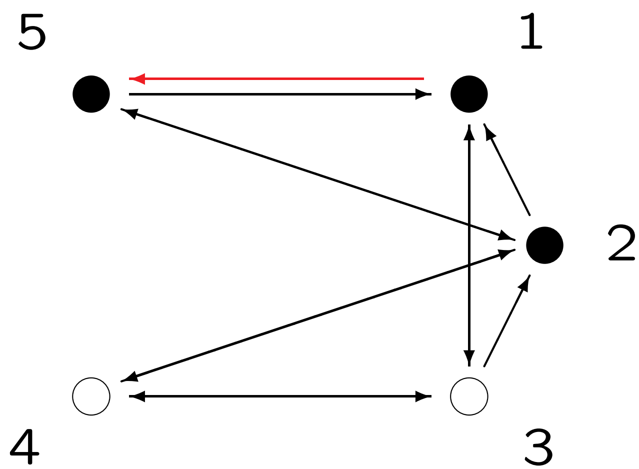
$t = 0.41$



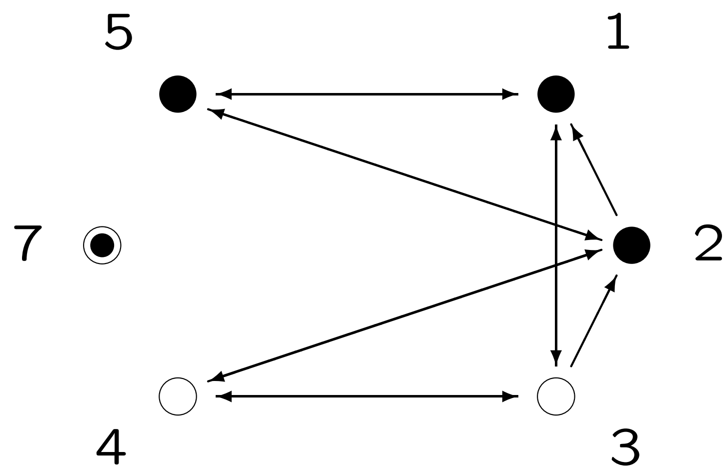
$t = 0.43$



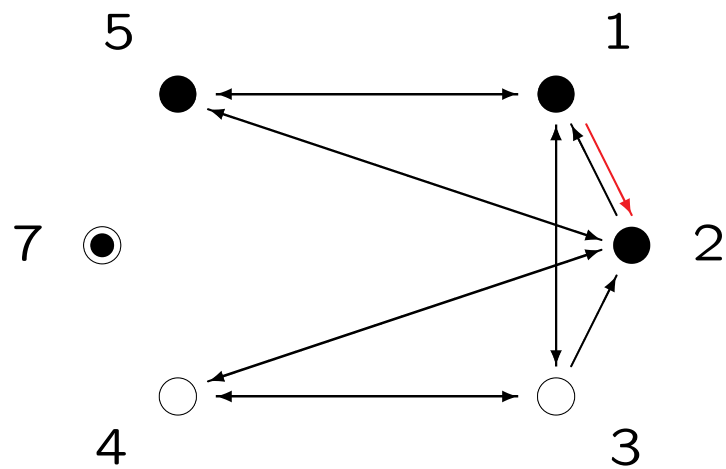
$t = 0.48$



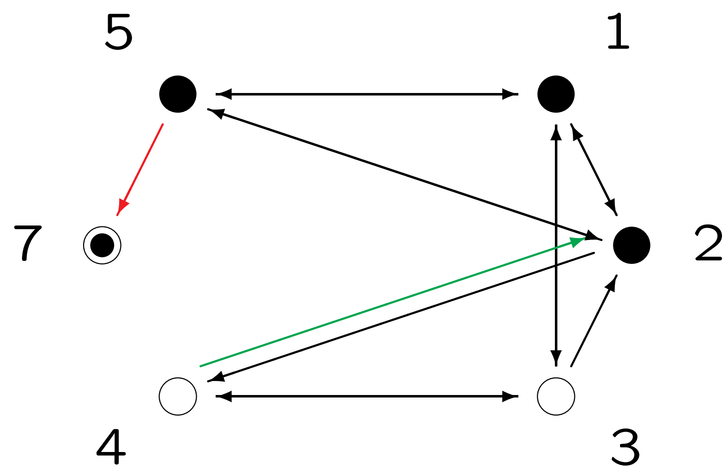
$t = 0.50$



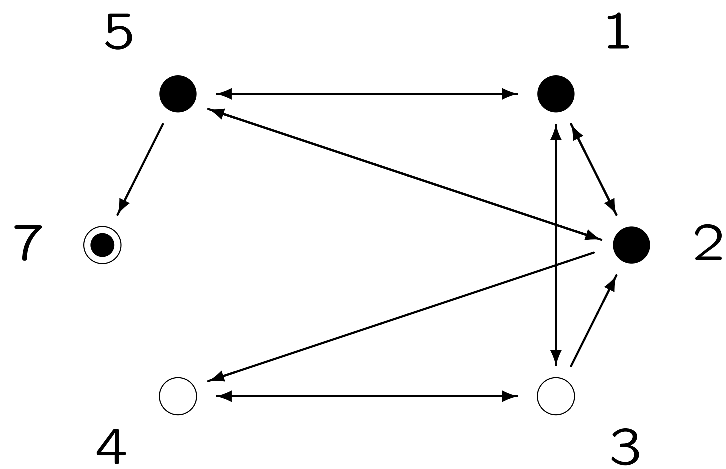
$t = 0.56$



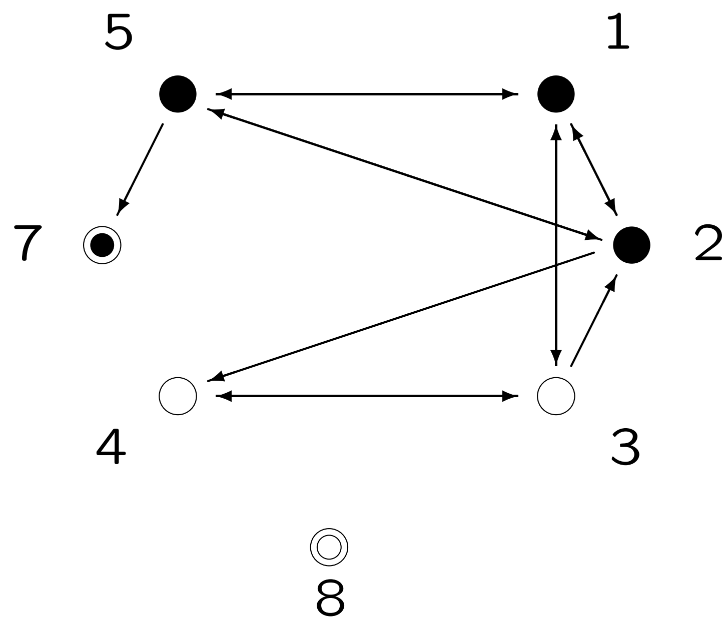
$t = 0.67$



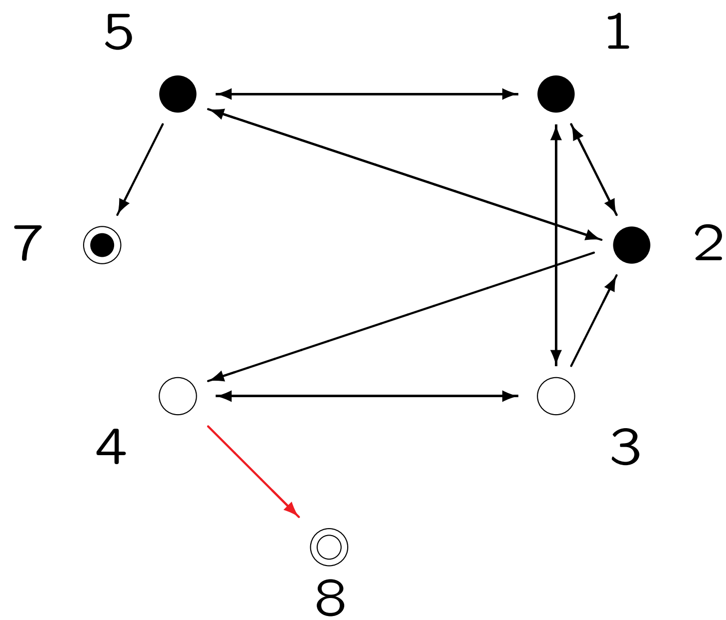
$t = 0.73$



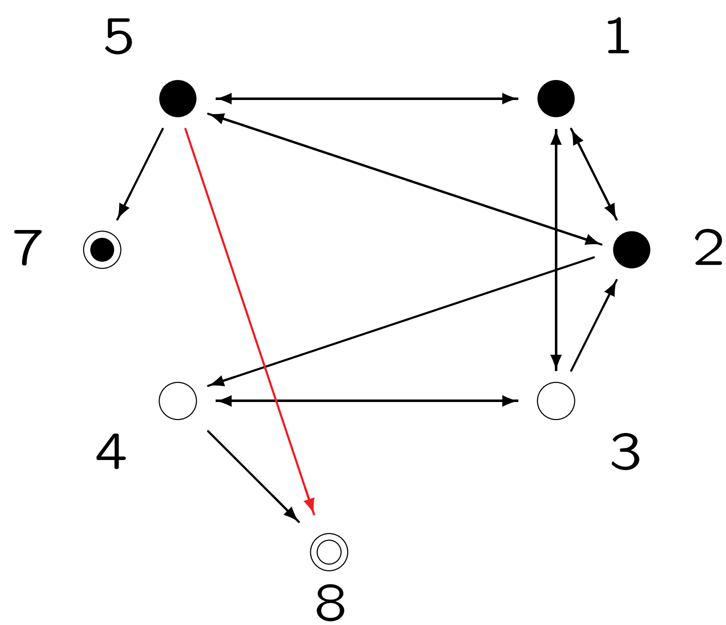
$t = 0.75$



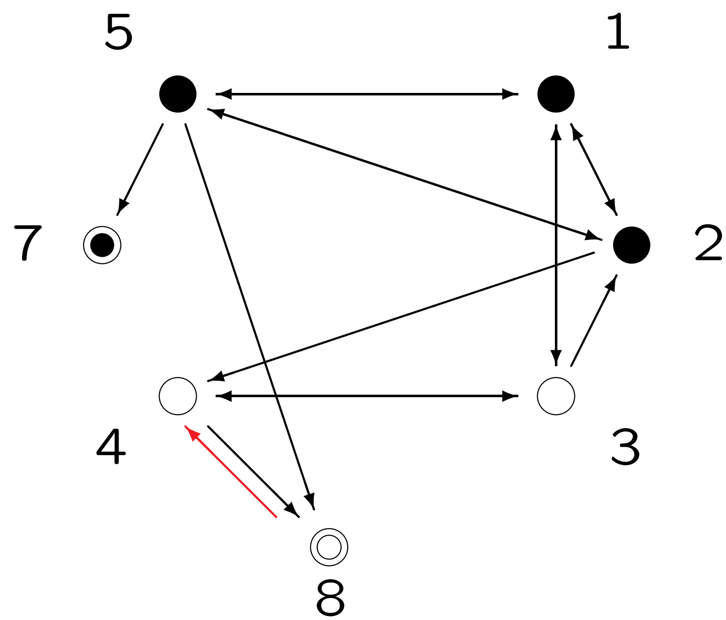
$t = 0.78$



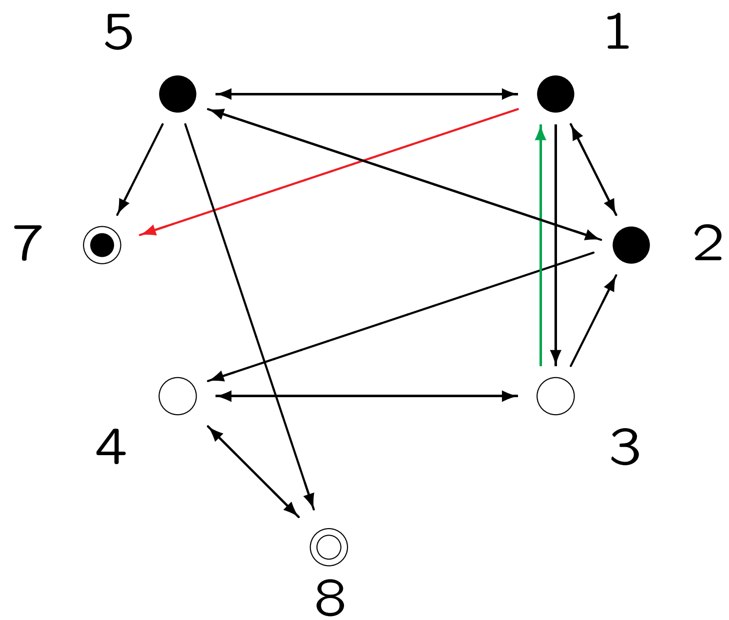
$t = 0.80$



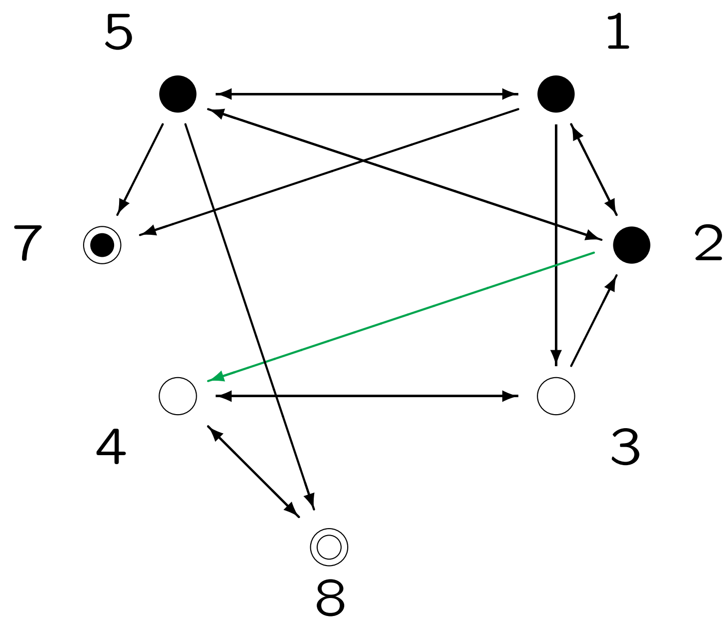
$t = 0.89$



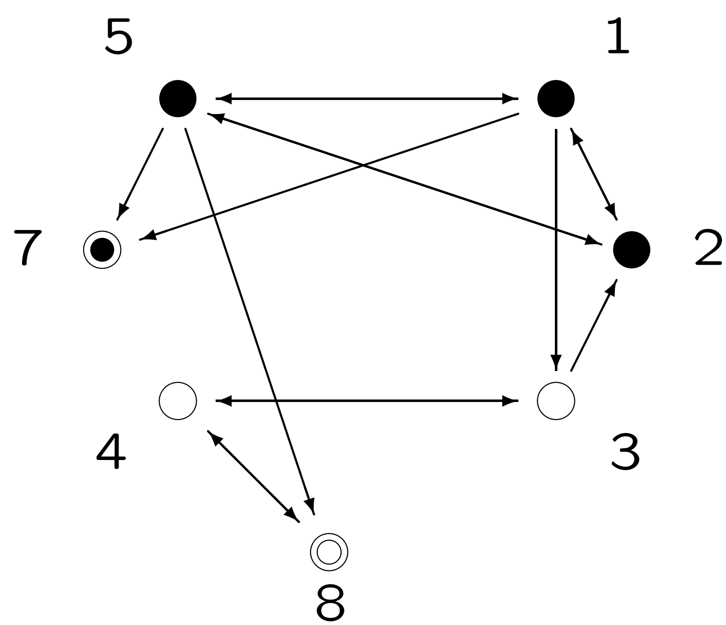
$t = 0.90$



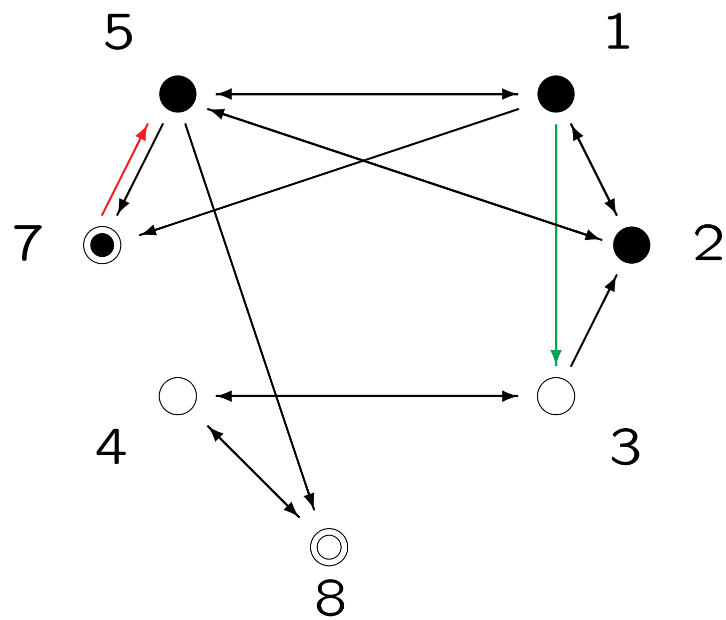
$t = 0.91$



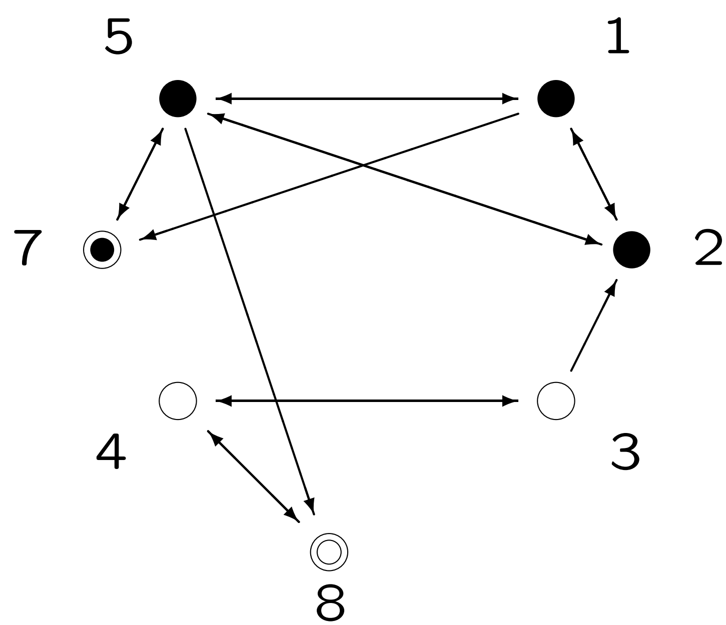
$t = 0.91$

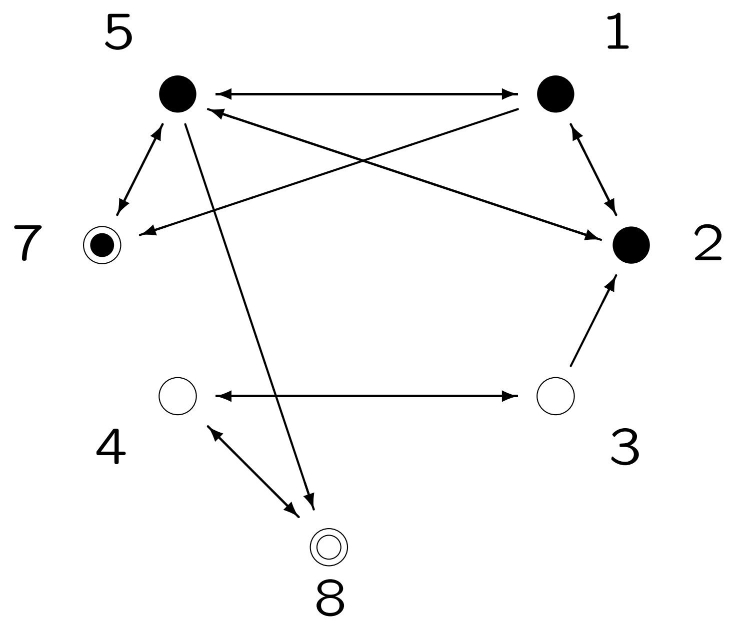


$t = 0.95$



$t = 0.98$





$t = 1$

STOP

Estimation

The model is estimated with the *Method of moments* using the *Robbins-Monro stochastic approximation algorithm*, based on the simulation of Markov chains

At each iteration step the simulated network (statistics) is compared with the observed network (statistics) at T_2, and the parameter estimates are updated

For details see Snijders (1996, 2001, 2005)

Example data: *Friendship* among university freshman (van de Bunt)

What are essential elements of how people determine and change their social ties?

1. Density, or number of ties ’outdegree’
2. Reciprocity
3. Transitivity:
   A friend of friend is also my friend
4. Indirect relations: number of actors to whom an actor is indirectly related (at distance 2)
5. characteristics of inidividuals: program, gender, smoking

StOCNET - module SIENA
(Simulation Investigation for Empirical Network Analysis).

Data: 4 observations at $t_2 - t_5$
of the dichotomized network (tie $=$ at least friendly relation)

Model 1 – start with a *baseline model*:

Objective function is a weighted sum of
number of ties (*outdegree*) and
number of reciprocated ties

The weights are the estimated parameters
Rates of change are constant between observations

StOCNET - SIENA output

```
Estimates and standard errors
0.1  Rate par. cond. var. period  1  1.2174 (0.1385)  rho = 3.3784
0.2  Rate par. cond. var. period  2  1.1498 (0.1332)  rho = 3.1576
0.3  Rate par. cond. var. period  3  1.4643 (0.1526)  rho = 4.3246

Other parameters:
  1. v: outdegree (density)          -0.9627  (0.1083)
  2. v: reciprocity                   1.6769  (0.1765)

Covariance matrix of estimates (correlations below diagonal):
     0.012     -0.008
    -0.417      0.031
```

Interpretation

constant rate parameters:
actors take on average 3 micro-steps in periods 1 and 2,
actors take on average 4 micro-steps in period 3
(micro-steps are occasions for change)

density parameter negative: -0.96
on average, costs of arbitrary ties is higher than their benefits

reciprocity effect positive but and strong: 1.67

Model 1 objective function:

$$f_i(x) = \sum_j \left(-0.96 x_{ij} + 1.67 x_{ij} x_{ji} \right)$$

Adding non-reciprocated tie gives -0.96
$\Rightarrow$ negative benefits

Adding reciprocated tie gives $-0.96 + 1.67 = 0.71$
$\Rightarrow$ positive benefits

Conclusion:
reciprocated ties are valued positively, and
actors will be reluctant to form unreciprocated ties

Model 2 interpretation

Add effects for *network closure*
actors prefer closed subgroups: positive *transitivity effect* (0.19)
and negative *indirect relations effect* (−0.33)

Model 2 objective function:

$$f_i(x) = -1.40 s_{i1}(x) + 1.57 s_{i2}(x) + 0.19 s_{i3}(x) - 0.33 s_{i5}$$

Example: Adding a reciprocated tie,
creating a transitive triplet (from the actor's viewpoint)
gives: $-1.40 + 1.57 + 0.19 + 0.33 = 0.69$

Effect	Model 2 Est.	Model 2 S.E.	Model 3 Est.	Model 3 S.E.	Model 4 Est.	Model 4 S.E.	Model 5 Est.	Model 5 S.E.
Rate effects								
Constant $t_2 - t_3$	4.52		4.83		5.14		3.38	
Constant $t_3 - t_4$	3.76		3.71		3.89		2.58	
Constant $t_4 - t_5$	4.48		4.72		4.86		2.79	
Outdegree							0.06	0.03
Structural effects								
Density	−1.40	0.14	−1.49	0.13	−1.55	0.15	−1.44	0.13
Reciprocity	1.57	0.22	1.36	0.25	0.98	0.35	1.18	0.26
Breaking recip. tie					−1.19	0.55	−0.75	0.37
Transitivity	0.19	0.04	0.20	0.04	0.21	0.04	0.18	0.03
Ind. relations	−0.33	0.09	−0.30	0.08	−0.32	0.09	−0.34	0.09
Covariate effects								
Sex ego			0.08	0.21				
Sex alter			0.16	0.19				
Sex sim.			0.40	0.17	0.26	0.15	0.25	0.18
Program sim.			0.81	0.18	0.84	0.24	0.86	0.18
Smoking sim.			0.36	0.06	0.37	0.17	0.37	0.14

Model 3 interpretation

Add *effects of Sex* (female = 1, male = 2)

1. *Similarity*: 0.40
   the actors prefer friendships with others from the same sex
2. *Popularity* (ego) and *Activity* (alter): not significant

Add *effects of Program*: 0.81
preference for actors within same program

Add *effects of Smoking*: 0.36
preference for actors with same smoking behavior

Extended model specification

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

Represents the "gratification" experienced by an 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 than for *termination* of ties

Maximize $f_i(\theta, x(i \leadsto j)) + g_i(\gamma, x, j) + U_i(t, x, j)$

$g_i(\gamma, x, j)$ is a *weighted sum* expressing the gratification for i when starting from the present network structure x as a consequence of changing his relation with j

Model 4: interpretation

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

Add effect for reciprocity: -1.19
$\Rightarrow$ Negative effect of breaking reciprocated tie

Total effect due to reciprocity conductive to *creating* a tie is 0.98
Total effect due to reciprocity conductive to *breaking* a tie is $-0.98 - 1.19 = -2.17$

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

Model 5: small positive effect (0.06) of outdegree
Actors with higher outdegrees tend to change their relations more often

Extended model specification

3. Networks as dependent and independent variables

Simultaneous endogeneous dynamics of networks and behavior,
or *co-evolution of networks and behavior*
e.g., friendship and smoking, or trust and stress

Repeated observations of:
the network (who is tied to whom) and *behavior* of all actors

Aim: disentangle effects *network* $\Rightarrow$ *behavior*
from effects *behavior* $\Rightarrow$ *network*

Methodology: *micro-steps* for change in network
and change in behavior.

Further issues

1. Model selection: score tests
2. Relations with more than 2 ordered values
3. Multivariate relations
4. Fit diagnostics and goodness-of-fit
5. Explained variation (‘R^2’)
6. Maximum likelihood estimation
7. Handling missing data