Stochastic Actor Oriented Models for Longitudinal Networks using RSiena

Ramiro Berardo

University of Wisconsin-Milwaukee

May 28th, 2014

First thing first

HUGE THANKS go to Tom Snijders
Most of the slides here are his
copyrighted material (used with permission).

All mistakes are mine!! (because I've made changes)

What we will cover today

What are Stochastic Actor-Oriented Models?

What are they good for?

What we will cover today

What are Stochastic Actor-Oriented Models?

What are they good for?

The basics of RSiena

What we will cover today

What are Stochastic Actor-Oriented Models?
What are they good for?

The basics of RSiena

Basic modeling for the evolution of networks

What we will cover today

What are Stochastic Actor-Oriented Models?
What are they good for?

The basics of RSiena

Basic modeling for the evolution of networks
Basic modeling for the coevolution of networks and behavior

What we will NOT cover today

Endowment and Creation functions

What we will NOT cover today

Endowment and Creation functions

Modeling for non-directed networks and multiplex networks

The procedures we will discuss are implemented in the R package

`R`

`S` imulation

`I` nvestigation for

`E` mpirical

`N` etwork

`A` nalysis

(frequently updated) with the website

`http://www.stats.ox.ac.uk/siena/`.

(programmed by Tom Snijders, Ruth Ripley, Krists Boitmanis; contributions by many others).

Examples of research questions in polisci that are amenable to social network analysis:

- Do voters extract political information from friend? What about “friends of friends"?

Examples of research questions in polisci that are amenable to social network analysis:

- Do voters extract political information from friend? What about “friends of friends"?
- How are social movements shaped? how do they affect individual behavior?

Examples of research questions in polisci that are amenable to social network analysis:

- Do voters extract political information from friend? What about “friends of friends"?
- How are social movements shaped? how do they affect individual behavior?
- How do advocacy coalitions form and evolve?
- etc.

In some of such questions, networks are *independent variables*.

In some others, they are *dependent variables*.

In some of such questions, networks are *independent variables*.

In some others, they are *dependent variables*.

Single observations of networks are snapshots,
therefore explaining them has limited importance.

In some of such questions, networks are *independent variables*.

In some others, they are *dependent variables*.

Single observations of networks are snapshots,
therefore explaining them has limited importance.

Longitudinal modeling offers more promise for understanding
how networks (and behavior) evolve.

1. Networks as dependent variables

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

1. Networks as dependent variables

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

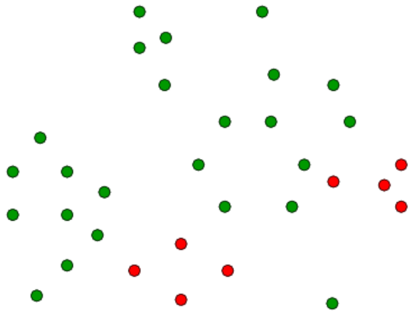
Data requirements:

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.

Example: Friendship among Freshmen

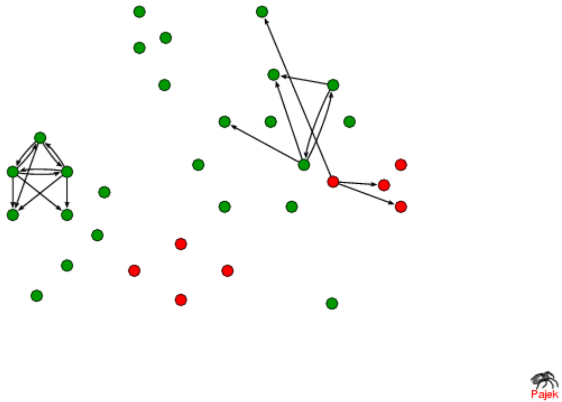
- Study of 32 freshmen in college,
  7 waves in 1 year.
  See van de Bunt, van Duijn, & Snijders,
  Computational & Mathematical Organization Theory,
  5 (1999), 167 – 192.

This data set can be pictured by the following graphs
(arrow stands for 'best friends').



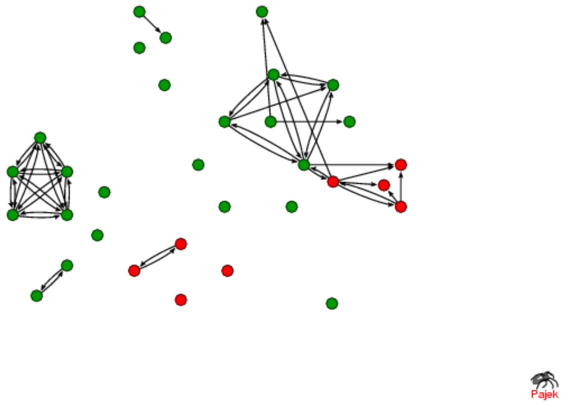
Friendship network time 1.

Average degree 0.0; missing fraction 0.0.



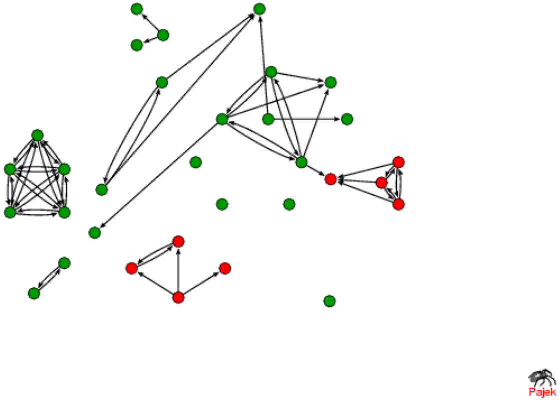
Friendship network time 2.

Average degree 0.7; missing fraction 0.06.



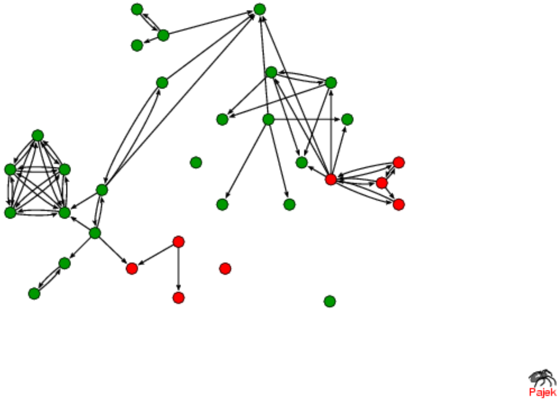
Friendship network time 3.

Average degree 1.7; missing fraction 0.09.



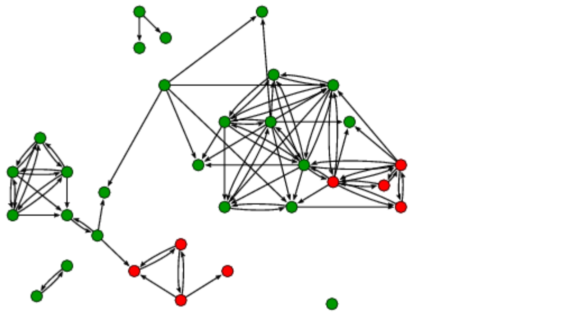
Friendship network time 4.

Average degree 2.1; missing fraction 0.16.



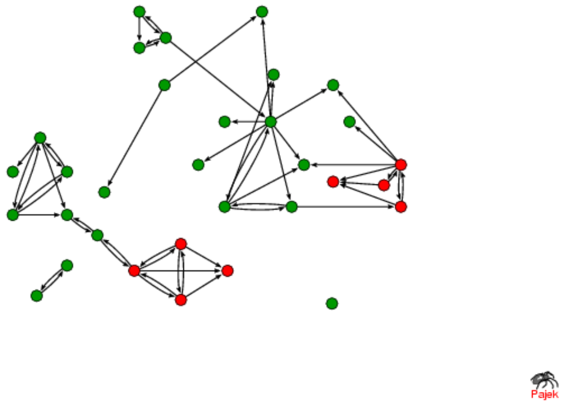
Friendship network time 5.

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.

Dynamics of social networks are complicated

Dynamics of social networks are complicated

“Network effects” are endogenous feedback effects:
e.g., reciprocity, transitivity, popularity, subgroup formation.

Dynamics of social networks are complicated

"Network effects" are endogenous feedback effects:
e.g., reciprocity, transitivity, popularity, subgroup formation.

We need models for network dynamics
that are flexible enough to represent
the complicated dependencies in such processes

Stochastic Actor Oriented Models ('SAOMs')

They allow us to model network evolution *(dependent var.)* as function of

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

simultaneously.

Stochastic Actor Oriented Models (‘SAOMs’)

They allow us to model network evolution *(dependent var.)* as function of

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

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

Principles for this approach to analysis of network dynamics:

1. use simulation models as *models for data*

Principles for this approach to analysis of network dynamics:

1. use simulation models as *models for data*
2. include a random influence in the simulation model to account for 'unexplained variability'

Principles for this approach to analysis of network dynamics:

1. use simulation models as *models for data*
2. include a random influence in the simulation model
   to account for ‘unexplained variability’
3. use methods of statistical inference
   for probability models implemented as simulation models

Principles for this approach to analysis of network dynamics:

1. use simulation models as *models for data*
2. include a random influence in the simulation model
   to account for ‘unexplained variability’
3. use methods of statistical inference
   for probability models implemented as simulation models
4. for panel data: employ a continuous-time model
   to represent unobserved endogenous network evolution

Principles for this approach to analysis of network dynamics:

1. use simulation models as *models for data*
2. include a random influence in the simulation model
   to account for ‘unexplained variability’
3. use methods of statistical inference
   for probability models implemented as simulation models
4. for panel data: employ a continuous-time model
   to represent unobserved endogenous network evolution
5. condition on the first observation and do not model it:
   no stationarity assumption.

Stochastic Actor-Oriented Model ('SAOM')

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.
   Matrix X is *adjacency matrix* of digraph.
   X_{ij} is a *tie indicator* or *tie variable*.

Stochastic Actor-Oriented Model ('SAOM')

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.
   Matrix X is *adjacency matrix* of digraph.
   X_{ij} is a *tie indicator* or *tie variable*.
2. Exogenously determined independent variables:
   actor-dependent covariates v, dyadic covariates w.
   These can be constant or changing over time.

Stochastic Actor-Oriented Model ('SAOM')

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.
   Matrix X is *adjacency matrix* of digraph.
   X_{ij} is a *tie indicator* or *tie variable*.
2. Exogenously determined independent variables:
   actor-dependent covariates v, dyadic covariates w.
   These can be constant or changing over time.
3. Continuous time parameter t,
   observation moments $t_1, \ldots, t_M$.

Stochastic Actor-Oriented Model ('SAOM')

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.
   Matrix X is *adjacency matrix* of digraph.
   X_{ij} is a *tie indicator* or *tie variable*.
2. Exogenously determined independent variables:
   actor-dependent covariates v, dyadic covariates w.
   These can be constant or changing over time.
3. Continuous time parameter t,
   observation moments $t_1, \ldots, t_M$.
4. Current state of network $X(t)$ is dynamic constraint for its own change process: Markov process.

'actor-oriented' = 'actor-based'

5. The actors control their outgoing ties.

‘actor-oriented’ = ‘actor-based’

5. The actors control their outgoing ties.
6. The ties have inertia: they are *states* rather than *events*.
   At any single moment in time,
   only one variable $X_{ij}(t)$ may change.

'actor-oriented' = 'actor-based'

5. The actors control their outgoing ties.
6. The ties have inertia: they are *states* rather than *events*.
   At any single moment in time,
   only one variable $X_{ij}(t)$ may change.
7. Changes are modeled as choices by actors
   in their outgoing ties, with probabilities
   depending on the '*objective function*'

The change process is decomposed into two sub-models, formulated on the basis of the idea that the actors *i* control their outgoing ties $(X_{i1}, \ldots, X_{in})$:

The change process is decomposed into two sub-models, formulated on the basis of the idea that the actors i control their outgoing ties $(X_{i1}, \ldots, X_{in})$:

1. waiting times until the next opportunity
   for a change made by actor i:
   rate functions;

The change process is decomposed into two sub-models, formulated on the basis of the idea that the actors i control their outgoing ties $(X_{i1}, \ldots, X_{in})$:

1. waiting times until the next opportunity
   for a change made by actor i:
   rate functions;

2. probabilities of changing (toggling) X_{ij},
   conditional on such an opportunity for change:
   objective functions.

The change process is decomposed into two sub-models, formulated on the basis of the idea that the actors *i* control their outgoing ties $(X_{i1}, \ldots, X_{in})$:

1. waiting times until the next opportunity for a change made by actor *i*: *rate functions*;

2. probabilities of changing (toggling) X_{ij}, conditional on such an opportunity for change: *objective functions*.

The distinction between rate function and objective function separates the model for *how many* changes are made from the model for *which* changes are made.

This decomposition between the timing model and the model for change can be pictured as follows:

At randomly determined moments t,
actors i have opportunity to change a tie variable X_{ij}:
micro step.
(Actors are also permitted to leave things unchanged.)
Frequency of micro steps is determined by *rate functions*.

When a micro step is taken,
the probability distribution of the result of this step
depends on the *objective function* :
higher probabilities of moving toward new states
that have higher values of the objective function.

Specification: rate function

'how fast is change / opportunity for change ?'

Rate of change of the network by actor *i* is denoted λ_i :
expected frequency of opportunities for change by actor *i*.

Simple specification: rate functions are constant within periods.

Specification: rate function

'how fast is change / opportunity for change ?'

Rate of change of the network by actor *i* is denoted λ_i :
expected frequency of opportunities for change by actor *i*.

Simple specification: rate functions are constant within periods.

Specification: objective function

‘what is the direction of change?’

Specification: objective function

'what is the direction of change?'

The objective function $f_i(\beta, x)$ indicates
preferred 'directions' of change.
β is a statistical parameter, i is the actor (node), x the network.

When actor i gets an opportunity for change,
he has the possibility to change *one* outgoing tie variable X_{ij},
or leave everything unchanged.

By $x^{(\pm ij)}$ is denoted the network obtained
when x_{ij} is changed ('toggled') into $1 - x_{ij}$.
Formally, $x^{(\pm ii)}$ is defined to be equal to x.

Conditional on actor i being allowed to make a change, the probability that X_{ij} changes into $1 - X_{ij}$ is

$$p_{ij}(\beta, x) = \frac{\exp\left(f_i(\beta, x^{(\pm ij)})\right)}{\sum_{h=1}^{n} \exp\left(f_i(\beta, x^{(\pm ih)})\right)} ,$$

and p_{ii} is the probability of not changing anything.

Higher values of the objective function indicate the preferred direction of changes.

One way of obtaining this model specification is to suppose that actors make changes such as to optimize
the objective function $f_i(\beta, x)$
plus a random disturbance that has a Gumbel distribution
(Multinomial Logit form of random utility models in econometrics!)

Actor i chooses the "best" j by maximizing

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

(with the formal definition $x^{(\pm ii)} = x$).

Objective functions will be defined as sum of:

1. *evaluation function* expressing satisfaction with network;

Objective functions will be defined as sum of:

1. *evaluation function* expressing satisfaction with network;

To allow asymmetry creation $\leftrightarrow$ termination of ties:

2. *creation function*
   expressing aspects of network structure
   playing a role only for creating new ties

Objective functions will be defined as sum of:

1. *evaluation function* expressing satisfaction with network;

To allow asymmetry creation $\leftrightarrow$ termination of ties:

2. *creation function*
   expressing aspects of network structure
   playing a role only for creating new ties
3. *endowment = maintenance function*
   expressing aspects of network structure
   playing a role only for maintaining existing ties

If creation function = endowment function,
then these can be jointly replaced by the evaluation function.
This is usual for starting modelling.

Model specification:

Simple specification: only evaluation function;
periodwise constant rate function.

Model specification:

Simple specification: only evaluation function; periodwise constant rate function.

Evaluation function f_i reflects network effects (endogenous) and covariate effects (exogenous).

Model specification:

Simple specification: only evaluation function;
periodwise constant rate function.

Evaluation function f_i reflects network effects (endogenous) and covariate effects (exogenous).
Convenient definition of evaluation function is a weighted sum

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

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

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

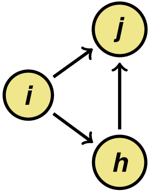
1. *out-degree effect*, controlling the density / average degree,
   $s_{i1}(x) = x_{i+} = \sum_j x_{ij}$

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

1. *out-degree effect*, controlling the density / average degree,
   $s_{i1}(x) = x_{i+} = \sum_j x_{ij}$
2. *reciprocity effect*, number of reciprocated ties
   $s_{i2}(x) = \sum_j x_{ij}\, x_{ji}$

Various potential effects representing network closure:

3. *transitive triplets effect* ('transTrip'), number of transitive patterns in i's ties $(i \to j,\ i \to h,\ h \to j)$
$s_{i3}(x) = \sum_{j,h} x_{ij}\, x_{ih}\, x_{hj}$

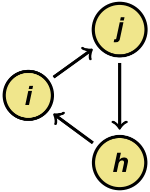


transitive triplet

4. *indirect ties effect*,
   this is the number of actors j to whom i is tied indirectly
   (through at least one intermediary: $x_{ih} = x_{hj} = 1$)
   but not directly $x_{ij} = 0$),
   = number of geodesic distances equal to 2,
   $s_{i4}(x) = \#\{j \mid x_{ij} = 0, \max_h(x_{ih}\, x_{hj}) > 0\}$

   Network closure will lead to relatively few
   geodesic distances equal to 2.

5. *three-cycle effect*,
number of three-cycles in i's ties
$(i \to j,\ j \to h,\ h \to i)$
$s_{i5}(x) = \sum_{j,h}\ x_{ij}\, x_{jh}\, x_{hi}$



three-cycle

This represents a kind of generalized reciprocity,
and absence of hierarchy.

6. *in-degree related popularity effect*, sum friends' in-degrees
$s_{i6}(x) = \sum_j x_{ij}\,\sqrt{x_{+j}} = \sum_j x_{ij}\sqrt{\sum_h x_{hj}}$
related to dispersion of in-degrees

6. *in-degree related popularity effect*, sum friends' in-degrees
   $s_{i6}(x) = \sum_j x_{ij} \sqrt{x_{+j}} = \sum_j x_{ij} \sqrt{\sum_h x_{hj}}$
   related to dispersion of in-degrees
7. *out-degree related popularity effect*,
   sum friends' out-degrees
   $s_{i7}(x) = \sum_j x_{ij} \sqrt{x_{j+}} = \sum_j x_{ij} \sqrt{\sum_h x_{jh}}$
   related to association in-degrees — out-degrees;

Overview of structural effects

1. Always: outdegree effect (like constant term in regression)

Overview of structural effects

1. Always: outdegree effect (like constant term in regression)
2. Almost always: reciprocity

Overview of structural effects

1. Always: outdegree effect (like constant term in regression)
2. Almost always: reciprocity
3. Triadic effects: transitivity, 3-cycles, reciprocity $\times$ transitivity, etc.

Overview of structural effects

1. Always: outdegree effect (like constant term in regression)
2. Almost always: reciprocity
3. Triadic effects: transitivity, 3-cycles,
   reciprocity $\times$ transitivity, etc.
4. Degree-related effects:
   inPop, outPop, etc.

Overview of structural effects

1. Always: outdegree effect (like constant term in regression)
2. Almost always: reciprocity
3. Triadic effects: transitivity, 3-cycles, reciprocity $\times$ transitivity, etc.
4. Degree-related effects: inPop, outPop, etc.

Of course, there are more.

Four kinds of evaluation function effect associated with actor covariate v_i.

6. *covariate-related popularity*, 'alter'
sum of covariate over all of i 's friends
$s_{i6}(x) = \sum_j x_{ij} v_j$;

Four kinds of evaluation function effect associated with actor covariate v_i.

6. *covariate-related popularity*, 'alter'
   sum of covariate over all of i 's friends
   $s_{i6}(x) = \sum_j x_{ij} \, v_j$;
7. *covariate-related activity*, 'ego'
   i's out-degree weighted by covariate
   $s_{i7}(x) = v_i \, x_{i+}$;

8. *covariate-related similarity*,
sum of measure of covariate similarity
between i and his friends,
$s_{i8}(x) = \sum_j x_{ij}\,\mathrm{sim}(v_i, v_j)$
where $\mathrm{sim}(v_i, v_j)$ is the similarity between v_i and v_j,
$$\mathrm{sim}(v_i, v_j) \;=\; 1 - \frac{|v_i - v_j|}{R_V} \;,$$
R_V being the range of V;

8. *covariate-related similarity*,
   sum of measure of covariate similarity
   between i and his friends,
   $s_{i8}(x) = \sum_j x_{ij} \, \mathrm{sim}(v_i, v_j)$
   where $\mathrm{sim}(v_i, v_j)$ is the similarity between v_i and v_j,
   $$\mathrm{sim}(v_i, v_j) \; = \; 1 - \frac{|v_i - v_j|}{R_V} \; ,$$
   R_V being the range of V;
9. *covariate-related interaction*, 'ego $\times$ alter'
   $s_{i9}(x) = v_i \sum_j x_{ij} \, v_j$;

The estimation of parameters is accomplished with
a three-phase iterative stochastic simulation algorithm

The estimation of parameters is accomplished with a three-phase iterative stochastic simulation algorithm

Phase I. The algorithm produces rough parameters that reproduce observed network configurations (or “target statistics")

The estimation of parameters is accomplished with a three-phase iterative stochastic simulation algorithm

Phase I. The algorithm produces rough parameters that reproduce observed network configurations (or “target statistics")

Phase II. The second phase progressively improves the value of the parameters through multiple simulations. The goal is to make the parameters more likely to produce statistics that equal those in the observed network.

The estimation of parameters is accomplished with a three-phase iterative stochastic simulation algorithm

Phase I. The algorithm produces rough parameters that reproduce observed network configurations (or “target statistics")

Phase II. The second phase progressively improves the value of the parameters through multiple simulations. The goal is to make the parameters more likely to produce statistics that equal those in the observed network.

Phase III. Checks whether the average statistics obtained in the simulations of phase 2 are close to the target statistics. If so, the model “converges".

Time to practice!

Download and open the
the R script on my website

Networks as dependent and independent variables

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

- voting behavior
- legislation co-sponsorship
- management of common-pool resources
- political activism
- war and trade

Networks as dependent and independent variables

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

- voting behavior
- legislation co-sponsorship
- management of common-pool resources
- political activism
- war and trade

In all of these, social interactions and behavior affect each other

Two-way influence between networks and behavior

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

Actors are influenced in their behavior, 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 behavior and other characteristics
(similarity, opportunities for future rewards, etc.).

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

Terminology

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

Thus, there is a feedback relation in the dynamics of relational networks and actor behavior / 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* $\Rightarrow$ *behavior*
  and the *behavior* $\Rightarrow$ *network* effects
  leads to ‘network autocorrelation’
  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.

Data

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

several discrete observation moments;

for each observation moment:

- network: who is tied to whom
- behavior of all actors

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

Notation:

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

In addition to the network X, associated to each actor i
there is a vector $Z_i(t)$ of actor characteristics
indexed by $h = 1, \ldots, H$.
Assumption: 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^{Z_h} for changes in behavior h),
the actors may change a tie or a behavior.

Probabilities of change are increasing functions of
objective functions of the new state,
defined specifically for network, f^X,
and for behavior, f^Z .

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

For network change, change probabilities are as before.

For the behaviors, the formula of the change probabilities is

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

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

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

Again, multinomial logit form.

Again, a 'maximizing' interpretation is possible.

Micro-step for change in network:

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:

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 behavior:

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

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

Optimizing interpretation:

When actor i 'may' change an outgoing tie variable to some other 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 behavior 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 behavior change
plus a random component representing unexplained influences.

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

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

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

Optimal behavior change:

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:

$$f_i^{Z_h}(\beta, x, z(i, h \rightsquigarrow v)) \quad + \quad \underbrace{U_i^{Z_h}(t, z, h, v)}_{\Uparrow \atop \text{random component}} .$$

(behavior is permitted to stay the same.)

Specification of the behavior model

Many different reasons why networks are important for behavior:

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 behavior;

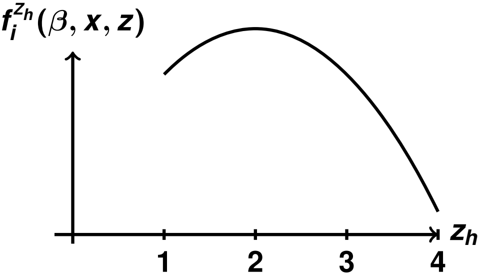
Theoretical elaboration will be helpful for a good data analysis.

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

For a negative quadratic tendency parameter,
the model for behavior is a unimodal preference model.



For positive quadratic tendency parameters ,
the behavior objective function can be bimodal
('positive feedback').

3. *behavior-related average similarity*,
   average of behavior 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;

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

10. *popularity-related tendency*, (in-degree)
$s_{i10}(x, z) = z_{ih}\, x_{+i}$

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

12. *popularity-related tendency*, (in-degree)
    $s_{i12}(x, z) = z_{ih} \, x_{+i}$
13. *activity-related tendency*, (out-degree)
    $s_{i13}(x, z) = z_{ih} \, x_{i+}$

12. *dependence on other behaviors* $(h \neq \ell)$,
    $s_{i12}(x, z) = z_{ih}\, z_{i\ell}$
13. *influence from other characteristics V*
    $s_{i13}(x, z) = z_{ih}\, \frac{1}{x_{i+}} \sum_j x_{ij}\, v_j$

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

Statistical estimation: networks & behavior

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 $\Rightarrow$ behavior and behavior $\Rightarrow$ networks,
both lead to 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:

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

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

Let’s go back to the RSiena script!

Some references about longitudinal models

- Tom A.B. Snijders,
  The Statistical Evaluation of Social Network Dynamics,
  Sociological Methodology 2001, 361–395;
- Tom A.B. Snijders, 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).
- 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.
- See SIENA manual and homepage.

Lots of new developments

- goodness of fit (Josh Lospinoso)
- diffusion of innovations (Charlotte Greenan)
- multivariate dependent networks
  e.g. signed networks: positive and negative ties
- networks with discrete ordered tie values
- two-mode dependent networks
- co-evolution of one-mode and two-mode networks
- multilevel networks (Johan Koskinen & Tom Snijders)
- etc.

A lot of material
– programs, manuals, papers, workshop announcements –
can be found at the Siena website:

`http://www.stats.ox.ac.uk/siena/`

There is also a user's group:

`http://groups.yahoo.com/groups/stocnet/`