Network Modeling

Viviana Amati Jürgen Lerner David Schoch

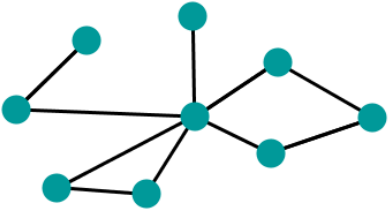
Dept. Computer & Information Science
University of Konstanz

Winter 2013/2014
(version 05 February 2014)

Outline

Outline

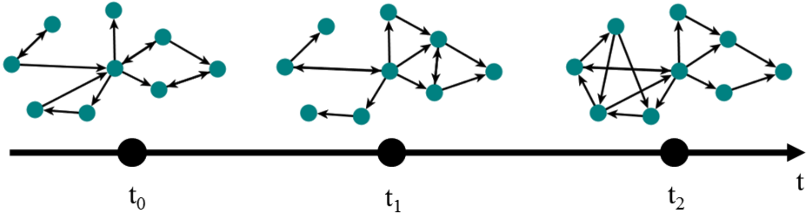
Where we are



Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties dependency
Planted partition	intra/inter group density	ties dependency
Preferential attachment	degree distribution	other structural properties
ERGM	class of models	reasonable representation

These are models for cross-sectional data

Where we are going



Network are dynamic by nature. How to model network evolution?

We need a model for longitudinal data

Networks are dynamic by nature: a real example

The *Teenage Friends and Lifestyle Study* analyzes smoking behavior and friendship

Data collection: (available from http://www.stats.ox.ac.uk/∼snijders/siena/)

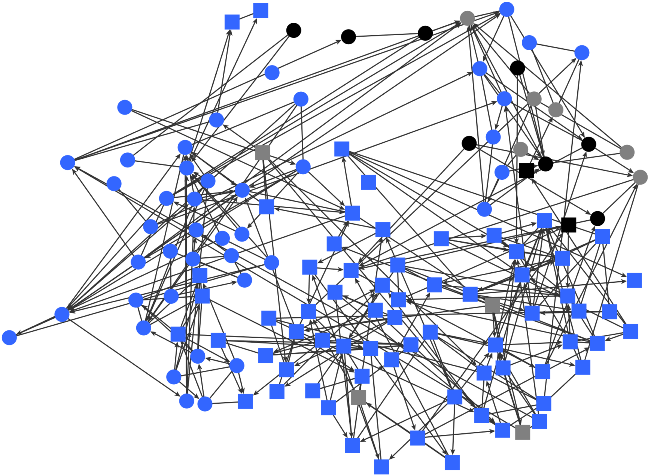
- One school year group monitored over 3 years;
- questionnaires at approximately one year interval:
  1. Friendship relation: each pupil could name up to 12 friends
  2. Individual information and lifestyle elements: gender, age, substances use, smoking of parents and siblings etc.

arrows = friendship relation

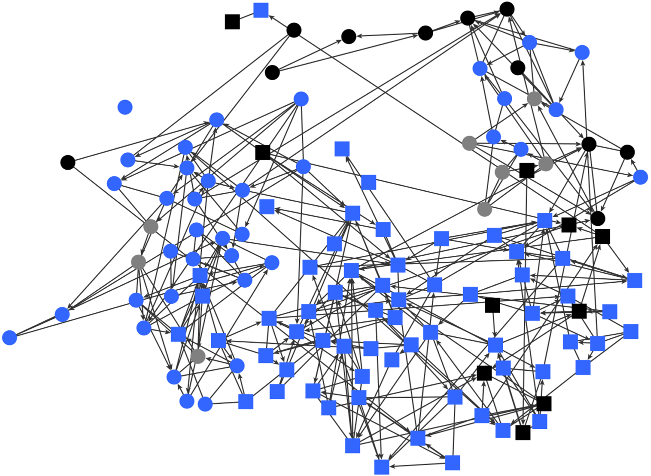
gender: circle = girl, square = boy

smoking behavior: blue = non, gray = occasional, black = regular

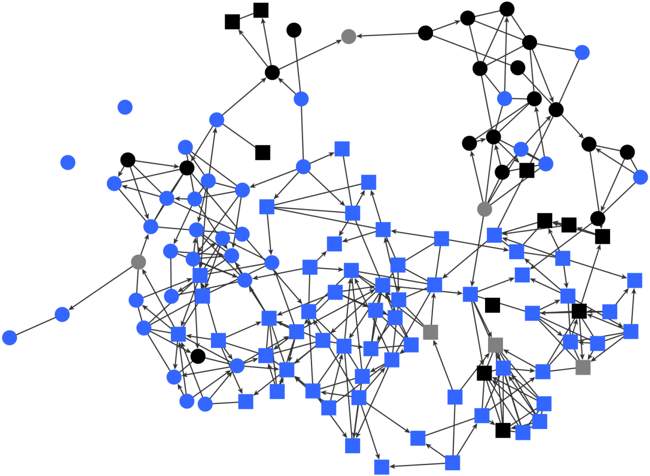
Networks are dynamic by nature: a real example



Networks are dynamic by nature: a real example



Networks are dynamic by nature: a real example

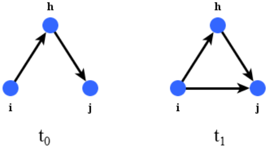


Some questions

Is there any tendency in friendship formation ...

- towards reciprocity?

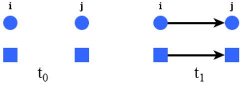
- towards transitivity?



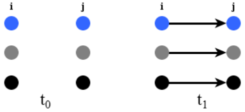
Some questions

Is there any homophily in friendship formation with respect to ...

- gender?



- smoking behavior?



Solution

Stochastic actor-oriented model (SAOM)

Aim

Explain network evolution as a result of

- endogenous variables: structural effects depending on the network only (e.g. reciprocity, transitivity, etc.)
- exogenous variables: actor-dependent and dyadic-dependent covariates (e.g. effect of a covariate on the existence of a tie or on homophily)

simultaneously

Background: random variables

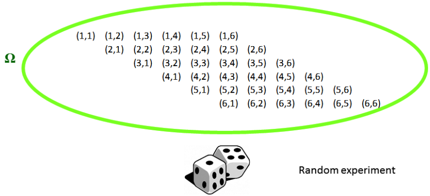
Definitions

Let (Ω, P) be a *probability space*

$\Omega =$ is a *set of possible outcomes*

$P : \Omega \to [0,1]$ is a *probability function* such that:

1. $P(\omega) \geq 0$
2. $\sum_{\omega \in \Omega} P(\omega) = 1$



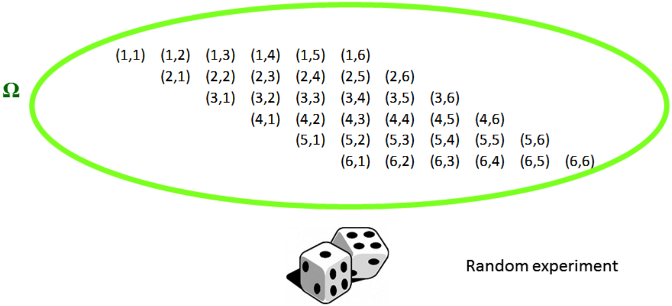
1. $P((a,b)) = \dfrac{1}{36}, \quad a, b \in \{1, 2, \cdots, 6\}$
2. $\displaystyle\sum_{\omega \in \Omega} P(\omega) = 1$

A (real-valued) *random variable* (r.v.) is a function $X : (\Omega, P) \to (\mathbb{R}, P)$.

Background: random variables (motivation)

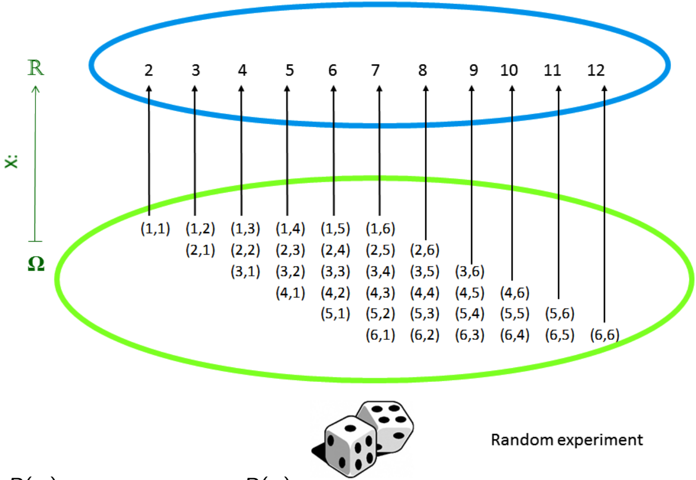
Example

X: sum of the two dice



Background: random variables (motivation)

Example

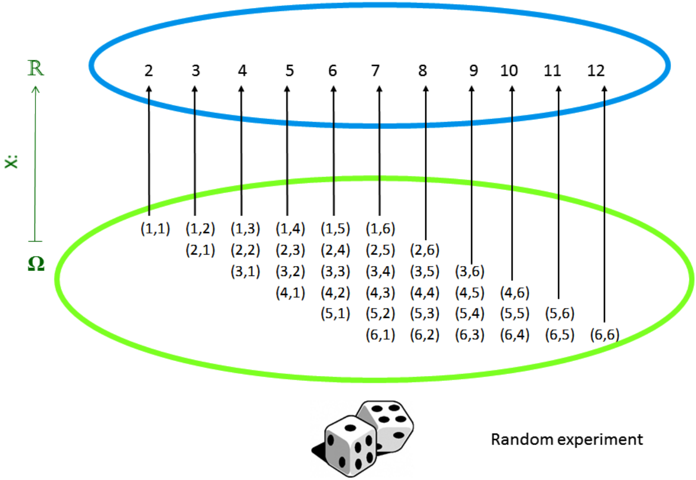


Given $P(\omega)$ we can compute $P(x)$:

$$P(X = 4) = P((1,3)) + P((2,2)) + P((3,1)) = 1/36 + 1/36 + 1/36 = 1/12$$

Background: random variables (motivation)

Example



N.b.:
- capital letters denote r.vs (e.g. X=sum of two dice)
- small letters denote the values assumed by a r.v. (e.g. x=2)
- The set of values that X can take is called range and will be denoted by $\mathcal{S}$ (e.g. $\mathcal{S} = \{2, 3, \ldots, 12\}$)

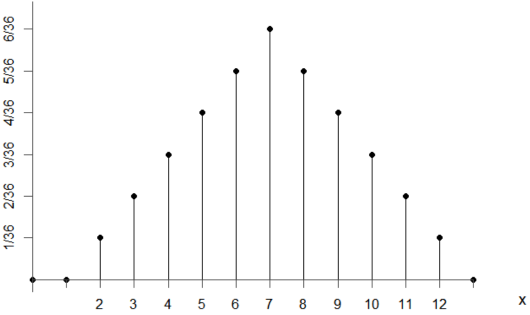
Background: discrete random variable

Definition

A r.v. X is defined to be *discrete* if $\mathcal{S}$ is countable.

The *probability mass function* (p.m.f) $\varphi_X(x) : \mathbb{R} \to [0,1]$ describes the values that X can take along with the probability associated with each value

x	$\varphi_X(x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36



Background: discrete random variable

Definition

A r.v. X is defined to be *discrete* if $\mathcal{S}$ is countable.

The *probability mass function* (p.m.f) $\varphi_X(x) : \mathbb{R} \to [0, 1]$ describes the values that a X can take along with the probability associated with each value

$$\varphi_X(x) = P(X = x)$$

The *cumulative distribution function* (c.d.f.) $F_X(x) : \mathbb{R} \to [0, 1]$ describes the probability that X takes value lower than x

$$F_X(x) = P(X \leq x) = \sum_{x' < x} P(X = x')$$

Examples

X=Sum of two dice

$$P(X \leq 3) = P(X = 2) + P(X = 3) = 1/36 + 2/36 = 1/12$$

Background: continuous random variable

Definition

A random variable X is called *(absolutely) continuous* if $\mathcal{S}$ is uncountable and there exists a function $f_X(x) : \mathbb{R} \to \mathbb{R}^+$ such that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)du \quad \forall x \in \mathbb{R}$$

$$P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f_X(x)dx = 1$$

$f_X(x)$ is the *probability density function* (p.d.f)

Examples

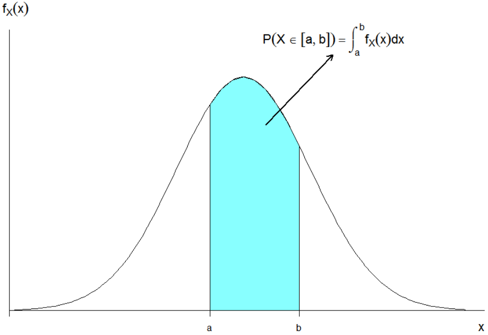
- $X =$ weight of people in a population
- $X =$ waiting time at a post office clerk
- ...

Background: continuous random variable

The p.d.f. $f_X(x)$ allows to compute all the probability statements about X. For instance, the probability that X takes values in $[a, b]$ is

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

Geometrical interpretation



Intuition suggests that

$$P(X = x) = \int_x^x f_X(u)du = 0$$

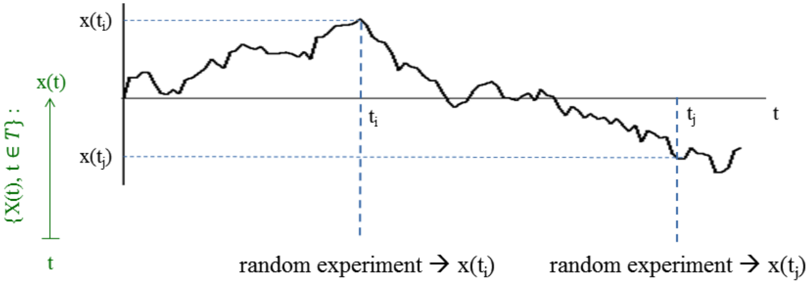
Thus, we cannot determine a continuous random variable via its "mass function"

Background: stochastic (or random) process

Definition

A stochastic process $\{X(t), t \in \mathcal{T}\}$ is a mapping

$$\forall t \in \mathcal{T} \mapsto X(t) : \Omega \to \mathbb{R}$$



Background: stochastic process

$\mathcal{T}$ = index set (usually interpreted as time)
$\mathcal{S}$ = state space

Different stochastic processes can be defined according to $\mathcal{S}$ and $\mathcal{T}$

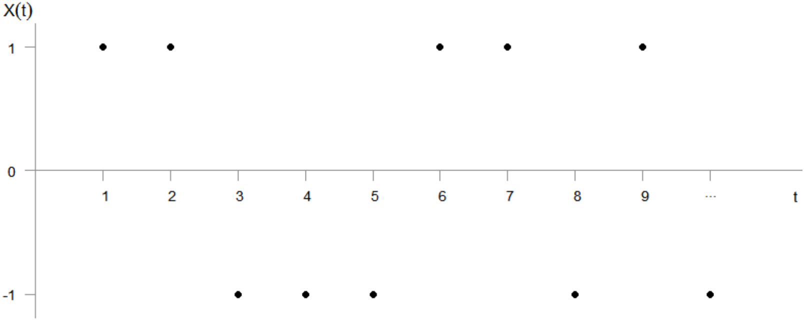
$\mathcal{S}$	$\mathcal{T}$	
	Countable (discrete)	**Uncountable (continuous)**
Countable (finite)	discrete-time with finite state space	continuous-time with finite state space
Uncountable (continuous)	discrete-time with continuous state space	continuous-time with continuous state space

Background: stochastic process

Example

$X(t)$ = the outcome of flipping a coin

$\mathcal{S} = \{-1, 1\}$, where -1 =tail 1 =head
$\mathcal{T} = \{1, 2, \cdots\}$



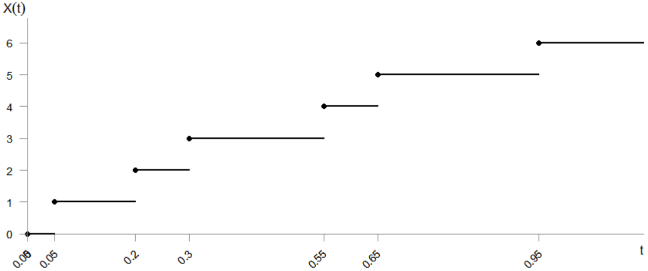
$\{X(t), t \in \mathcal{T}\}$ is a discrete-time stochastic process with a finite state space

Background: stochastic process

Example

$X(t) =$ the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

$\mathcal{S} = \{0, 1, 2, \cdots\}$
$\mathcal{T} = [0, 12]$



$\{X(t), t \in \mathcal{T}\}$ is a continuous-time stochastic process with a finite state space

Background: continuous-time Markov Chain

Definition

$\{X(t),\ t \in \mathcal{T}\}$ has the *Markov property* if:

$\forall\ x \in \mathcal{S}$ and $\forall\ t_i < t_j$

$$P(X(t_j) = x(t_j) \mid X(t) = x(t)\ \ \forall\ \ t \leq t_i) = P(X(t_j) = x(t_j) \mid X(t_i) = x(t_i))$$

Definition

A *continuous-time* Markov chain $\{X_t,\ t \geq 0\}$ is a stochastic process having

1. **finite state**
2. **continuous-time**
3. the **Markovian property**

Background: continuous-time Markov Chain

Example

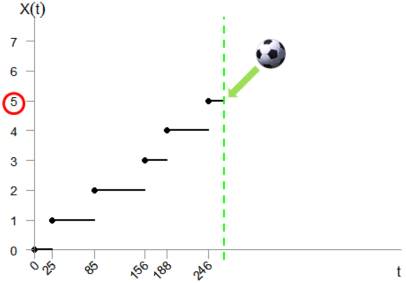
$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

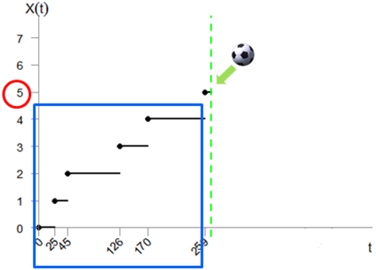
$$\{X(t),\ t \geq 0\} \text{ is a continuous-time Markov chains}$$

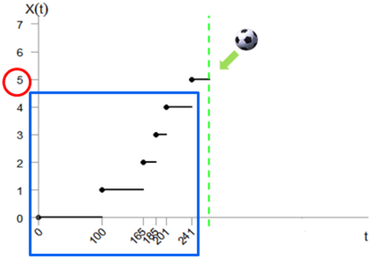
Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$
   B = total number of goals scored during the career
2. the time is continuous: [0,T]
   T = time of retirement
3. the process $\{X(t),\ t \geq 0\}$ has the Markov property

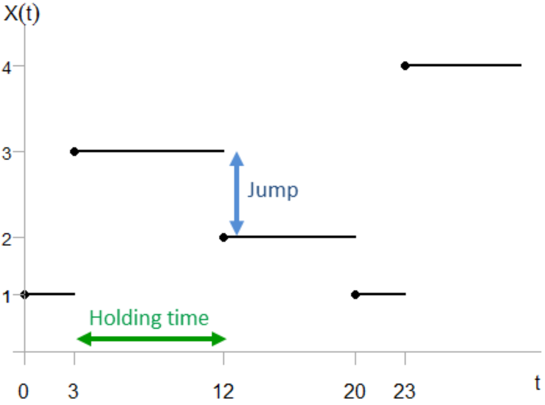
Background: Markov property







Background: describing a continuous-time Markov chain



Background: describing a continuous-time Markov chain

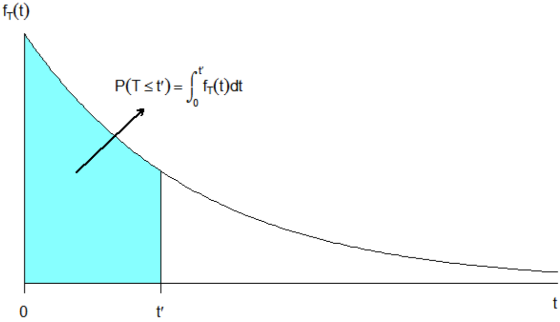
Holding time

T = amount of time the chain spends in state i (Exponential r.v.)

$$f_T(t) = \lambda_i e^{-\lambda_i t}, \qquad \lambda_i > 0, \quad t > 0$$

$f_T(t) : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$P(T \leq t') = \int_0^{t'} f_T(t)dt = 1 - e^{-\lambda_i t'} \qquad \forall t \geq 0$$



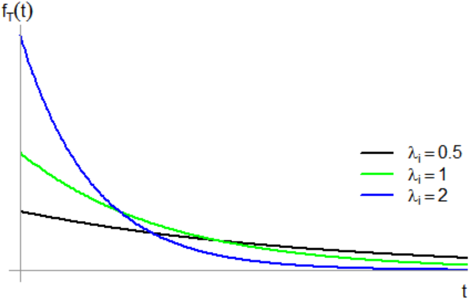
Background: describing a continuous-time Markov chain

Holding time

T = amount of time the chain spends in state i (Exponential r.v.)

$$f_T(t) = \lambda_i e^{-\lambda_i t}, \qquad \lambda_i > 0, \quad t > 0$$

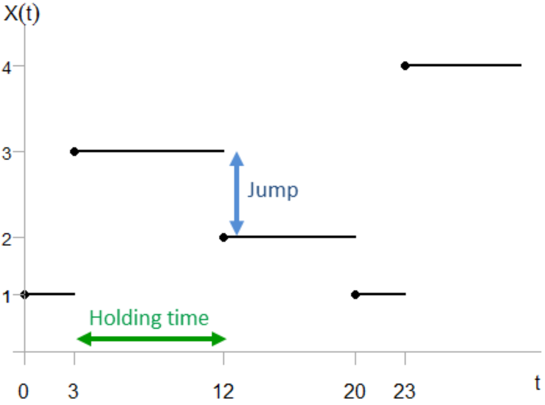
λ_i is the rate parameter



The Exponential r.v. has the *memoryless property*

$$P(T > s + t \mid T > t) = P(T > s) \quad \forall\, s, t > 0$$

Background: describing a continuous-time Markov chain



Background: describing a continuous-time Markov chain

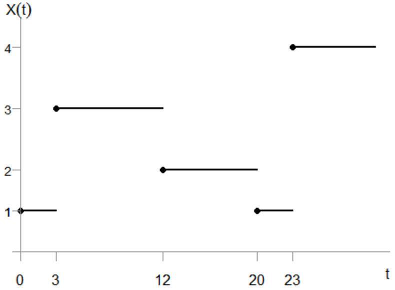
Jump chain

$P = (p_{ij} : i,j \in \mathcal{S}) = \textit{jump matrix}$

$$p_{ij} = P(X(t') = j | X(t) = i, \ \textit{the opportunity to leave } i)$$

$$p_{ij} \geq 0 \qquad \sum_{j \in \mathcal{S}} p_{ij} = 1 \qquad \forall i,j \in S$$

Background: describing a continuous-time Markov chain



Example

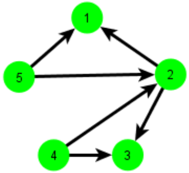
$$P = \begin{bmatrix} 0.1 & 0 & 0.6 & 0.3 \\ 0.8 & 0.1 & 0.1 & 0 \\ 0.05 & 0.5 & 0.05 & 0.4 \\ 0.6 & 0.1 & 0.15 & 0.15 \end{bmatrix}$$

Outline

Recall: adjacency matrix and directed relations

Social network: a set of actors $\mathcal{N}$ + a relation $\mathcal{R}$

Graph $= G(\mathcal{N}, \mathcal{R})$



Adjacency matrix$=X$

-	0	0	0	0
1	-	1	0	0
0	0	-	0	0
0	1	1	-	0
1	1	0	0	-

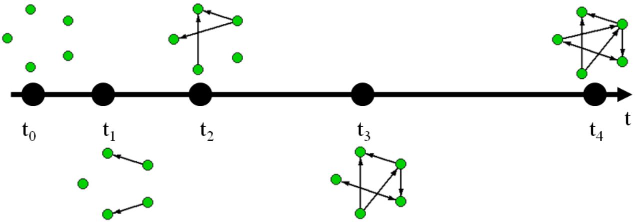
Directed relation:

$\not\Rightarrow$

$i \to j$

$j \to i$

Data



Longitudinal (or panel) network data = M (≥ 2) repeated observations on a network

$$x(t_0),\ x(t_1),\ \ldots,\ x(t_m),\ \ldots,\ x(t_{M-1}),\ x(t_M)$$

- set of actors $\mathcal{N} = \{1, 2, \ldots, n\}$
- a non reflexive and directed relation $\mathcal{R}$
- actor covariates V (gender, age,social status, ...)

Model definition: assumptions

Network evolution is the outcome of a **stochastic process** specified by the following assumptions:

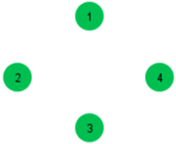
1. *Ties are state*:
   a tie is a state with a tendency to endure over time

2. *Distribution of the process*:
   $\{X(t), t \in \mathcal{T}\}$ is a continuous time Markov Chain defined on:
   - the state space $\mathcal{X}$
   - the set of actors $\mathcal{N}$

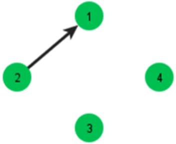
Model definition: assumptions

Finite state space: $\mathcal{X}$ is the set of all possible adjacency matrices defined on $\mathcal{N}$

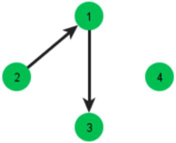
$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \textit{ is a countable set}$$



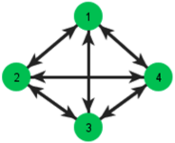
	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

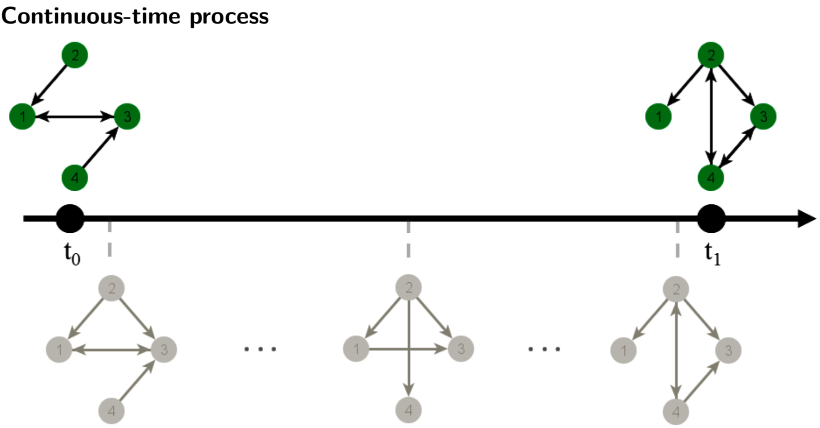


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	1	1	1
2	1	-	1	1
3	1	1	-	1
4	1	1	1	-

Model definition: assumptions



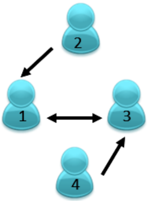
Latent process: the network evolves in continuous-time but we observed it only at discrete time points

Model definition: assumptions

Markov property: the current state of the network determines probabilistically its further evolution

Model definition: assumptions

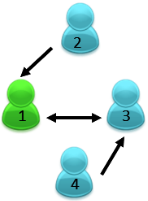
3. *Opportunity to change*: at any given moment t one actor has the opportunity to change



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

3. *Opportunity to change*: at any given moment t one actor has the opportunity to change

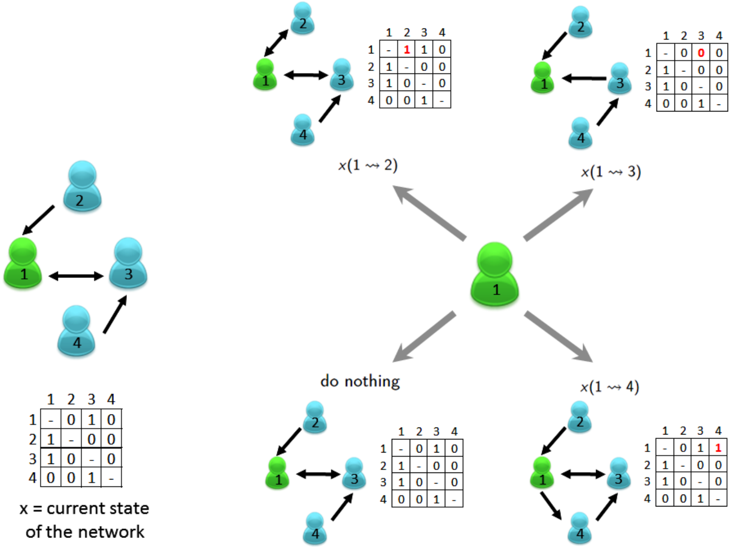


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t

   (Notation: $x(i \rightsquigarrow j)$ means that actor i changes his outgoing tie towards j)



Model definition: assumptions

5. *Actor-oriented perspective*: actors control their outgoing ties

    - change in ties are made by the actor who sends the ties

    - decisions are made according to the position of the actor in the network, his attributes and the characteristics of the others

      Aim: maximize *a utility function*

    - actors have complete knowledge about the network and all the other actors

    - the maximization is based on immediate returns (myopic actors)

Model definition: assumptions (recap)

1. *Ties are states*

2. *The evolution process is a continuous-time Markov chain*

3. *At any given moment t one probabilistically selected actor has the opportunity to change*

4. *No more than one tie can change at any given moment t*

5. *Actor-oriented perspective*

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps

Micro-step	Continuous-time Markov chain
- the time at which *i* had the opportunity to change	- the waiting time until the next opportunity for a change made by an actor *i* (holding time)
- the precise change *i* made	- the probability of changing the link x_{ij} given that *i* is allowed to change (jump chain)

Distribution of the holding time: *rate function*

Transition matrix of the jump chain: *objective function*

Model definition: rate function

How fast is the opportunity for changing?

Waiting time between opportunities of change for actor $i \sim Exp(\lambda_i)$

λ_i is called the rate function

Simple specification: all actors have the same rate of change λ

$$P(i \text{ has the opportunity of change}) = \frac{1}{n} \quad \forall i \in \mathcal{N}$$

Model definition: rate function

How fast is the opportunity for changing?

More complex specification

Actors may change their ties at different frequencies $\lambda_i(\alpha, x, v)$

Example

"Young girls might change their ties more frequently"

$$\lambda_i(\alpha, x, v) = \alpha_{age} * v_{age} + \alpha_{gender} * v_{gender}$$

It follows

$$P(i \ has \ the \ opportunity \ of \ change) = \frac{\lambda_i(\alpha, x, v)}{\sum_{j=1}^{n} \lambda_j(\alpha, x, v)}$$

Model definition: rate function

How fast is the opportunity for changing?

In the following we assume that:

- all actors have the same rate of change

  $\Longrightarrow \quad \lambda$ is constant over the actors

- the frequencies at which actors have the opportunity to make a change depends on time

  $\Longrightarrow \lambda$ is not constant over time

As a consequence, we must specify $M-1$ rate functions

$$\lambda_1, \; \cdots, \; \lambda_{M-1}$$

Model definition: objective function

Which tie is changed?

Changing a tie means turning it into its opposite:

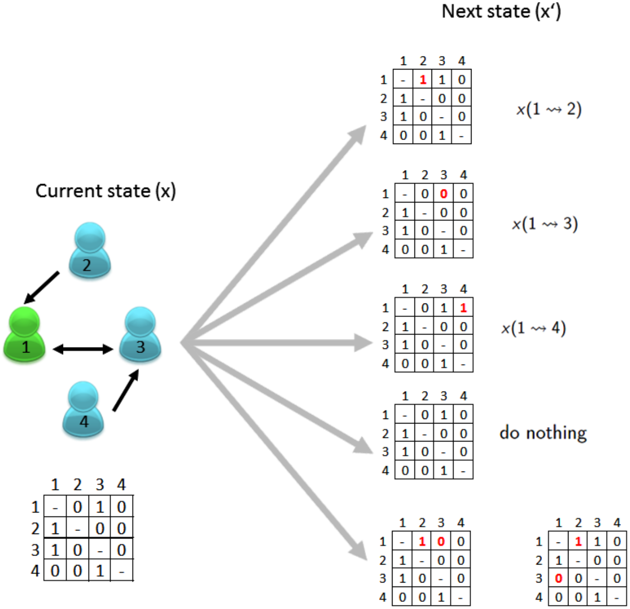
$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

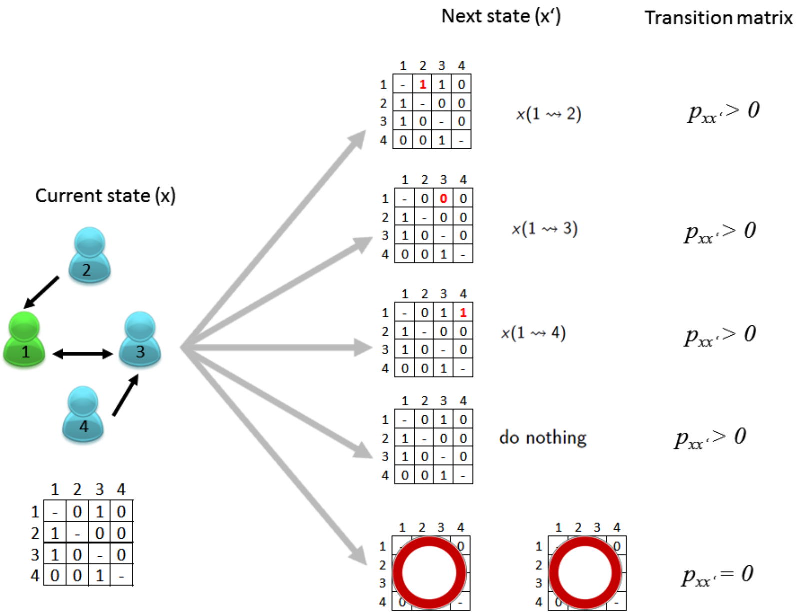
Given that i has the opportunity to change:

Possible choices of i	Possible reachable states
$n - 1$ changes	$n - 1$ networks $x(i \rightsquigarrow j)$
1 non-change	1 network equal to x

Model definition: objective function



Model definition: objective function



Background: random utility model

Setting:
decision makers who face a choice between N-alternatives

Notation:

i denotes the decision maker

$J = \{1, \ldots, j, \ldots, N\}$ choice set
J is **exhaustive** and choices are **mutually exclusive**

Assumption:
the decision makers obtain a certain level of profit from each alternative.
The profit is modeled by the *utility function* $U_{ij} : J \to \mathbb{R}$

Decision rule: i chooses the alternative j that assures him the highest profit, i.e.

$$j \ : \ max_{j \in J} \ U_{ij}$$

Background: random utility model

The researcher does not observe the decision maker's utility, but only:

- $n \times A$ matrix x of attributes of each alternative j (as faced by i)
- $B \times 1$ vector v_i of attributes of i

Since, there are factors that the researcher cannot observe, the utility function is decomposed as

$$U_{ij} = F_{ij}(\beta, \gamma, x_{ij}, v_i) + \mathcal{E}_{ij}$$

where:

- F_{ij} is the deterministic part of the utility (observed!)

$$F_{ij}(\beta, \gamma, x_{ij}, v_i) = \sum_a \beta_a x_{ija} + \sum_{b=1} \gamma_b v_{ib}, \quad \beta_a, \ \gamma_b \in \mathbb{R}, x_{ij}$$

- $\mathcal{E}_{ij}$: random term (not observed!)
  The random term are independent and identically distributed.

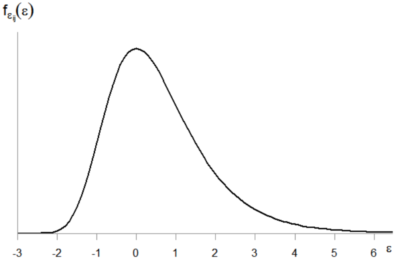
Consequence: The researcher can only "guess" i's choice

Background: random utility model

Decision probabilities:

it is assumed that $\mathcal{E}_{ij}$ is Gumbel distributed

$$f_{\mathcal{E}_{ij}}(\epsilon) = e^{-\epsilon} e^{-e^{-\epsilon}} \qquad \epsilon \in \mathbb{R}$$



so that the probability that i chooses the alternative j is given by

$$p_{ij} = P(U_{ij} > U_{ih},\ \forall\ h \in J) = \frac{e^{F_{ij}}}{\sum_{h=1}^{N} e^{F_{ih}}}$$

Model definition: objective function

Actors change their ties in order to maximize a utility function

$$u_i(\beta, x(i \rightsquigarrow j)) = f_i(\beta, x(i \rightsquigarrow j), v_i, v_j) + \mathcal{E}_{ij}$$

- $f_i(\beta, x(i \rightsquigarrow j), v_i, v_j)$ is the *objective function*
- $\mathcal{E}_{ij}$ is assumed to be distributed as a Gumbel r.v.

Consequence: the probability that i changes his outgoing tie towards j is:

$$p_{ij} = \frac{exp\left(f_i(\beta, x(i \rightsquigarrow j), v_i, v_j))\right)}{\sum_{h=1}^{n} exp\left(f_i(\beta, x(i \rightsquigarrow h), v_i, v_j)\right)}$$

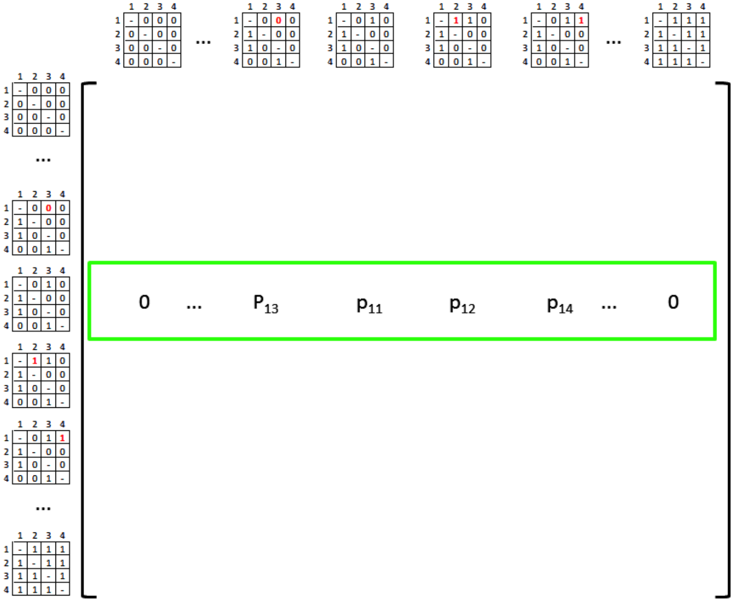
Probabilities interpretation:

p_{ij} is the probability that i changes the tie towards j

p_{ii} is the probability of not changing

Model definition: objective function

Example



Model definition: objective function

The objective function is defined as a linear combination

$$f_i(\beta, x(i \rightsquigarrow j), v_i, v_j) = \sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow j), v_i, v_j)$$

- $s_{ik}(x(i \rightsquigarrow j), v_i, v_j)$ is called effect
- $\beta_k \in \mathbb{R}$ is a statistical parameter

N.b.
In the following, we will write:

- x' instead of $x(i \rightsquigarrow j)$
- $s_{ik}(x', v)$ instead of $s_{ik}(x(i \rightsquigarrow j), v_i, v_j)$

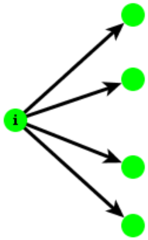
to simplify the notation

Objective function specification

Endogenous effects = *dependent on the network structures*

- Outdegree effect

$$s_{i\_out}(x') = \sum_j x'_{ij}$$



- Reciprocity effect

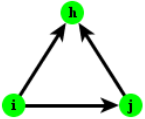
$$s_{i\_rec}(x') = \sum_j x'_{ij} x'_{ji}$$

Objective function specification

Endogenous effects = *dependent on the network structures*

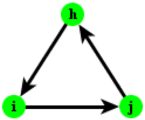
- Transitive effect

$$s_{i\_trans}(x') = \sum_{j,h} x'_{ij} x'_{ih} x'_{jh}$$



- three cycle-effect

$$s_{i\_cyc}(x') = \sum_{j,h} x'_{ij} x'_{jh} x'_{hi}$$



Objective function specification

Exogenous effects = *related to actor's attributes*

Example

- Friendship among pupils:

  Smoking: non, occasional, regular

  Gender: boys, girls

- Trade/Trust (Alliances) among countries:

  Geographical area: Europe, Asia, North-America,...

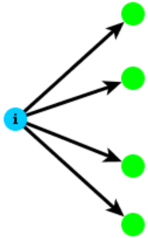
  Worlds: First, Second, Third, Fourth

Objective function specification

Exogenous effects *(individual covariate)*

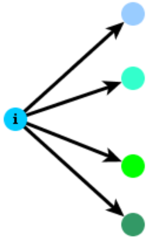
- covariate-ego

$$s_{i\_cego}(x', v) = \sum_j x'_{ij} v_i$$



- covariate-alter

$$s_{i\_calt}(x', v) = \sum_j x'_{ij} v_j$$

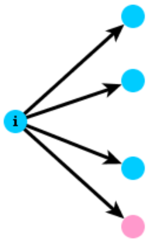


Objective function specification

Exogenous effects *(dyadic covariate)*

- covariate-related similarity

$$s_{i\_csim}(x', v) = \sum_j x'_{ij} \left(1 - \frac{|v_i - v_j|}{R_V}\right)$$



where R_V is the range of V and $\left(1 - \frac{|v_i - v_j|}{R_V}\right)$ is called *similarity score*

Remark:

when V is a binary covariate, the covariate-related similarity can be written in the following way:

$$s_{i\_csim}(x', v) = \sum_j x'_{ij} \mathbb{I}\{v_i = v_j\}$$

Objective function specification

Which effects must be included in the objective function?

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Example

Friendship network

Theory		Effect
the friend of my friend is also my friend	$\Rightarrow$	transitive effect
girls trust girls boys trust boys	$\Rightarrow$	covariate-related similarity

Parameter interpretation

1. Parameter interpretation: β_k quantifies the role of $s_{ik}(x')$ in the network evolution.

   - $\beta_k = 0$: $s_{ik}(x')$ plays no role in the network dynamics
   - $\beta_k > 0$: higher probability of moving into networks where $s_{ik}(x')$ is higher
   - $\beta_k < 0$: higher probability of moving into networks where $s_{ik}(x')$ is lower

2. The preferences driving the choice of the actors have the same intensities over time

$$\Longrightarrow \quad \beta_1, \cdots, \beta_K \text{ are constant over time}$$

Parameter interpretation

The procedures for estimating the parameters of the SAOM are implemented in a R library called ***RSiena***

(SIENA = **S**imulation **I**nvestigation for **E**mpirical **N**etwork **A**nalysis)

The R script "estimation.R" contains the R commands to implement the estimation procedure in R and the folder "tfls.zip" includes the data files.

Example data: an excerpt from the "Teenage Friends and Lifestyle Study" data set:

- Networks: relation = friendship
  actors = 129 pupils present at all three measurement points
- Covariates: gender (1 = Male, 2 = Female)
  smoking behavior (1 = no, 2= occasional, 3 = regular)

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Rate parameter: expected frequency, between two consecutive network observations, with which actors get the opportunity to change a network tie

- about 9 opportunities for change in the first period
- about 7 opportunities for change in the second period

The estimated rate parameters will be higher than the observed number of changes per actor (why?)

Parameter interpretation: a very simple model

Interpreting the objective function parameters:

The parameter β_k quantifies the role of the effect s_{ik} in the network evolution.

$\beta_k = 0$ s_{ik} plays no role in the network dynamics

$\beta_k > 0$ higher probability of moving into networks where s_{ik} is higher

$\beta_k < 0$ higher probability of moving into networks where s_{ik} is lower

Which β_k are "significantly" different from 0?

E.g. $\beta_{rec} = 0.13$ is "significantly" different from 0?

Parameter interpretation: a very simple model

Hypothesis test:

1. State the hypotheses.
   - The *null hypothesis* (H_0):
     the observed increase or decrease in the number of network configurations related to a certain effect results purely from chance

$$H_0 : \beta_k = 0$$

   - The *alternative hypothesis* (H_1):
     the observed increase or decrease in the number of network configurations related to a certain effect is influenced by some non-random cause.

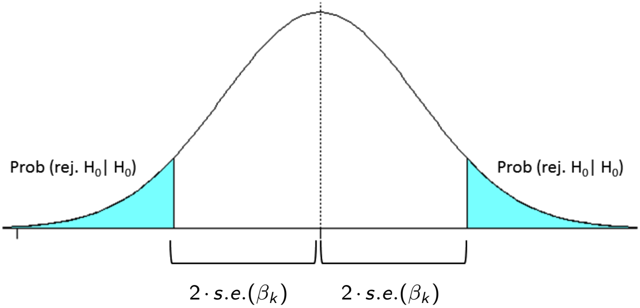
$$H_1 : \beta_k \neq 0$$

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left| \frac{\beta_k}{s.e.(\beta_k)} \right| \geq 2 & \textit{reject } H_0 \\ \left| \frac{\beta_k}{s.e.(\beta_k)} \right| < 2 & \textit{fail to reject } H_0 \end{cases}$$



$H_0 : \beta_k = 0$

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Objective function parameters:

- outdegree parameter: the observed networks have low density

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Objective function parameters:

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties

Parameter interpretation: a very simple model

In more detail

$$\beta_{out} \sum_{j=1}^{n} x'_{ij} + \beta_{rec} \sum_{j=1}^{n} x'_{ij} x'_{ji} = -2.4147 \sum_{j=1}^{n} x'_{ij} + 2.7106 \sum_{j=1}^{n} x'_{ij} x'_{ji}$$

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

$$-2.4147 + 2.7106 = 0.2959$$

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

$$-2.4147$$

Conclusion: reciprocated ties are valued positively and non-reciprocated ties are valued negatively by actors

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

- the statement "the friend of my friend is also my friend" is almost always true $\rightarrow$ transitive triplets effect

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- pupils prefer to establish friendship relations with others that are similar to themselves → covariate similarity

This effect must be controlled for the sender and receiver effects of the covariate.

- Covariate ego effect

- Covariate alter effect

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties
- transitivity parameter: preference for being friends with friends'friends

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- sex alter: gender does not affect actor popularity
- sex ego: gender does not affect actor activity
- sex similarity: tendency to choose friends with the same gender

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.

- All actor covariates are centered: $\overline{v} = 1.434$ is the mean of the covariate

$$v_i - \overline{v} = \begin{cases} -0.434 & \textit{for boys} \\ \\ 0.566 & \textit{for girls} \end{cases}$$

- The contribution of x_{ij} to the objective function is

$$\beta_{ego}(v_i - \overline{v}) + \beta_{alter}(v_j - \overline{v}) + \beta_{same}\left(\mathbb{I}\{v_i = v_j\} - sim_v\right) =$$

$$= 0.1571(v_i - \overline{v}) - 0.1513(v_j - \overline{v}) + 0.9191\left(\mathbb{I}\{v_i = v_j\} - 0.5048\right)$$

where sim_v is the average of the similarity score.

Parameter interpretation: a more complex model

	Male	Female
Male	0.4526	-0.618
Female	-0.309	0.4584

Table : Gender-related contributions to the objective function

Conclusions: Preference for intra-gender relationships.

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- smoke alter: smoking behavior does not affect actor popularity
- smoke ego: smoking behavior not affect actor activity
- smoke similarity: tendency to choose friends with the same smoking behavior

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{v} = 1.310$ is the mean of the covariate

$$v_i - \overline{v} = \begin{cases} -0.310 & \textit{for no smokers} \\ 0.690 & \textit{for occasional smokers} \\ 1.690 & \textit{for regular smokers} \end{cases}$$

- The contribution of x_{ij} to the objective function is

$$\beta_{ego}(v_i - \overline{v}) + \beta_{alter}(v_j - \overline{v}) + \beta_{same}\left(1 - \frac{|v_i - v_j|}{R_v} - sim_v\right) =$$

$$= 0.0714(v_i - \overline{v}) + 0.1055(v_j - \overline{v}) + 0.3724\left(1 - \frac{|v_i - v_j|}{2} - 0.7415\right)$$

Parameter interpretation: a more complex model

	no	occasional	regular
no	0.0414	-0.0734	-0.1882
occasional	-0.0393	0.2183	0.1035
regular	-0.1200	0.1376	0.3952

Table : Smoking-related contributions to the objective function

Conclusions:

- preference for similar alters
- this tendency is strongest for high values on smoking behavior

Simulating network evolution

Aim: given $x(t_0)$ and fixed parameter values, provide $x^{sim}(t_1)$ according to the process behind the SAOM

$$\Downarrow$$

reproduce a possible series of micro-steps between t_0 and t_1

Input

$n =$ number of actors

$\lambda =$ rate parameter (given)

$\beta = (\beta_1, \ldots, \beta_k) =$ objective function parameters (given)

$x(t_0) =$ network at time t_0 (given)

Output

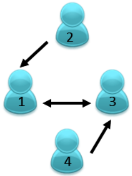
$x^{sim}(t_1) =$ network at time t_1

Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
    $dt \sim Exp(n\lambda)$
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
        $x \leftarrow x(i \rightsquigarrow j)$
    else
        $x \leftarrow x$
    $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(t_0)$

$n = 4$

$\lambda = 1.5$

$\beta = (\beta_{out}, \beta_{rec}, \beta_{trans})$
$= (-1, 0.5, -0.25)$

Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
    $dt \sim Exp(n\lambda)$
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
        $x \leftarrow x(i \rightsquigarrow j)$
    else
        $x \leftarrow x$
    $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

Generate the time elapsed between t_0 and the first opportunity to change

The more intuitive way to generate dt is:

- generate the waiting time for each actor i

$$w_i \sim Exp(\lambda)$$

- $dt = \min_{1 \le i \le n} \{w_i\}$

but this requires the generation of n numbers.

Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
  $dt \sim Exp(n\lambda)$
  $i \sim Uniform(1, \dots, n)$
  $j \sim Multinomial(p_{i1}, \dots, p_{in})$
  if $i \neq j$ **then**
    $x \leftarrow x(i \rightsquigarrow j)$
  else
    $x \leftarrow x$
  $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

Generate the time elapsed between t_0 and the first opportunity to change

To avoid the generation of n numbers, we use the following result:

If

$$W_i \sim Exp(\lambda_i), \quad 1 \leq i \leq n$$

and $W_1, \dots, W_n$ are mutually independent, then

$$DT = \min\{W_1, \dots, W_n\} \sim Exp(\sum_{i=1}^{n} \lambda_i)$$

e.g. $dt = 0.0027$

Simulating network evolution

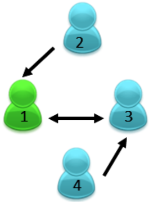
Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
  $dt \sim Exp(n\lambda)$
  $i \sim Uniform(1, \ldots, n)$
  $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
  if $i \neq j$ **then**
    $x \leftarrow x(i \rightsquigarrow j)$
  else
    $x \leftarrow x$
  $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

Select the actor i who has the opportunity to change

e.g. $i = 1$



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(t_0)$

Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while *condition* $= TRUE$ **do**
  $dt \sim Exp(n\lambda)$
  $i \sim Uniform(1, \ldots, n)$
  $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
  if $i \neq j$ **then**
    $x \leftarrow x(i \rightsquigarrow j)$
  else
    $x \leftarrow x$
  $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

Select j, the actor towards i is going to change his outgoing tie

$i \rightarrow j$	f_i	p_{ij}
$1 \rightarrow 1$	-1.75	0.15
$1 \rightarrow 2$	-1.00	0.31
$1 \rightarrow 3$	-3.25	0.03
$1 \rightarrow 4$	-0.5	0.51

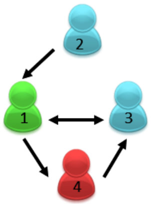
Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
  $dt \sim Exp(n\lambda)$
  $i \sim Uniform(1, \ldots, n)$
  $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
  if $i \neq j$ **then**
    $x \leftarrow x(i \rightsquigarrow j)$
  else
    $x \leftarrow x$
  $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

e.g. $j = 4$



	1	2	3	4
1	-	0	1	1
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 4)$

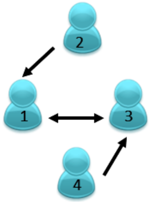
Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
  $dt \sim Exp(n\lambda)$
  $i \sim Uniform(1, \dots, n)$
  $j \sim Multinomial(p_{i1}, \dots, p_{in})$
  if $i \neq j$ **then**
    $x \leftarrow x(i \rightsquigarrow j)$
  else
    $x \leftarrow x$
  $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

e.g. $j = 1$



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 1)$

Simulating network evolution

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
  $dt \sim Exp(n\lambda)$
  $i \sim Uniform(1, \ldots, n)$
  $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
  if $i \neq j$ **then**
    $x \leftarrow x(i \rightsquigarrow j)$
  else
    $x \leftarrow x$
  $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

t = time
dt = holding time between consecutive opportunities to change
$\sim$ = generated from

e.g. $t = 0 + 0.0027$

Simulating network evolution

Two different stopping rules:

1. *Unconditional* simulation:

   the simulation of the network evolution carries on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:

   Simulation runs on until

$$\sum_{\substack{i,j=1\\ i\neq j}}^{n} \left| x_{ij}^{obs}(t_1) - x_{ij}(t_0) \right| = \sum_{\substack{i,j=1\\ i\neq j}}^{n} \left| x_{ij}^{sim}(t_1) - x_{ij}(t_0) \right|$$

   This criterion can be generalized conditioning on any other explanatory variable.

Simulating network evolution

Use of simulations:

- simulating the network evolution between two consecutive time points

  N.b.
  For simulations of 3 or more waves ($M \geq 2$), the simulations for wave $m+1$ start at the simulated network for wave m.

- provide possible scenarios of the network evolution according to different values of the parameters of the SAOM

- estimate the parameters of the SAOM

- evaluate the goodness of fit of the model

Estimating the parameter of the SAOM

Problem

Given the longitudinal network data

$$x(t_0),\ x(t_1), \ldots,\ x(t_M)$$

and a parametrization of the SAOM

$$\theta = (\lambda_1,\ \ldots, \lambda_M, \beta_1,\ \ldots, \beta_K)$$

we want to estimate θ in a plausible way.

Solution

Different estimation methods are available:

1. Method of Moments (MoM)
2. Maximum Likelihood Estimation (MLE)

Background: Method of Moments (MoM)

Definition

Let X be a random variable with probability distribution depending on a parameter θ.
Let $(x_1, \ldots, x_q)$ a sample of q observations from the r.v. X.

The **expected value (mean or moment)** of X, denoted by $E_\theta[X]$, is defined by:

$$E_\theta[X] = \sum_{x \in \mathcal{S}} x \cdot \varphi(x, \theta)$$

if X is discrete with p.m.f $\varphi(x, \theta)$ and

$$E_\theta[X] = \int_{x \in \mathcal{S}} x \cdot f(x, \theta) dx$$

if X is continuous with p.d.f $f(x, \theta)$

The **sample counterpart** of $E_\theta[X]$, denoted by μ, is defined by:

$$\mu = \frac{1}{q} \sum_{i=1}^{q} x_i$$

Background: Method of Moments (MoM)

Definition

The method of moment estimator for θ is found by equating the expected value $E_\theta[X]$ to its sample counterpart μ

$$E_\theta[X] = \mu$$

and solving the resulting equation for the unknown parameter.
The estimate for θ is denoted by $\widehat{\theta}$.

In practice:

1. Compute the expected value $E_\theta[X]$
2. Compute the sample counterpart $\mu = \frac{1}{q}\sum_{i=1}^{q} x_i$
3. Solve the moment equation $E_\theta[X] = \mu$ for θ

Motivation

One can observe that the expected value of a certain distribution usually depends on the parameter θ

Background: Method of Moments (MoM)

Example

Let W be the r.v. describing the waiting times between two consecutive opportunities for change for an actor in a network evolution process described by the SAOM.

A sample is reported in the following table:

	1	2	3	4	5	6	7	8	9	10
w_i	0.33	0.08	0.06	0.01	0.04	0.11	0.03	0.18	0.02	0.07

Estimate the rate parameter λ according to the MoM.

From the assumptions of the SAOM follows that $W \sim Exp(\lambda)$

$$f_W(w) = \lambda e^{-\lambda w} \qquad \lambda, w > 0$$

Background: Method of Moments (MoM)

Example

1. The expected value of W is:

$$\begin{aligned} E_\lambda[W] &= \int_0^{+\infty} w \cdot f_W(w)dw = \int_0^{+\infty} w \cdot \lambda e^{-\lambda w} dw \\ &= \underbrace{\left[-w \cdot e^{-\lambda w}\right]_0^{+\infty} - \int_0^{+\infty} -e^{-\lambda w} dw}_{integration\ \ by\ \ parts} \\ &= 0 - \left[-\frac{1}{\lambda} e^{-\lambda w}\right]_0^{+\infty} = \frac{1}{\lambda} \end{aligned}$$

Background: Method of Moments (MoM)

Example

	1	2	3	4	5	6	7	8	9	10
w_i	0.33	0.08	0.06	0.01	0.04	0.11	0.03	0.18	0.02	0.07

2. The sample counterpart is:

$$\mu = \frac{1}{10}\sum_{i=1}^{10} w_i = \frac{0.93}{10} = 0.093$$

3. The estimate for λ is the solution of:

$$E_\lambda[W] = \mu$$
$$\frac{1}{\lambda} = \mu$$

namely

$$\widehat{\lambda} = \frac{1}{\mu} = \frac{1}{0.093} = 10.75$$

Background: Generalizations of MoM

The principle of the MoM can be easily generalized to any function $s : \mathcal{S} \longmapsto \mathbb{R}$.

1. Expected value of $s(X)$:

$$E_\theta[s(X)] = \sum_{x \in \mathcal{S}} s(x)\varphi(x,\theta)$$

$$E_\theta[s(X)] = \int\limits_{x \in \mathcal{S}} s(x) f(x,\theta) dx$$

2. Corresponding sample moment:

$$\gamma = \frac{1}{q} \sum_{i=1}^{q} s(x_i)$$

3. Moment equation:

$$E_\theta[s(X)] = \gamma$$

The functions $s(X)$ are called *statistics*

Background: Generalizations of MoM

The MoM can be applied also in situations where $\theta = (\theta_1, \ldots, \theta_p)$.

1. Definition of p statistics $(s_1(X), \ldots, s_p(X))$

2. Definition of p moment conditions:

$$E_\theta[s_1(X)] = \gamma_1$$

$$E_\theta[s_2(X)] = \gamma_2$$

$$\ldots$$

$$E_\theta[s_p(X)] = \gamma_p$$

3. Solving the resulting equations for the unknown parameters

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1, \ldots, \lambda_M, \beta_1, \ldots, \beta_K)$$

In practice:

1. find $M + K$ statistics
2. set the theoretical expected value of each statistic equal to its sample counterpart
3. solve the resulting system of equations with respect to θ.

For simplicity, let us assume to have observed a network at two time points t_0 and t_1 and to condition the estimation on the first observation $x(t_0)$

1. Defining the statistics

The rate parameter λ describes the frequency at which changes can potentially happen.

$$s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{i,j=1}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

Reason

	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
s_λ	94	135	171

$\Rightarrow$ higher values of λ leads to higher values of s_λ

1. Defining the statistics

The parameter β_k quantifies the role played by each effect in the network evolution.

$$s_k(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{ik}(X(t_1))$$

Example

Let us consider the outdegree:

$$s_{out}(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{i\_out}(X(t_1)) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}(t_1)$$

	$\beta_{out} = -2.5$	$\beta_{out} = -2$	$\beta_{out} = -1.5$
s_{out}	195	214	234

$\Rightarrow$ higher values of β_{out} leads to higher values of s_{out}

1. Defining the statistics

Generalizing to M periods:

- Statistics for the rate function parameters

$$s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

...

$$s_{\lambda_M}(X(t_M), X(t_{M-1})|X(t_{M-1}) = x(t_{M-1})) = \sum_{i,j=1}^{n} \left| X_{ij}(t_M) - X_{ij}(t_{M-1}) \right|$$

- Statistics for the objective function parameters:

$$\sum_{m=1}^{M} s_{mk}(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{M} s_{mk}(X(t_m))$$

2. Setting the moment equations

The MoM estimator for θ is defined as the solution of the system of $M+K$ equations

$$\begin{cases} E_\theta\left[s_{\lambda_m}(X(t_m),X(t_{m-1})|X(t_{m-1})=x(t_{m-1}))\right]=s_{\lambda_m}(x(t_m),x(t_{m-1})) \\ E_\theta\left[\sum_{m=1}^{M} s_{mk}(X(t_m)|X(t_{m-1})=x(t_{m-1}))\right]=\sum_{m=1}^{M} s_{mk}(x(t_m)) \end{cases}$$

with $m=1,\ldots,M$ and $k=1,\cdots,K$

2. Setting the moment equations

Example

Let us assume to have observed a network at two time points

t_0

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

t_1

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

We want to model the network evolution according to the outdegree, the reciprocity and the transitivity effects

$$\theta = (\lambda, \beta_{out}, \beta_{rec}, \beta_{trans})$$

2. Setting the moment equations

Example

Statistics:

$$s_{\lambda}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{4} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

$$s_{out}\left(X(t_1)|X(t_0) = x(t_0)\right) = \sum_{i,j=1}^{4} X_{ij}(t_1)$$

$$s_{rec}\left(X(t_1)|X(t_0) = x(t_0)\right) = \sum_{i,j=1}^{4} X_{ij}(t_1) X_{ji}(t_1)$$

$$s_{trans}\left(X(t_1)|X(t_0) = x(t_0)\right) = \sum_{i,j,h=1}^{4} X_{ij}(t_1) X_{ih}(t_1) X_{jh}(t_1)$$

2. Setting the moment equations

Example

t_0:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

t_1:

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

Observed values of the statistics:

$$s_\lambda = 5$$

$$s_{out} = 6 \quad s_{rec} = 4 \quad s_{trans} = 2$$

2. Setting the moment equations

Example

We look for the value of θ that satisfies the system:

$$\begin{cases} E_\theta\left[s_\lambda(X(t_1), X(t_0)|X(t_0) = x(t_0))\right] = 5 \\ E_\theta\left[s_{out}\left(X(t_1)|X(t_0) = x(t_0)\right)\right] = 6 \\ E_\theta\left[s_{rec}\left(X(t_1)|X(t_0) = x(t_0)\right)\right] = 4 \\ E_\theta\left[s_{trans}\left(X(t_1)|X(t_0) = x(t_0)\right)\right] = 2 \end{cases}$$

3. Solving the moment equations

Simplified notation:

- S: $(M+K)$-dimensional vector of statistics
- s: $(M+K)$-dimensional vector of the observed values of the statistics

Consequently, the system of moment equations can be written as

$$E_\theta[S] = s$$

or equivalently as

$$E_\theta[S - s] = 0$$

Problem: analytical procedures cannot be applied to solve this system

Solution: stochastic approximation method
i.e. an iterative stochastic algorithm that attempt to find zeros of functions which cannot be analytically computed.

3. Solving the moment equations: stochastic approximation method

Given an initial guess θ_0 for the parameter θ, the procedure can be roughly depicted as follows:

$$
\begin{array}{ccccc}
\theta_0 & \xrightarrow{\textit{approximation}} & E_{\theta_0}[S-s] & \xrightarrow{\textit{update}} & \theta_1 \\
\theta_1 & \xrightarrow{\textit{approximation}} & E_{\theta_1}[S-s] & \xrightarrow{\textit{update}} & \theta_2 \\
\ldots & \xrightarrow{\textit{approximation}} & \ldots & \xrightarrow{\textit{update}} & \ldots \\
\theta_{i-1} & \xrightarrow{\textit{approximation}} & E_{\theta_{i-1}}[S-s] & \xrightarrow{\textit{update}} & \theta_i \\
\ldots & \xrightarrow{\textit{approximation}} & \ldots & \xrightarrow{\textit{update}} & \ldots
\end{array}
$$

until a certain criterion is satisfied

3. Solving the moment equations: stochastic approximation method

Example

Let us consider the *"Teenage Friends and Lifestyle Study"* data set.

We model the network evolution according to the following parameter

$$\theta = (\lambda_1, \lambda_2, \beta_{out}, \beta_{rec}, \beta_{trans})$$

The MoM equations are:

$$\begin{cases} E_\theta\left[s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0))\right] = 477 \\ E_\theta\left[s_{\lambda_2}(X(t_2), X(t_1)|X(t_1) = x(t_1))\right] = 437 \\ E_\theta\left[s_{out}(X(t_1)|X(t_0) = x(t_0))\right] = 909 \\ E_\theta\left[s_{rec}(X(t_1)|X(t_0) = x(t_0))\right] = 548 \\ E_\theta\left[s_{trans}(X(t_1)|X(t_0) = x(t_0))\right] = 1146 \end{cases}$$

3. Solving the moment equations: stochastic approximation method

Example

- Guess $\theta_0 = (7.45, 6.83, -1.61, 0, 0)$

- Simulate the network evolution 1000 times according to $\widehat{\theta}_0$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 605.745$ $\quad$ $\overline{S}_{\lambda_2} = 573.715$

  $\overline{S}_{\beta_{out}} = 1151.886$ $\quad$ $\overline{S}_{\beta_{rec}} = 141.406$ $\quad$ $\overline{S}_{\beta_{trans}} = 270.118$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = 128.745$ $\quad$ $\overline{S}_{\lambda_2} - 437 = 136.715$

  $\overline{S}_{\beta_{out}} - 909 = 242.886$ $\quad$ $\overline{S}_{\beta_{rec}} - 548 = -406.594$ $\quad$ $\overline{S}_{\beta_{trans}} - 1146 = -875.882$

3. Solving the moment equations: stochastic approximation method

Example

- Guess $\theta_1 = (7.1, 6.75, -1.70, 1.20, 0.25)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_1}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 549.787$ $\quad$ $\overline{S}_{\lambda_2} = 532.551$

  $\overline{S}_{\beta_{out}} = 1478.988$ $\quad$ $\overline{S}_{\beta_{rec}} = 517.450$ $\quad$ $\overline{S}_{\beta_{trans}} = 1062.537$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = 72.787$ $\quad$ $\overline{S}_{\lambda_2} - 437 = 95.551$

  $\overline{S}_{\beta_{out}} - 909 = 569.988$ $\quad$ $\overline{S}_{\beta_{rec}} - 548 = -30.550$ $\quad$ $\overline{S}_{\beta_{trans}} - 1146 = -83.463$

3. Solving the moment equations: stochastic approximation method

Example

- Guess $\theta_2 = (7.10, 6.75, -2.20, 1.40, 0.35)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_2}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 446.853$ $\overline{S}_{\lambda_2} = 437.166$

  $\overline{S}_{\beta_{out}} = 1025.729$ $\overline{S}_{\beta_{rec}} = 414.484$ $\overline{S}_{\beta_{trans}} = 698.734$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = -30.147$ $\overline{S}_{\lambda_2} - 437 = 0.166$

  $\overline{S}_{\beta_{out}} - 909 = 116.729$ $\overline{S}_{\beta_{rec}} - 548 = -133.516$ $\overline{S}_{\beta_{trans}} - 1146 = -447.266$

and so on...

3. Solving the moment equations: stochastic approximation method

Example

- Guess $\theta_i = (10.71, 8.79, -2.63, 2.16, 0.46)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_i}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 476.022$ $\quad$ $\overline{S}_{\lambda_2} = 436.983$

  $\overline{S}_{\beta_{out}} = 906.809$ $\quad$ $\overline{S}_{\beta_{rec}} = 545.578$ $\quad$ $\overline{S}_{\beta_{trans}} = 1147.795$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = -0.978$ $\quad$ $\overline{S}_{\lambda_2} - 437 = -0.017$

  $\overline{S}_{\beta_{out}} - 909 = -2.191$ $\quad$ $\overline{S}_{\beta_{rec}} - 548 = -2.422$ $\quad$ $\overline{S}_{\beta_{trans}} - 1146 = 1.795$

3. Solving the moment equations: stochastic approximation method

1. Approximation

Definition

Let X be a random variable with distribution function $f_X(x)$.
The Monte Carlo method consists in:

1. generating a sample $(x_1, \cdots, x_q)$ from the distribution function $f_X(x)$
2. computing $s(x_l), \ l = 1, \ldots, \ q$
3. approximating the expected value with the empirical average, i.e.:

$$\overline{S} = \frac{1}{q} \sum_{l=1}^{q} s(x_l)$$

Reason

It can be proved that

$$\overline{S} \rightarrow E[s(X)]$$

as $q \rightarrow \infty$

3. Solving the moment equations: stochastic approximation method

1. Approximation

1. Given $x(t_0)$ and θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

2. For each sequence compute the value $S^{(l)}$ taken by S

3. Approximate the expected value by

$$\overline{S} = \frac{1}{q} \sum_{l=1}^{q} S^{(l)} \rightarrow E_\theta[S]$$

3. Solving the moment equations: stochastic approximation method

1. Approximation

Example

Approximating $E_\theta[s_{out}(X(t_1)|X(t_0) = x(t_0))]$ for the *"Teenage Friends and Lifestyle Study"* data set

1. Given:

   - $x(t_0)$
   - $\theta = (\lambda_1 = 10.69, \lambda_2 = 8.82, \beta_{out} = -2.63, \beta_{rec} = 2.17, \beta_{trans} = 0.46)$

   simulate the network evolution $q = 1000$ times

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ x^{(q)}(t_M)$$

3. Solving the moment equations: stochastic approximation method

1. Approximation

Example

2. Compute the value assumed by S_{out} for each sequence of networks

$$S_{out}^{(l)} = \sum_{m=1}^{M-1} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}^{(l)}(t_m)$$

sim	1	2	3	4	5	6	7	8	...
Nr. Edges	942	874	1047	881	865	866	999	948	...

3. Solving the moment equations: stochastic approximation method

1. Approximation

Example

3. Approximate the expected value by

$$\overline{S}_{out} = \frac{1}{q} \sum_{i=1}^{q} S_{out}^{(l)}$$

$$\overline{S}_{out} = \frac{942 + 874 + 1047 + 881 + 865 + 866 + 999 + 948 + \dots}{1000} \approx 912$$

3. Solving the moment equations: stochastic approximation method

2. Updating rule

The (modified) Robbins-Monro (RM) algorithm
Iterative algorithm to find the solution to

$$E_\theta[S] = s$$

The value of θ is iteratively updated according to:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \widehat{D}^{-1} \left(E_{\widehat{\theta}_i}[S] - s \right)$$

where:

- a_i is a series such that

$$\lim_{i \to \infty} a_i = 0 \qquad \sum_{i=1}^{\infty} a_i = \infty \qquad \sum_{i=1}^{\infty} a_i^2 < \infty$$

- $\widehat{D}$ is a diagonal matrix with elements

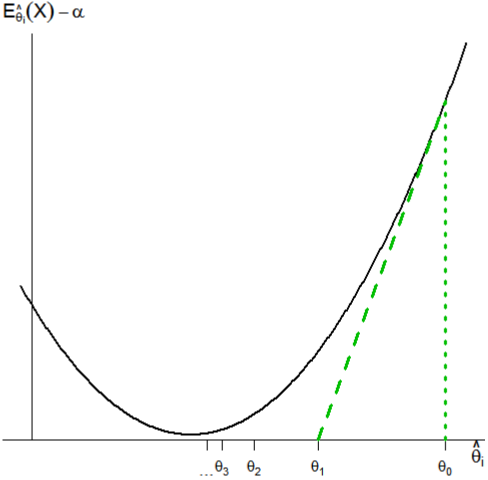
$$\widehat{D} = \frac{\partial}{\partial \widehat{\theta}_i} E_{\widehat{\theta}_i}[S]$$

3. Solving the moment equations: stochastic approximation method

2. Updating rule

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \widehat{D}^{-1} \left(E_{\widehat{\theta}_i}[S] - s \right)$$

Intuitively:

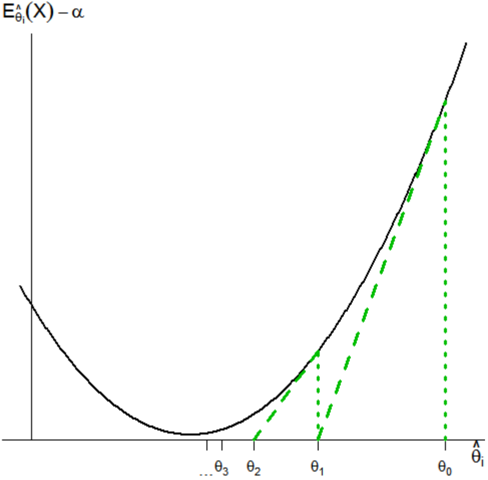


3. Solving the moment equations: stochastic approximation method

2. Updating rule

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \widehat{D}^{-1} \left(E_{\widehat{\theta}_i}[S] - s \right)$$

Intuitively:

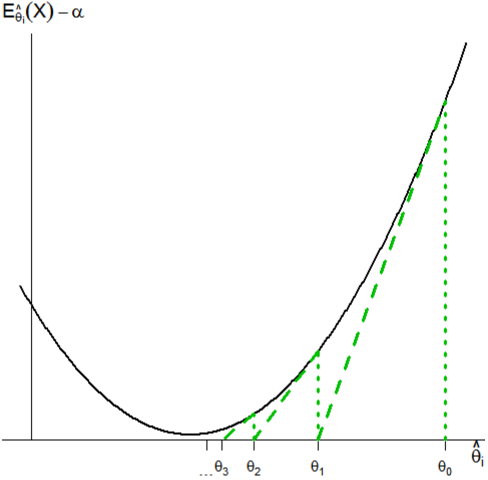


3. Solving the moment equations: stochastic approximation method

2. Updating rule

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \widehat{D}^{-1}\left(E_{\widehat{\theta}_i}[S] - s\right)$$

Intuitively:

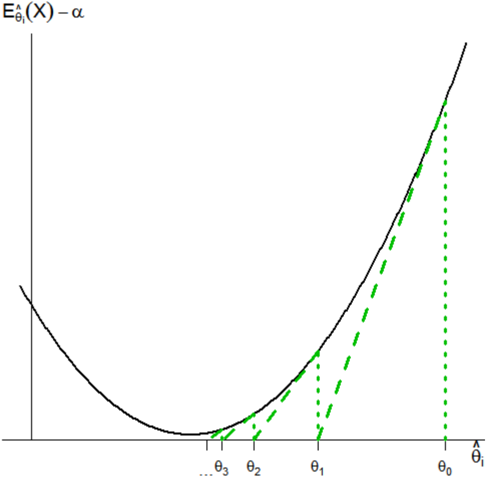


3. Solving the moment equations: stochastic approximation method

2. Updating rule

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \widehat{D}^{-1} \left(E_{\widehat{\theta}_i}[S] - s \right)$$

Intuitively:



Estimating the parameter of the SAOM

Issue

Given

$$x(t_0),\ x(t_1), \ldots,\ x(t_M)$$

and a parametrization of the SAOM

$$\theta = (\lambda_1,\ \ldots, \lambda_{M-1}, \beta_1,\ \ldots, \beta_K)$$

we want to estimate θ in a plausible way.

Different estimation methods are available:

1. Method of Moments:
   an estimation for θ is the value $\widehat{\theta}$ that solves:

   $$E_\theta[S - s] = 0$$

2. Maximum-likelihood estimation:
   what is the most likely value of θ that could have generated the observed data?

Background: the Maximum-likelihood estimation (MLE)

Definition

Suppose that X is a r.v. with p.m.f $\varphi(x,\theta)$, if X is discrete, or with p.d.f. $f(x,\theta)$, if X is continuous, where $\theta \in \Theta\mathbb{R}^k$. Let $x = (x_1, x_2, \ldots, x_q)$ be the observed value of a random sample of size q.

The **likelihood function** associated with the observed data is:

$$L(\theta) : \Theta \to \mathbb{R}; \quad \theta \longmapsto P_\theta(x_1, \ldots, x_q)$$

defined as:

$$L(\theta) = \prod_{i=1}^{q} \varphi(x_i, \theta)$$

if X is discrete

$$L(\theta) = \prod_{i=1}^{q} f(x_i, \theta)$$

if X is continuous

A parameter vector $\widehat{\theta}$ maximizing L:

$$\widehat{\theta} = arg \max_{\theta \in \Theta} L(\theta)$$

is called a **maximum likelihood estimate** for θ

Background: the Maximum-likelihood estimation (MLE)

In practice, it is easier to compute $\widehat{\theta}$ using the *log-likelihood function*, i.e. $log(L(\theta))$

$$\widehat{\theta} = arg \max_{\theta \in \Theta} log(L(\theta))$$

N.b.

The logarithm is a monotonic increasing function

Background: the Maximum-likelihood estimation (MLE)

Example

Let W be the r.v. describing the waiting times between two consecutive opportunities for change for an actor in a network evolution process described by the SAOM.
A sample is reported in the following table:

	1	2	3	4	5	6	7	8	9	10
w_i	0.33	0.08	0.06	0.01	0.04	0.11	0.03	0.18	0.02	0.07

Estimate the rate parameter λ according to the MLE.

From the assumptions of the SAOM follows that $W \sim Exp(\lambda)$

$$f_W(w, \lambda) = \lambda e^{-\lambda w} \qquad \lambda, w > 0$$

Background: the Maximum-likelihood estimation (MLE)

Example

Finding an estimate for θ requires:

1. computing the (log-)likelihood of the evolution process
2. maximizing the (log-)likelihood

1. Computing the likelihood of the evolution process

$$L(\lambda) = \prod_{i=1}^{q} f_W(w_i, \lambda) = \prod_{i=1}^{q} \lambda e^{-\lambda w_i} = \lambda^q e^{-\lambda \sum_{i=1}^{q} w_i}$$

$$log(L(\lambda)) = log\left(\lambda^q e^{-\lambda \sum_{i=1}^{q} w_i}\right) = q \cdot log(\lambda) - \lambda \sum_{i=1}^{q} w_i$$

Background: the Maximum-likelihood estimation (MLE)

Example

Finding an estimate for θ requires:

1. computing the (log-)likelihood of the evolution process
2. maximizing the (log-)likelihood

2. Maximizing the (log-)likelihood

$$\begin{aligned}
\frac{\partial}{\partial\lambda} log(L(\lambda)) &= 0 \\
\frac{q}{\lambda} - \sum_{i=1}^{q} w_i &= 0 \\
\lambda &= \frac{q}{\sum_{i=1}^{q} w_i} \quad (\textit{stationary point})
\end{aligned}$$

Checking that this stationary point is a maximum

$$\frac{\partial^2}{\partial\lambda^2} log(L(\lambda)) = -\frac{q}{\lambda^2} < 0$$

Therefore, $\widehat{\lambda} = 10.75$

Estimating the parameter of the SAOM using MLE

Let

- $$\mathcal{F} = \{F(\theta), \theta \in \Theta \subseteq \mathbb{R}^k\}$$

  be a collection of SAOMs parametrized by $\theta \in \Theta \subseteq \mathbb{R}^k$

- $x(t_0), \ldots, x(t_M)$ be the observed network data
- $V_1, \ldots, V_H$ be the observed actor attributes

The likelihood function associated with the observed data is:

$$L : \Theta \to \mathbb{R}; \quad \theta \longmapsto P_\theta(x(t_0), \ldots, x(t_M))$$

1. Computing the (log-)likelihood of the evolution process

For semplicity, let us consider only two observations $x(t_0)$ and $x(t_1)$

The model assumptions allow to decompose the process in a series of micro-steps:

$$\{(T_r, i_r, j_r), r = 1, \ \dots, R\}$$

- T_r: time point for an opportunity for change

$$t_0 < T_1 < \dots < T_R < t_1$$

- i_r: actor who has the opportunity to change
- j_r: actor towards whom the tie is changed

Given the sequence $\{(T_r, i_r, j_r), r = 1, \dots, R\}$, the likelihood of the evolution process

$$logL(\theta) = log\left(\prod_{r=1}^{R} P_\theta((T_r, i_r, j_r))\right) \propto log\left(\frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{n} p_{i_r j_r}(\beta, x(T_r))\right)$$

2. Maximizing the (log-)likelihood

Example

Let us consider the *"Teenage Friends and Lifestyle Study"* data set.

We model the network evolution according to the following parameter

$$\theta = (\lambda_1, \lambda_2, \beta_{out}, \beta_{rec}, \beta_{trans})$$

We look for $\widehat{\theta}$ such that:

$$\begin{cases} \frac{\partial}{\partial \lambda_1} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \lambda_2} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \beta_{out}} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \beta_{rec}} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \beta_{trans}} log(L(\theta)) = 0 \end{cases}$$

2. Maximizing the (log-)likelihood

Problem:

we cannot observe the complete data, i.e., the complete series of micro-steps that lead from $x(t_0)$ to $x(t_1)$, from $x(t_1)$ to $x(t_2)$, ...

$$\Downarrow$$

we cannot compute the L of the observed data

$$\Downarrow$$

a stochastic approximation method must be applied.

2. Maximizing the (log-)likelihood

Given an initial guess θ_0 for the parameter θ, the procedure can be roughly depicted as follows:

$$\begin{array}{ccccc}
\theta_0 & \xrightarrow{\textit{approximation}} & \frac{\partial}{\partial\theta} log(L(\theta_0)) & \xrightarrow{\textit{update}} & \theta_1 \\
\theta_1 & \xrightarrow{\textit{approximation}} & \frac{\partial}{\partial\theta} log(L(\theta_1)) & \xrightarrow{\textit{update}} & \theta_2 \\
\dots & \xrightarrow{\textit{approximation}} & \dots & \xrightarrow{\textit{update}} & \dots \\
\theta_{i-1} & \xrightarrow{\textit{approximation}} & \frac{\partial}{\partial\theta} log(L(\theta_{i-1})) & \xrightarrow{\textit{update}} & \theta_i \\
\dots & \xrightarrow{\textit{approximation}} & \dots & \xrightarrow{\textit{update}} & \dots
\end{array}$$

until a certain criterion is satisfied

2. Maximizing the (log-)likelihood

1. Approximation

To approximate the (log-)likelihood we use the *augmented data method*

Definition

The *augmented data* (or *sample path*) consist of the sequence of tie changes that brings the network from $x(t_0)$ to $x(t_1)$

$$(i_1, j_1), \ldots, (i_R, j_R)$$

Formally:

$$\underline{v} = \{(i_1, j_1), \ldots, (i_R, j_R)\} \in \mathcal{V}$$

where $\mathcal{V}$ is the set of all sample paths connecting $x(t_0)$ and $x(t_1)$.

We can approximate the (log-)likelihood function (and then the score function) of the observed data using the probability of $\underline{v}$

$$log P(\underline{v}|x(t_0), x(t_1)) \propto log \left(\frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{n} p_{i_r j_r}(\beta, x(T_r)) \right)$$

2. Maximizing the (log-)likelihood

2. Updating rule

We would like to solve the equation:

$$\frac{\partial}{\partial\theta} log(L(\theta)) = 0$$

Given $\widehat{\theta}_i$ and the corresponding approximation of the score function:

$$\frac{\partial}{\partial\theta} log(L(\widehat{\theta}_i; v_m^{(i)}))$$

we update the parameter estimate using the Robbins-Monro step

$$\theta_{i+1} = \theta_i + a_i D^{-1} \frac{\partial}{\partial\theta} log(L(\widehat{\theta}_i; v_m^{(i)}))$$

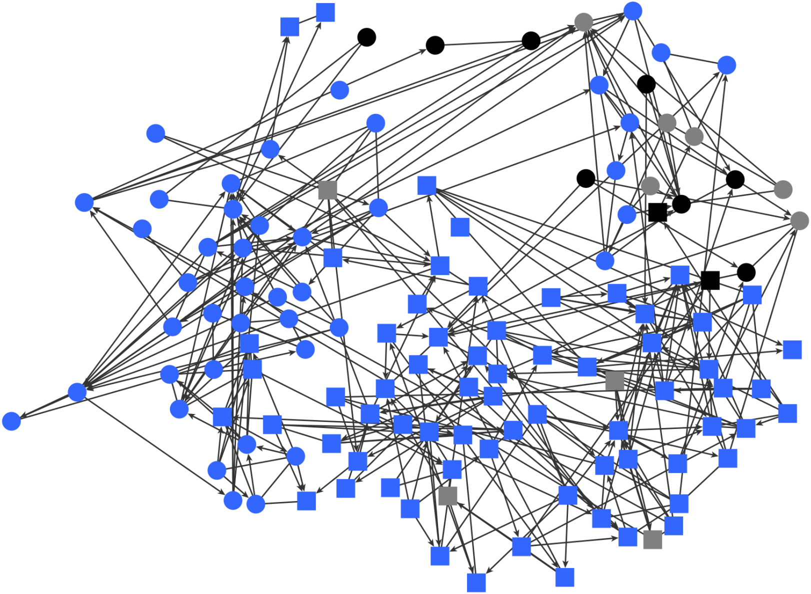
where D is a diagonal matrix with elements

$$D^{-1} = \left[\frac{\partial^2}{\partial\theta^2} log(L(\widehat{\theta}_i; v_m^{(i)})) \right]^{-1}$$

Outline

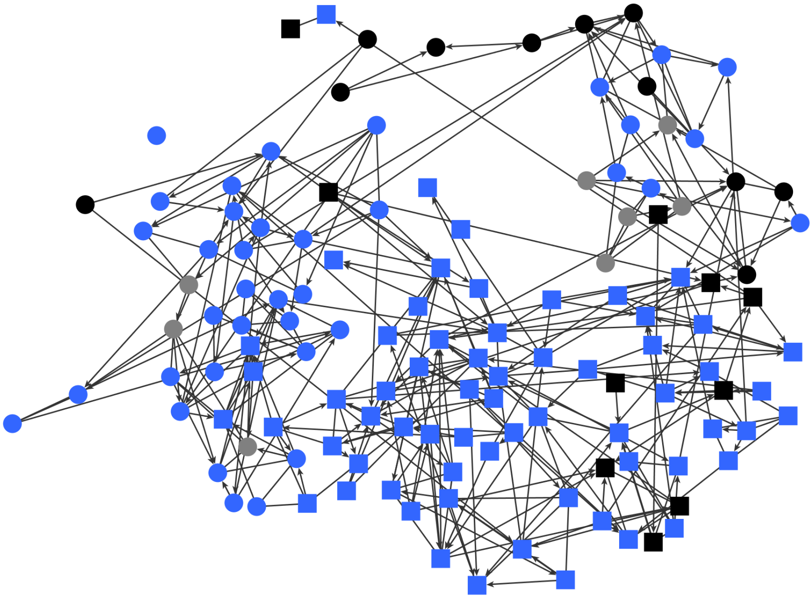
Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



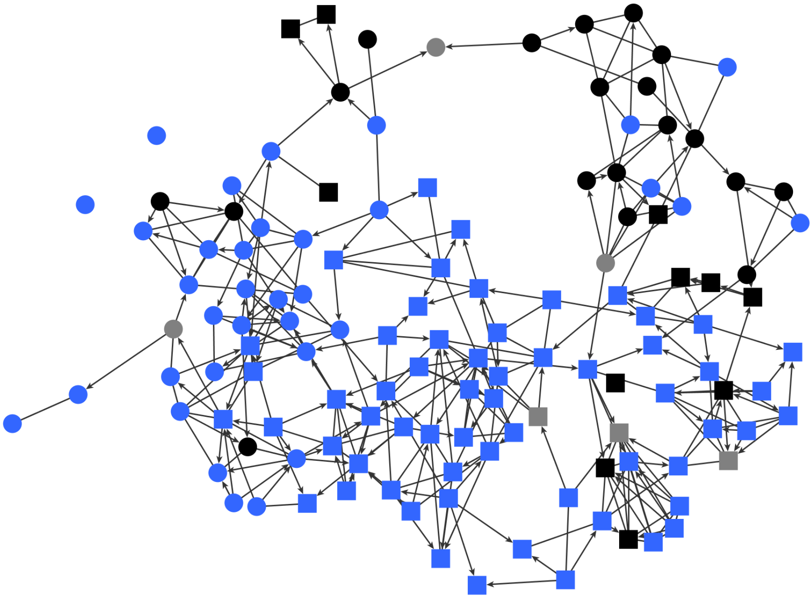
Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



Motivation

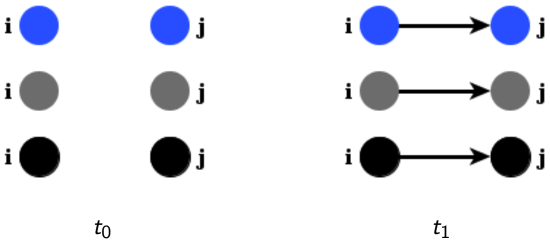
1. Social network dynamics can depend on actors'characteristics.

Selection process: relationship *partners* are selected according to their characteristics

Example

Homophily: the formation of relations based on the similarity of two actors

E.g. smoking behavior



Motivation

2. Changeable actors'characteristics can depend on the social network

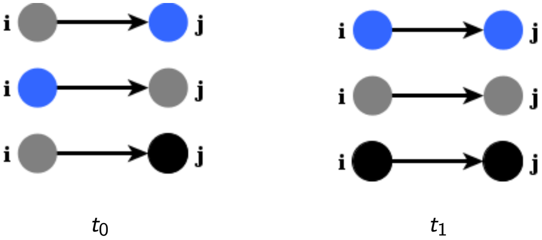
E.g.: opinions, attitudes, intentions, etc. - we use the word behavior for all of these!

Influence process: actors adjust their characteristics according to the characteristics of other actors to whom they are tied

Example

Assimilation/contagion: connected actors become increasingly similar over time

E.g. smoking behavior



Competing explanatory stories

Homophily and assimilation give rise to the same outcome (similarity of connected individuals)

$$\Downarrow$$

study of influence requires the consideration of selection and vice versa.

Fundamental question: is this similarity caused mainly by influence or mainly by selection?

Extending the SAOM for the co-evolution of networks and behaviors

Competing explanatory stories

Example

Similarity in smoking:

Selection: “a smoker may tend to have smoking friends because, once somebody is a smoker, he or she is likely to meet other smokers in smoking areas and thus has more opportunities to form friendship ties with them”

Influence: “the friendship with a smoker may have made an actor smoking in the first place”

Longitudinal network-behavior panel data

1. a network x represented by its adjacency matrix

2. a series of actors' attributes:
   - H constant covariates $V_1, \cdots, V_H$
   - L behavior covariates $Z_1(t), \cdots, Z_L(t)$
     Behavior variables are ordinal categorical variables.

Longitudinal network-behavior panel data: networks and behaviors observed at $M \geq 2$ time points $t_1, \cdots, t_M$

$$(x, z)(t_0),\ (x, z)(t_1), \cdots,\ (x, z)(t_M)$$

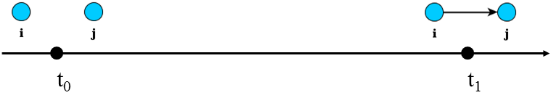
and the constant covariates $V_1, \cdots, V_H$.

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*

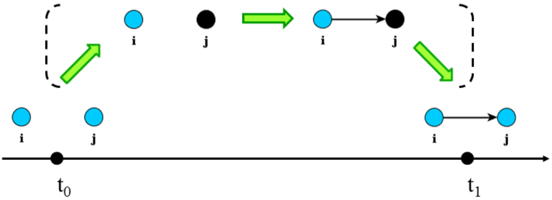


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

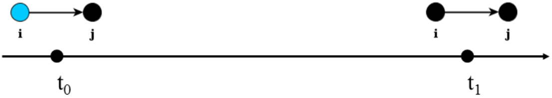
     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*



Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors
   $$|C| = 2^{n(n-1)} \times B^n$$
   where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network and behavior
   - *Continuous-time:*

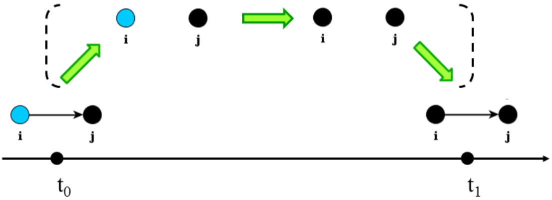


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space* $\mathcal{C}$*:* all the possible configurations arising from the combination of network and behaviors

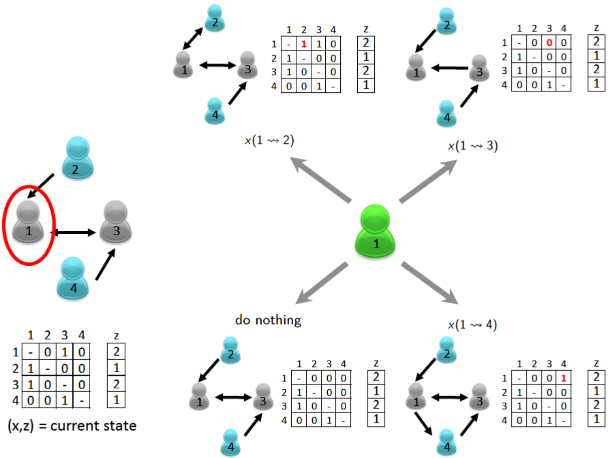
     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network and behavior
   - *Continuous-time:*



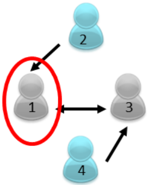
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.

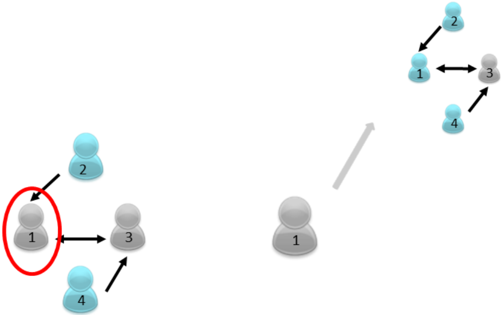


	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

(x,z) = current state

Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



	1	2	3	4	z
1	-	0	1	0	1
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

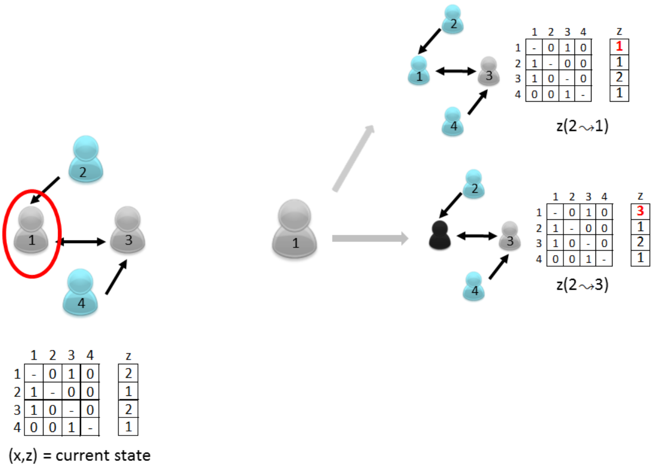
z(2⇝1)

	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

(x,z) = current state

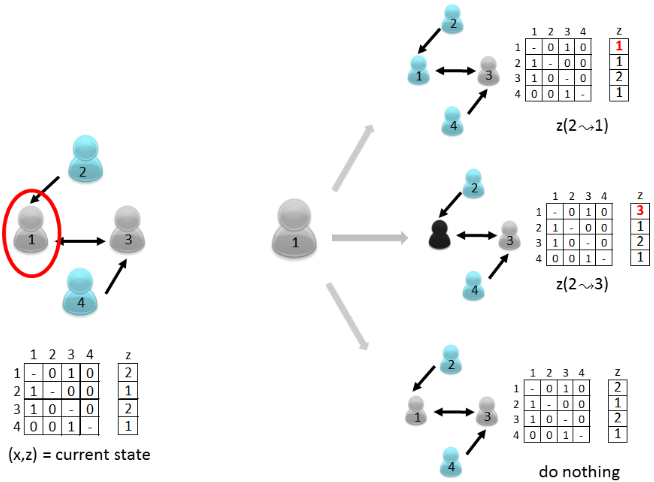
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



Assumptions

3. *Absence of co-occurrence.*
   At each instant t, only one actor has the opportunity to change (one of his outgoing ties or his behavior)

4. *Actor-oriented perspective.*
   Actors control their outgoing ties as well as their own behavior.

   - the actor decide to change one of his outgoing ties or his behavior trying to maximize *a utility function*
   - two distinct objective functions: one for the network and one for the behavior change
   - actors have complete knowledge about the network and the behaviors of all the the other actors
   - the maximization is based on immediate returns (myopic actors)

Model definition

The co-evolution process is decomposed into a series of micro-steps:

- **network micro-step:** the opportunity of changing one network tie and the corresponding tie changed
- **behavior micro-step:** the opportunity of changing a behavior and the corresponding unit changed in behavior

$$\Downarrow$$

every micro-step requires the identification of a focal actor who gets the opportunity to make a change and the identification of the change outcome

	Occurrence	*Preference*
Network changes	Network rate function	Network objective function
Behavioral changes	Behavioral rate function	Behavioral objective function

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

Practically always, one type of decision will be made more frequently than the other

Example

In the joint study of friendship and smoking behavior at high school, we would expect more frequent changes in the network than in behavior

The rate functions

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda_i^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda_i^{beh})$$

Waiting time for a new micro-step
$T_i^{net \vee beh}$ = the waiting time until i gets the opportunity to make any change

$$T_i^{net \vee beh} \sim Exp(\lambda_{tot})$$

where

$$\lambda_{tot} = \sum_i (\lambda_i^{net} + \lambda_i^{beh})$$

The rate functions (simplest specification)

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda^{beh})$$

Waiting time for a new micro-step
$T_i^{net \vee beh}$ = the waiting time until i gets the opportunity to make any change

$$T_i^{net \vee beh} \sim Exp(\lambda_{tot})$$

where

$$\lambda_{tot} = n(\lambda^{net} + \lambda^{beh})$$

The rate functions (simplest specification)

Probabilities for an actor to make a micro-step

$$P(i\ \ can\ \ make\ \ a\ \ network\ \ micro-step) = \frac{\lambda^{net}}{\lambda_{tot}}$$

$$P(i\ \ can\ \ make\ \ a\ \ behavioral\ \ micro-step) = \frac{\lambda^{beh}}{\lambda_{tot}}$$

Probabilities for a micro-step

$$P(network\ \ micro-step) = \frac{n\lambda^{net}}{\lambda_{tot}} = \frac{\lambda^{net}}{\lambda^{net} + \lambda^{beh}}$$

$$P(behavioral\ \ micro-step) = \frac{n\lambda^{beh}}{\lambda_{tot}} = \frac{\lambda^{beh}}{\lambda^{net} + \lambda^{beh}}$$

The objective functions

Why must we specify two different objective functions?

- The network objective function represents how likely it is for i to change one of his outgoing ties
- The behavioral objective function represents how likely it is for the actor i the current level of his behavior

Network utility function

$$u_i^{net}(\beta, x(i \rightsquigarrow j), z, v) = f_i^{net}(\beta, x(i \rightsquigarrow j), z, v) + \mathcal{E}_{ij}$$
$$= \sum_{k=1}^{K} \beta_k s_{ik}^{net}(x, z, v) + \mathcal{E}_{ij}$$

The objective functions

Behavioral utility function

$$\begin{aligned} u_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v) &= f_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v) + \mathcal{E}_{ll'} \\ &= \sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l'), v) + \mathcal{E}_{ll'} \end{aligned}$$

where

- $s_{iw}^{beh}(x, z(l \rightsquigarrow l'), v)$ are effects
- γ_w are statistical parameters
- $\mathcal{E}_{ll'}$ is a random term (Gumbel distributed)

The probability that an actor i changes his own behavior by one unit is:

$$p_{ll'}(i) = \frac{exp\left(f_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v)\right)}{\sum_{l'' \in \{l+1, l-1, l\}} exp\left(f_i^{beh}(\gamma, z(l \rightsquigarrow l''), x, v)\right)}$$

$p_{ll}(i)$ is the probability that i does not change his behavior.

N.b. In the following we will write z' instead of $z(l \rightsquigarrow l')$

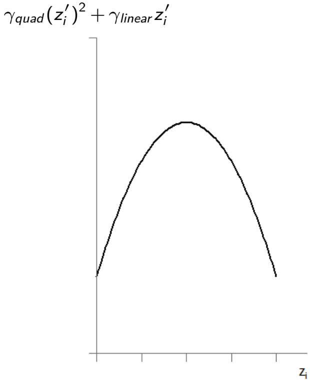
The objective functions

The specification of the behavioral objective function

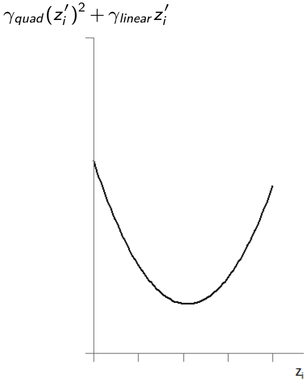
- *Basic shape effects*

$$s^{beh}_{i\_linear}(x, z', v) = z'_i \qquad s^{beh}_{i\_quadratic}(x, z', v) = (z'_i)^2$$

The basic shape effects must be always included in the model specification



$\gamma_{quad} < 0$



$\gamma_{quad} > 0$

The objective functions

The specification of the behavioral objective function

- *Classical influence effects*

1. The *average similarity effect*

$$s^{beh}_{i\_avsim}(x, z', v) = \frac{1}{x_{i+}} \sum_{j=1}^{n} x_{ij}(sim_{z'}(ij) - \widehat{sim}_z)$$

where

$$sim_{z'}(ij) = 1 - \frac{\left| z'_i - z'_j \right|}{R_z}$$

R_z is the range of the behavior z and $\widehat{sim}_z$ is the mean similarity value

N.b.: $z'_j = z_j$

The objective functions

The specification of the behavioral objective function

- *Classical influence effects*

1. The *average similarity effect*

$$s^{beh}_{i\_avsim}(x, z', v) = \frac{1}{\left(\sum_{j=1}^{n} x_{ij}\right)} \sum_{j=1}^{n} x_{ij}\left(sim_{z'}(ij) - \widehat{sim}_z\right)$$

where

$$sim_{z'}(ij) = 1 - \frac{\left|z'_i - z'_j\right|}{R_z}$$

R_z is the range of the behavior z and $\widehat{sim}_z$ is the mean similarity value

2. The *total similarity effect*

$$s^{beh}_{i\_totsim}(x, z', v) = \sum_{j=1}^{n} x_{ij}\left(sim_{z'}(ij) - \widehat{sim}_z\right)$$

N.b.: $z'_j = z_j$

The objective functions

The specification of the behavioral objective function

- *Position-dependent influence effects*

  Network position could also have an effect on the behavior of dynamics

  1. *outdegree effect*

$$s^{beh}_{i\_out}(x, z', v) = z'_i \sum_{j=1}^{n} x_{ij}$$

  2. *indegree effect*

$$s^{beh}_{i\_ind}(x, z', v) = z'_i \sum_{j=1}^{n} x_{ji}$$

- *Effects of other actor variables.*
  For each actor's attribute a main effect on the behavior can be included in the model

Simulating the co-evolution of networks and behavior

Aim: given $(x,z)(t_0)$ and fixed parameter values, provide $(x,z)^{sim}(t_1)$ according to the process behind the SAOM

$$\Downarrow$$

reproduce a possible series of network and behavior micro-steps between t_0 and t_1

Input

$n =$ number of actors

$\lambda^{net} =$ network rate parameter (given)

$\lambda^{beh} =$ behavior rate parameter (given)

$\beta = (\beta_1, \ldots, \beta_K) =$ objective function parameters (given)

$\gamma = (\gamma_1, \ldots, \gamma_W) =$ objective function parameters (given)

$(x,z)(t_0) =$ network and behavior at time t_0 (given)

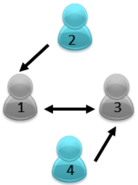
Output

$(x,z)^{sim}(t_1) =$ network and behavior at time t_1

Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition*=*TRUE* **do**
  $dt^{net} \sim Exp(n\lambda^{net})$
  $dt^{beh} \sim Exp(n\lambda^{beh})$
  if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
    $i \sim Uniform(1, \dots, n)$,
    $j \sim Multinomial(p_{i1}, \dots, p_{in})$
    if $i \neq j$ **then**
      $x \leftarrow x(i \rightsquigarrow j)$
      $t \leftarrow t + dt^{net}$
  else
    $i \sim Uniform(1, \dots, n)$,
    $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
    if $l \neq l'$ **then**
      $z \leftarrow z(l \rightsquigarrow l')$
      $t \leftarrow t + dt^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$



	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

$(x, z)(t_0)$

$n = 4$

$\lambda^{net} = 1.5$

$\lambda^{net} = 1$

$\beta = (\beta_{out}, \beta_{rec}, \beta_{trans})$
$=(-1, 0.5, -0.25)$

$\gamma = (\gamma_{linear}, \gamma_{quadratic}, \gamma_{avsim})$
$=(-2, 1, 0.25)$

Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$

```
t ← 0; x ← x(t0); z ← z(t0)
while condition=TRUE do
    dt^net ~ Exp(nλ^net)
    dt^beh ~ Exp(nλ^beh)
    if min{dt^net, dt^beh} = dt^net then
        i ~ Uniform(1, ..., n),
        j ~ Multinomial(p_i1, ..., p_in)
        if i ≠ j then
            x ← x(i ⇝ j)
        t ← t + dt^net
    else
        i ~ Uniform(1, ..., n),
        l' ~ Multinomial(p_l(l−1), p_ll', p_l(l+1))
        if l ≠ l' then
            z ← z(l ⇝ l')
            t ← t + dt^beh
x^sim(t1) ← x
z^sim(t1) ← z
return x^sim(t1), z^sim(t1)
```

Generating the waiting time:

- dt^{net} for a tie change
- dt^{beh} for a behavior change

Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition*=*TRUE* **do**
    $dt^{net} \sim Exp(n\lambda^{net})$
    $dt^{beh} \sim Exp(n\lambda^{beh})$
    if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
        $i \sim Uniform(1, \ldots, n)$
        $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
        if $i \neq j$ **then**
            $x \leftarrow x(i \rightsquigarrow j)$
        $t \leftarrow t + dt^{net}$
    else
        $i \sim Uniform(1, \ldots, n)$
        $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
        if $l \neq l'$ **then**
            $z \leftarrow z(l \rightsquigarrow l')$
            $t \leftarrow t + dt^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Which micro-step is going to happen?

If

$$dt^{net} < dt^{beh}$$

then a network micro-step takes place

The following steps are the same of those in Algorithm 1

Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition=TRUE* **do**
  $dt^{net} \sim Exp(n\lambda^{net})$
  $dt^{beh} \sim Exp(n\lambda^{beh})$
  if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
      $x \leftarrow x(i \rightsquigarrow j)$
    $t \leftarrow t + dt^{net}$
  else
    $i \sim Uniform(1, \ldots, n)$
    $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
    if $l \neq l'$ **then**
      $z \leftarrow z(l \rightsquigarrow l')$
    $t \leftarrow t + dt^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Which micro-step is going to happen?

If

$$dt^{beh} < dt^{net}$$

then a behavior micro-step takes place

Simulating the co-evolution of networks and behavior

Algorithm 2:

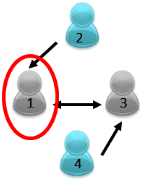
Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$

$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition=TRUE* **do**
  $dt^{net} \sim Exp(n\lambda^{net})$
  $dt^{beh} \sim Exp(n\lambda^{beh})$
  if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
    $i \sim Uniform(1, \dots, n)$
    $j \sim Multinomial(p_{i1}, \dots, p_{in})$
    if $i \neq j$ **then**
      $x \leftarrow x(i \rightsquigarrow j)$
    $t \leftarrow t + dt^{net}$
  else
    $i \sim Uniform(1, \dots, n)$
    $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
    if $l \neq l'$ **then**
      $z \leftarrow z(l \rightsquigarrow l')$
    $t \leftarrow t + dt^{beh}$

$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Select the actor i who has the opportunity to change his behavior

e.g. i=*1*



	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

$(x, z)(t_0)$

Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition=TRUE* **do**
    $dt^{net} \sim Exp(n\lambda^{net})$
    $dt^{beh} \sim Exp(n\lambda^{beh})$
    if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
        $i \sim Uniform(1, \ldots, n)$
        $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
        if $i \neq j$ **then**
            $x \leftarrow x(i \rightsquigarrow j)$
        $t \leftarrow t + dt^{net}$
    else
        $i \sim Uniform(1, \ldots, n)$
        $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
        if $l \neq l'$ **then**
            $z \leftarrow z(l \rightsquigarrow l')$
        $t \leftarrow t + dt^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Select the level l' towards i is going to adjust his behavior

$l \rightarrow l'$	f_i^{beh}	$p_{ll'}$
$2 \rightarrow 1$	0.017	0.017
$2 \rightarrow 2$	0.052	0.052
$3 \rightarrow 3$	0.930	0.931

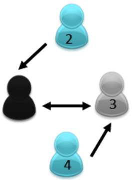
Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition=TRUE* **do**
  $dt^{net} \sim Exp(n\lambda^{net})$
  $dt^{beh} \sim Exp(n\lambda^{beh})$
  if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
      $x \leftarrow x(i \rightsquigarrow j)$
    $t \leftarrow t + dt^{net}$
  else
    $i \sim Uniform(1, \ldots, n)$
    $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
    if $l \neq l'$ **then**
      $z \leftarrow z(l \rightsquigarrow l')$
    $t \leftarrow t + dt^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Select the level l' towards i is going to adjust his behavior

e.g. $l'=3$



	1	2	3	4	z
1	-	0	1	0	3
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

$(x, z(l \rightarrow l'))$

Simulating the co-evolution of networks and behavior

Algorithm 2:

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition*=*TRUE* **do**
  $dt^{net} \sim Exp(n\lambda^{net})$
  $dt^{beh} \sim Exp(n\lambda^{beh})$
  if $\min\{dt^{net}, dt^{beh}\} = dt^{net}$ **then**
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
      $x \leftarrow x(i \rightsquigarrow j)$
    $t \leftarrow t + dt^{net}$
  else
    $i \sim Uniform(1, \ldots, n)$
    $l' \sim Multinomial(p_{l(l-1)}, p_{ll'}, p_{l(l+1)})$
    if $l \neq l'$ **then**
      $z \leftarrow z(l \rightsquigarrow l')$
    $t \leftarrow t + dt^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Simulating the co-evolution of networks and behavior

1. *Unconditional* simulation:

   simulation carries on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:
   - simulation runs on until

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}^{obs}(t_1) - X_{ij}(t_0) \right| = \sum_{i,j=1}^{n} \left| X_{ij}^{sim}(t_1) - X_{ij}(t_0) \right|$$

   - simulation runs on until

$$\sum_{i=1}^{n} \left| z_i^{obs}(t_1) - z_i(t_0) \right| = \sum_{i=1}^{n} \left| z_i^{sim}(t_1) - z_i(t_0) \right|$$

Example

Example data: excerpt from the "Teenage Friends and Lifestyle Study" data set

We will use the SAOM for the co-evolution of networks and behaviors to distinguish influence from selection.

1. Do pupils select friends based on similar smoking behavior?
2. Are pupils influenced by friends to adjust to their smoking behavior?

Dependent variables: friendship networks and smoking behavior

Covariate: gender

Precondition of the analysis

To find out whether it makes sense to analyze the data with a co-evolution model one should check whether:

1. the data are sufficiently informative

$$J = \frac{N_{11}}{N_{11} + N_{01} + N_{10}} > 0.3 \qquad \textit{Jaccard index}$$

Tie changes between subsequent observations:

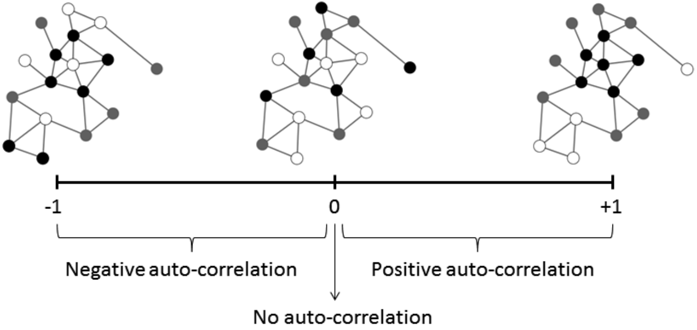
periods	0 => 0	0 => 1	1 => 0	1 => 1	Distance	Jaccard	Missing
1 ==> 2	15827	237	240	208	477	0.304	0 (0%)
2 ==> 3	15839	228	209	236	437	0.351	0 (0%)

Precondition of the analysis

2. there is interdependence between network and behavioral variables

$$I = \frac{n \sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_{i}(z_i - \overline{z})^2\right)} \qquad \textit{Moran index}$$

where $\overline{z}$ is the mean of z over all the periods



Precondition of the analysis

The computation of the index *I* for the data leads to

0.244	0.258	0.341

Conclusion:
there is considerable dependence between networks and behaviors and it is reasonable to apply the SAOM

```
moran1 <- nacf(net1,tobacco[,1],lag.max=1,neighborhood.type = "out",
        type="moran",mode="digraph")
moran2 <- nacf(net2,tobacco[,2],,lag.max=1,neighborhood.type = "out",
        type="moran",mode="digraph")
moran3 <- nacf(net3,tobacco[,3],,lag.max=1,neighborhood.type = "out",
        type="moran",mode="digraph")

moranInd <- c(moran1[2],moran2[2],moran3[2])
```

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Network rate parameters:

- about 9 opportunities for a network change in the first period
- about 7 opportunities for a network change in the second period

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Network objective function parameters:

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Behavioral rate parameters:

- about 4 opportunities for a behavioral change in the first period
- about 4 opportunities for a behavioral change in the second period

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Behavioral objective function parameters:

attractiveness of different behavioral levels based on the current structure of the network and the behavior of the others

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{z} = 1.377$ is the mean of the covariate

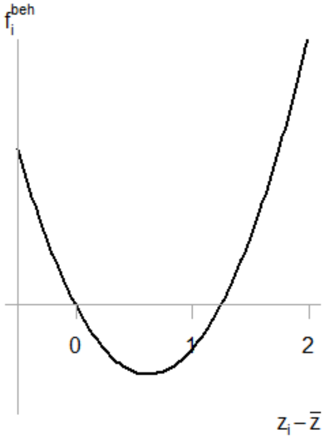
$$z_i - \overline{z} = \begin{cases} -0.377 & \textit{for no smokers} \\ 0.623 & \textit{for occasional smokers} \\ 1.623 & \textit{for regular smokers} \end{cases}$$

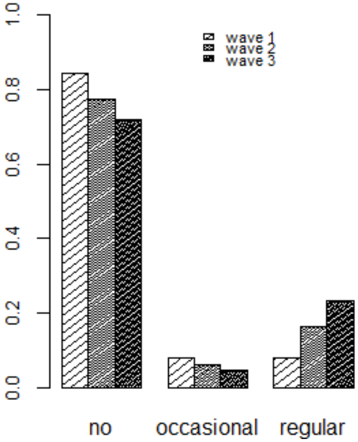
- The contribution to the behavioral objective function is

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 =$$

$$= -3.5464(z_i - \overline{z}) + 2.8464(z_i - \overline{z})^2$$

Parameter interpretation: a baseline model





U-shaped changes in the behavior are drawn to the extreme of the range

A more complex model

The baseline model does not provide any information about selection and influence processes:

- the network dynamics are explained by the preference towards creating and reciprocating ties
- the behavior dynamics are described only by the distribution of the behavior in the population

If we want to distinguish selection from influence we should include in the objective functions specification:

- the effects that capture the dependence of social network dynamics on actor's characteristic
- the effects that capture the dependence of behavior dynamics on social network

A more complex model

Effects for the dependence of network dynamics on actor's characteristic

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

This effect must be controlled for the sender and receiver effects of the covariate.

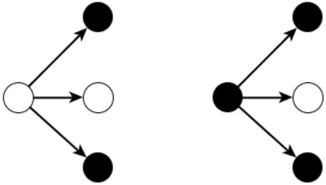
- Covariate ego effect

- Covariate alter effect

A more complex model

Effects for the dependence of behavior dynamics on network

- pupils tend to adjust their smoking behavior according to the behaviors of their friends → average similarity effect



This effect must be controlled for the indegree and the outdegree effects

- Indegree effect

- Outdegree effect

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

Network objective function parameters:

tendency towards reciprocity, transitivity and homophily with respect to gender

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

Network objective function parameters:

pupils selected others with similar smoking behavior as friends

$\rightarrow$ evidence for selection process

A more complex model

The contribution to the network objective function is given by:

$$\beta_{ego}(z_i - \overline{z}) + \beta_{alter}(z_j - \overline{z}) + \beta_{same}\left(1 - \frac{|z_i - z_j|}{R_z} - sim_z\right) =$$

$$= 0.0665(z_i - 1.377) + 0.1121(z_j - 1.377) + 0.5114(1 - \frac{|z_i - z_j|}{R_z} - 0.7415)$$

z_i/z_j	no	occasional	regular
no	0.0648	-0.0787	-0.2223
occasional	-0.1243	0.2435	0.0999
regular	-0.3135	0.0543	0.4221

- preference for similar alters
- this tendency is strongest for high values on smoking behavior

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shape	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters:

U-shaped distribution of the smoking behavior

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shape	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters:

indegree and outdegree effects are not significant

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shape	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters:

pupils are influenced by the smoking behavior of the others

$\rightarrow$ evidence for influence process

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573(z_i - \overline{z}) + 2.8406(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

where $sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_Z} = 1$

Example

a) i adjusts his behavior to "no-smoker" when all of his friends are no-smokers

$$sim_z(ij) = 1 - \frac{|1-1|}{2} = 1$$

$$-3.3573(1 - 1.377) + 2.8406(1 - 1.377)^2 + 3.4361(1 - 0.7415) = 2.56$$

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573(z_i - \overline{z}) + 2.8406(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

where $sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_Z} = 1$

Example

b) i adjusts his behavior to "no-smoker" when all of his friends are occasional smokers

$$sim_z(ij) = 1 - \frac{|1-2|}{2} = 0.5$$

$$-3.3573(1 - 1.377) + 2.8406(1 - 1.377)^2 + 3.4361(0.5 - 0.7415) = 0.84$$

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573(z_i - \overline{z}) + 2.8406(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

where $sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_Z} = 1$

Example

b) i adjusts his behavior to "no-smoker" when all of his friends are regular smokers

$$sim_z(ij) = 1 - \frac{|1-3|}{2} = 0$$

$$-3.3573(1 - 1.377) + 2.8406(1 - 1.377)^2 + 3.4361(0 - 0.7415) = -0.88$$

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573_{linear}(z_i - \overline{z}) + 2.8406_{quadratic}(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

z_j / z_i	no	occasional	regular
no	2.56	-1.82	-0.51
occasional	0.84	-0.10	1.20
regular	-0.88	-1.82	2.92

- the focal actor prefers to have the same behavior as all these friends (except for the occasional smokers)

- friends do not smoke at all: the preference toward imitating their behavior is less strong

The parameter estimation (MoM)

Aim: given the longitudinal data

$$(x,z)(t_0),\dots,(x,z)(t_M) \qquad V_1,\dots,V_H$$

estimate the parameters for the co-evolution model

- M rate parameters for the network rate function

$$\lambda_1^{net},\ \dots,\ \lambda_M^{net}$$

- M rate parameters for the behavior rate function

$$\lambda_1^{beh},\ \dots,\ \lambda_M^{beh}$$

- K and W parameters for the network objective function and for the behavior objective function, respectively

$$f_i^{net}(\beta,x',z,v)=\sum_{k=1}^{K}\beta_k s_{ik}^{net}(x',z,v) \qquad f_i^{beh}(\gamma,x',z,v)=\sum_{w=1}^{W}\gamma_w s_{iw}^{beh}(x,z',v)$$

The parameter estimation (MoM)

We can estimate the $2M + K + W$-dimensional parameter θ using the MoM

In practice:

1. find $2M + K + W$ statistics
2. set the theoretical expected value of each statistic equal to its sample counterpart
3. solve the resulting system of equations

   $$E_\theta[S - s] = 0$$

   with respect to θ

The parameter estimation (MoM)

Statistics:

- Network rate parameters for the period m

$$s^{net}_{\lambda_m}(X(t_m), X(t_{m-1}) | X(t_{m-1})) = \sum_{i,j=1}^{n} \left| X_{ij}(t_m) - X_{ij}(t_{m-1}) \right|$$

- Behavior rate parameters for the period m

$$s^{beh}_{\lambda_m}(Z(t_m), Z(t_{m-1}) | Z(t_{m-1})) = \sum_{i=1}^{n} |Z_i(t_m) - Z_i(t_{m-1})|$$

$m = 1, \ldots, M$

The parameter estimation (MoM)

Statistics:

- Network objective function effects

$$\sum_{m=1}^{M} s_{mk}^{net}\left((X,Z,V)(t_m)|(X,Z,V)(t_{m-1})\right) = \sum_{m=1}^{M} s_{mk}^{net}\left((X,Z,V)(t_m),(X,Z,V)(t_{m-1})\right)$$

- Behavior objective function effects

$$\sum_{m=1}^{M} s_{mw}^{beh}\left((X,Z,V)(t_m)|(X,Z,V)(t_{m-1})\right) = \sum_{m=1}^{M} s_{mw}^{beh}\left((X,Z,V)(t_m),(X,Z,V)(t_{m-1})\right)$$

The parameter estimation (MoM)

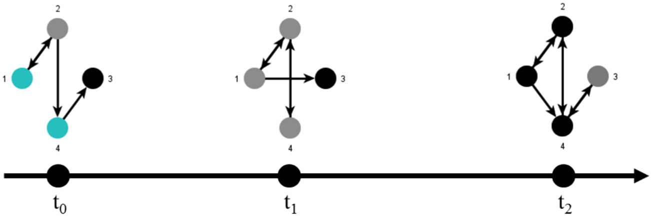
Consequently the MoM estimator for θ is provided by the solution of:

$$\begin{cases} E_\theta\left[s^{net}_{\lambda_m}(X(t_M), X(t_{m-1})|X(t_{m-1}))\right] = s^{net}_{\lambda_m}(x(t_m), x(t_{m-1})) & m = 1, \dots, M \\ E_\theta\left[s^{beh}_{\lambda_m}(Z(t_m), Z(t_{m-1})|Z(t_{m-1}))\right] = s^{beh}_{\lambda_m}(z(t_m), z(t_{m-1})) & m = 1, \dots, M \\ E_\theta\left[\sum_{m=1}^{M} s^{net}_{mk}((X, Z, V)(t_m))\right] = \sum_{m=1}^{M} s^{net}_{mk}((x, z, v)(t_m)) & k = 1, \dots, K \\ E_\theta\left[\sum_{m=1}^{M} s^{beh}_{mw}((X, Z, V)(t_m)))\right] = \sum_{m=1}^{M} s^{beh}_{mw}((x, z, v)(t_m)) & w = 1, \dots, W \end{cases}$$

The parameter estimation (MoM)

Example

Let us assume to have observed a network at $M = 3$ time points



t_0:

	1	2	3	4	beh
1	-	1	0	0	0
2	1	-	0	1	1
3	0	0	-	0	2
4	0	0	1	-	0

t_1:

	1	2	3	4	beh
1	-	1	1	0	1
2	1	-	0	1	1
3	0	0	-	0	2
4	0	1	0	-	1

t_2:

	1	2	3	4	beh
1	-	1	0	1	2
2	1	-	0	1	2
3	0	0	-	1	1
4	0	1	1	-	2

We want to model the network evolution according to the outdegree, the reciprocity, the linear shape and the quadratic shape effects

$$\theta = (\lambda_1^{net}, \lambda_2^{net}, \lambda_1^{beh}, \lambda_2^{beh}, \beta_{out}, \beta_{rec}, \gamma_{linear}, \gamma_{quadratic})$$

The parameter estimation (MoM)

Example

Statistics for the network evolution:

$$s_{\lambda_1^{net}}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{4} \left|X_{ij}(t_1) - X_{ij}(t_0)\right|$$

$$s_{\lambda_2^{net}}(X(t_2), X(t_1)|X(t_1) = x(t_1)) = \sum_{i,j=1}^{4} \left|X_{ij}(t_2) - X_{ij}(t_1)\right|$$

$$\sum_{m=1}^{M-1} s_{out}\,(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{2} \sum_{i,j=1}^{4} X_{ij}(t_m)$$

$$\sum_{m=1}^{M-1} s_{rec}\,(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{2} \sum_{i,j=1}^{4} X_{ij}(t_m) X_{ji}(t_m)$$

The parameter estimation (MoM)

Example

Statistics for the behavior evolution:

$$s_{\lambda_1^{beh}}(Z(t_1), Z(t_0) | Z(t_0) = z(t_0)) = \sum_{i=1}^{4} |Z_i(t_1) - Z_i(t_0)|$$

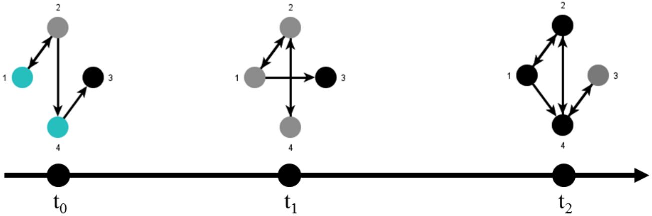
$$s_{\lambda_2^{beh}}(Z(t_2), Z(t_1) | Z(t_1) = z(t_1)) = \sum_{i=1}^{4} |Z_i(t_2) - Z_i(t_1)|$$

$$\sum_{m=1}^{M} s_{linear}(Z(t_m) | Z(t_{m-1}) = z(t_{m-1})) = \sum_{m=1}^{2} \sum_{i=1}^{4} z_i(t_m)$$

$$\sum_{m=1}^{M} s_{quadratic}(Z(t_m) | Z(t_{m-1}) = z(t_{m-1})) = \sum_{m=1}^{2} \sum_{i=1}^{4} z_i^2(t_m)$$

The parameter estimation (MoM)

Example



t_0:

	1	2	3	4	beh
1	-	1	0	0	0
2	1	-	0	1	1
3	0	0	-	0	2
4	0	0	1	-	0

t_1:

	1	2	3	4	beh
1	-	1	1	0	1
2	1	-	0	1	1
3	0	0	-	0	2
4	0	1	0	-	1

t_2:

	1	2	3	4	beh
1	-	1	0	1	2
2	1	-	0	1	2
3	0	0	-	1	1
4	0	1	1	-	2

$s_{\lambda_1^{net}} = 3$ $\quad$ $s_{\lambda_2^{net}} = 4$

$s_{\lambda_1^{beh}} = 2$ $\quad$ $s_{\lambda_2^{beh}} = 4$

$s_{out} = 5 + 7 = 12$ $\quad$ $s_{rec} = 4 + 6 = 10$

$s_{linear} = 5 + 7 = 12$ $\quad$ $s_{quadratic} = 7 + 13 = 20$

The parameter estimation (MoM)

Example

We look for the value of θ that satisfies the system:

$$\begin{cases} E_\theta\left[S_{\lambda_1^{net}}\right] = 3 \\ E_\theta\left[S_{\lambda_2^{net}}\right] = 4 \\ E_\theta\left[S_{\lambda_1^{beh}}\right] = 2 \\ E_\theta\left[S_{\lambda_2^{beh}}\right] = 4 \\ E_\theta[S_{out}] = 12 \\ E_\theta[S_{rec}] = 10 \\ E_\theta[S_{linear}] = 12 \\ E_\theta[S_{quadratic}] = 20 \end{cases}$$

The parameter estimation (MoM)

In a more compact notation, we look for the value of θ that satisfies the system:

$$E_\theta[S - s] = 0$$

but we know that we cannot solve it analytically.

The soultion is again provided by the Robbins-Monro algorithm.

Outline

Creating and deleting ties

Terminating a tie is not just the opposite of creating a tie

Example

- the loss in terminating a tie is greater than the reward in creating one
- transitivity plays an important role especially in creating ties

This is modeled by adding to the objective function one of the two components:

1. the *creation function*
2. the *endowment function*

The creation function

Models the gain in satisfaction incurred when a network tie is created:

$$c_i(\delta, x) = \sum_k \delta_k s_{ik}(x)$$

where

- δ_k are statistical parameters
- $s_{ik}(x)$ are the effects

The utility function for an actor i when he creates a new tie is provided by:

$$u_i(x) = f_i(\beta, x) + c_i(\delta, x) + \epsilon_i(t, x, j)$$

The creation function is zero for the dissolution of ties

The endowment function

Models the loss in satisfaction incurred when a network tie is deleted

$$e_i(\eta, x) = \sum_k \eta_k s_{ik}(x)$$

where

- η_k are statistical parameters
- $s_{ik}(x)$ are the effects

The utility function for an actor i when he deletes a tie is provided by:

$$u_i(x) = f_i(\beta, x) + e_i(\eta, x) + \epsilon_i(t, x, j)$$

The endowment function is zero for the creation of ties

Creating and deleting ties - Remarks

- creation and deletion functions must not be included when ties mainly are created or terminated

- it could also happen that increasing a behavior is not the same as decreasing a behavior. Thus, there are also:

  1. the *creation behavior function*

  2. the *endowment behavior function*

  but their usage is still under investigation

Creating and deleting ties

Example

Example data: excerpt from the “Teenage Friends and Lifestyle Study” data set

We estimate the SAOM for investing the evolution of friendship networks according to:

- outdegree
- reciprocity
- transitivity
- reciprocity for the endowment function

```
myeff < − includeEffects(myeff,transTrip)
myeff < − includeEffects(myeff,recip,type=”endow”)
myeff
mymodel < − sienaModelCreate(useStdInits = FALSE, projname = ’tfls’)
model1 < − siena07(mymodel, data = mydata, effects=myeff,useCluster=TRUE,
nbrNodes=2, initC=TRUE,clusterString=rep(”localhost”, 2))
```

Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	6.70	0.73	
Rate parameter period 2	5.81	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	2.23	0.58	3.85
transitive triplets	0.44	0.03	14.55

The utility function for an actor i when he deletes a tie is provided by:

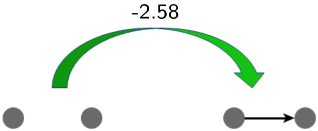
$$
\begin{aligned}
u_i(x) &= f_i(\beta, x) + e_i(\eta, x) + \epsilon_i(t, x, j) = \\
&= \beta_{out} s_{i\_out}(x) + \beta_{rec} s_{i\_rec}(x) + \beta_{trans} s_{i\_trans}(x) + \eta_{rec} s_{i\_rec}(x) \\
&= -2.58 s_{i\_out}(x) + 3.23 s_{i\_rec}(x) + 0.44 s_{i\_trans}(x) - 2.23 s_{i\_rec}(x)
\end{aligned}
$$

Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	2.23	0.58	3.85
transitive triplets	0.44	0.03	14.55

Ties formation/deletion

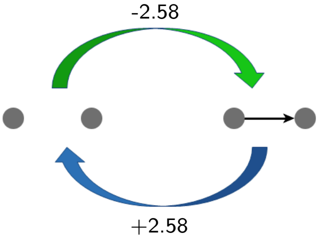


Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	2.23	0.58	3.85
transitive triplets	0.44	0.03	14.55

Ties formation/deletion

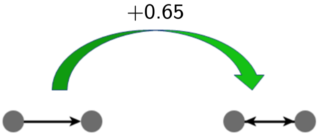


Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	2.23	0.58	3.85
transitive triplets	0.44	0.03	14.55

Reciprocation/ending reciprocation

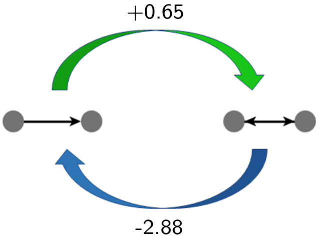


Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	2.23	0.58	3.85
transitive triplets	0.44	0.03	14.55

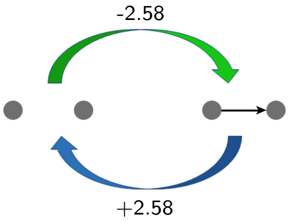
Reciprocation/ending reciprocation



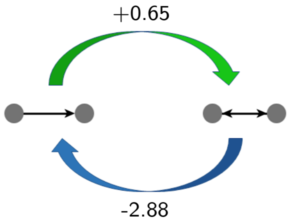
Creating and deleting ties

Example

Ties formation/deletion



Reciprocation/ending reciprocation



Conclusions:

1. formation of reciprocal ties is more rewarding than the formation of a non-reciprocal tie

2. dissolution of reciprocal ties is more costly than the dissolution of a non-reciprocal tie and the creation of a reciprocal tie

Non-directed relations

For directed relation we assumed that:

1. an actor gets the opportunity to make a change
2. he decided for the change that assures him the highest payoff

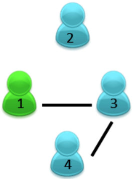
Are this assumptions still reliable when we consider undirected relations such as: collaboration, trade, strategic alliance?

- Yes, if one actor (*dictator*) can impose a decision about a tie to another
- No, if there is coordination or negotiation about a tie

Non-directed relations

1. Dictatorial choice: i chooses his action and imposes his decision to j

Actor 1 gets the opportunity to change



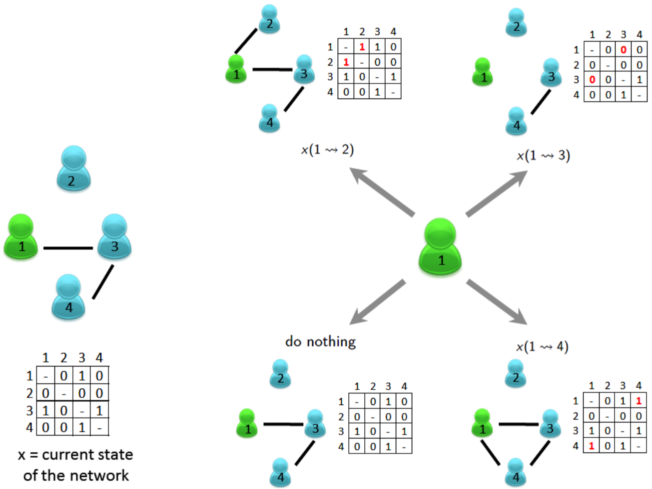
	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

Non-directed relations

1. Dictatorial choice: i chooses his action and imposes his decision to j

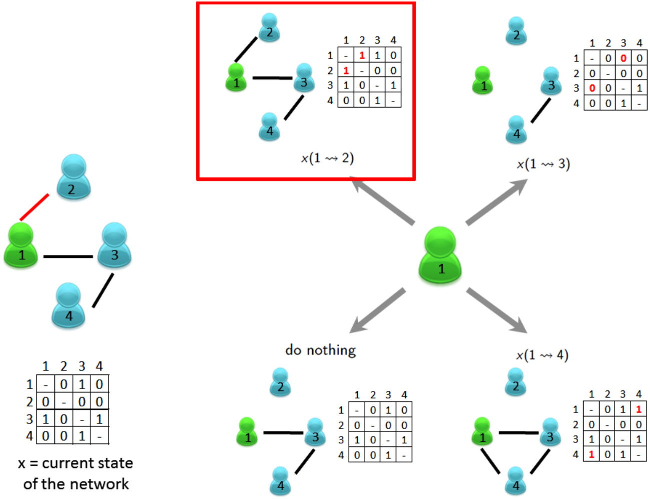
Actor 1 evaluates the alternatives and the corresponding objective functions



Non-directed relations

1. Dictatorial choice: *i* chooses his action and imposes his decision to *j*

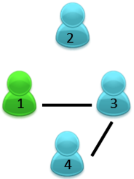
E.g. actor 1 imposes his choice to actor 2



Non-directed relations

2. Mutual agreement: both actors must agree

Actor 1 gets the opportunity to change



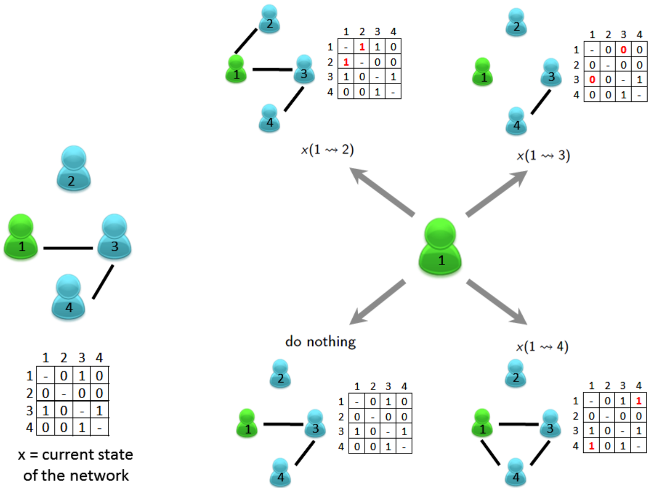
	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

Non-directed relations

2. Mutual agreement: both actors must agree

Actor 1 evaluates the alternatives and the corresponding objective functions



	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

$x(1 \rightsquigarrow 2)$:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	1
4	0	0	1	-

$x(1 \rightsquigarrow 3)$:

	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	1
4	0	0	1	-

do nothing:

	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

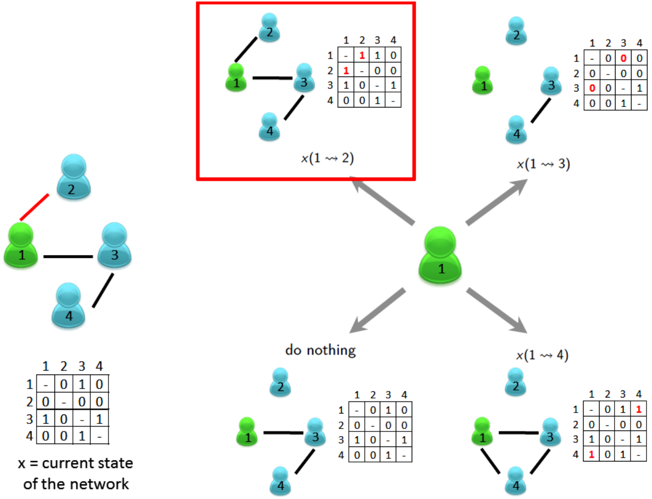
$x(1 \rightsquigarrow 4)$:

	1	2	3	4
1	-	0	1	1
2	0	-	0	0
3	1	0	-	1
4	1	0	1	-

Non-directed relations

2. Mutual agreement: both actors must agree

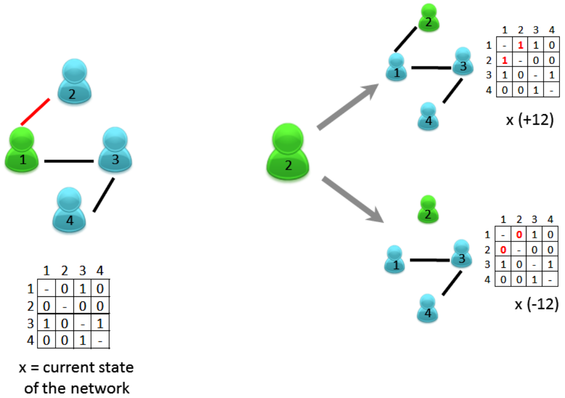
Actor 1 suggests to modify the tie towards actor 2



Non-directed relations

2. Mutual agreement: both actors must agree

Actor 2 evaluates the proposal of actor 1



	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	1
4	0	0	1	-

	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

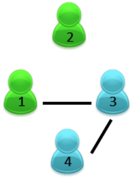
and accepts it with probability

$$P(2\ accepts\ tie\ proposal) = \frac{exp(f_2(x^{+12}))}{exp(f_2(x^{+12})) + exp(f_2(x^{-12}))}$$

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

1. Dictatorial choice: one actor can impose the decision (e.g. actor 1)



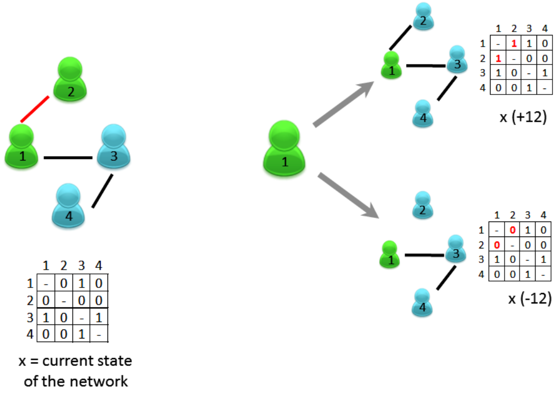
	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

1. Dictatorial choice: one actor can impose the decision (e.g. actor 1)



Current state x:

	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x (+12):

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	1
4	0	0	1	-

x (-12):

	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

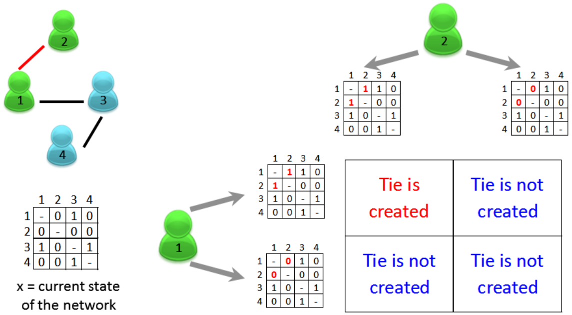
Actor 1 chooses his action with probability

$$P(1 \ imposes \ a \ tie \ on \ 2) = \frac{exp(f_1(x^{+12}))}{exp(f_1(x^{+12})) + exp(f_1(x^{-12}))}$$

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

2. Mutual agreement: both actors propose a tie



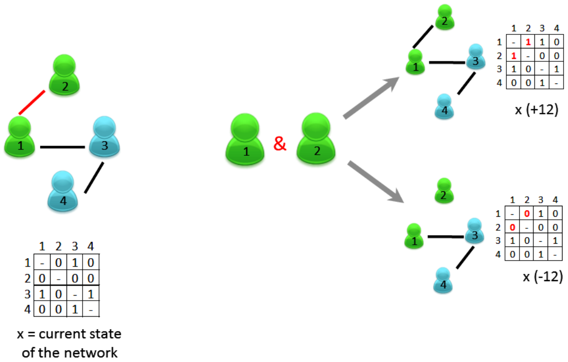
Actor 1 and 2 created a tie with probability

$$P(+12) = \frac{exp(f_1(x^{+12}))}{exp(f_1(x^{+12})) + exp(f_1(x^{-12}))} \frac{exp(f_2(x^{+12})}{exp(f_2(x^{+12})) + exp(f_2(x^{-12}))}$$

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

3. Compensatory: the decision is made on the combined interest



Actor 1 and 2 choose their action with probability

$$P(+12) = \frac{exp(f_1(x^{+12}) + f_2(x^{+12}))}{exp(f_1(x^{+12}) + f_2(x^{+12})) + exp(f_1(x^{-12}) + f_2(x^{-12}))}$$

And others...

- Improving the estimation procedures (MLE)
- New estimation procedures (bayesian estimation)
- Goodness of fit of the model
- Model selection
- Time-heterogeneity tests
- Missing data
- Analysis of multiple relations
- ...

Outline

Introduction

The Stochastic actor-oriented model

Extending the model: analyzing the co-evolution of networks and behavior

Something more on the SAOM

ERGMs and SAOMs

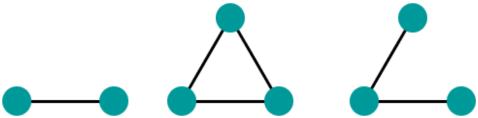
Recap: ERGMs

ERGMs are models for cross-sectional data:

they return the probability of an observed graph (network) $G \in \mathcal{G}$ as a function of statistics $g_i(G)$ and statistical parameters θ_i

$$P(G) = \frac{exp\left(\sum_{i=1}^{k} \theta_i \cdot g_i(G)\right)}{\kappa(\theta)}$$

Examples of statistics $g_i(G)$ are:



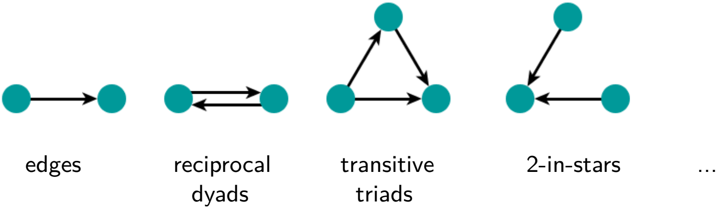
edges triangles 2-stars ...

Recap: ERGMs

ERGMs are also defined for directed graphs:

the mathematical formulation is the same but the effects take into account the direction of ties

Examples of statistics $g_i(G)$ for a directed network are:



Recap: SAOMs

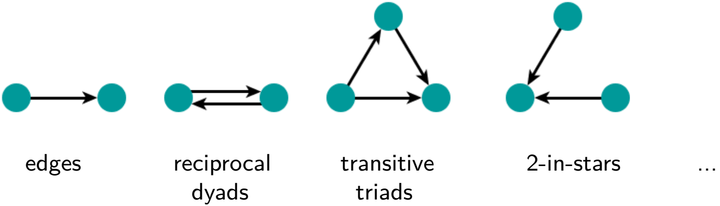
SAOMs are models for longitudinal data:

SAOMs try to explain the evolution of the network over time, assuming that network changes happen according to a continuous-time Markov chain modeled by:

- the rate function λ
- the objective function

$$f_i(\beta, x(i \rightsquigarrow j), v_i, v_j) = \sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow j))$$

where the statistics $s_{ik}(x(i \rightsquigarrow j))$ are:



SAOMs and ERGMs

Although ERGMs and SAOMs have different aims and require different data, the same statistics are used as explanatory variables in both models.

This might suggest the existence of a "statistical" relation between ERGMs and SAOMs

We are going to prove that:

1. ERGMs are the limiting distribution of the process described by a certain specification of SAOMs
2. ERGMs are the limiting distribution of the process described by a tie-based version of SAOMs

Background: intensity matrix

Definition

Let $\{X(t), t \in \mathcal{T}\}$ be a continuous-time Markov chain whose transition probabilities are defined by:

$$P(X(t_j) = \widetilde{x} | X(t) = x(t), \forall t \leq t_i) = P(X(t_j) = x | X(t_i) = x)$$

for each pair $(x, \widetilde{x})$.

There exists a function $q : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ such that

$$\begin{cases} q(x,\widetilde{x}) = \lim\limits_{dt \to 0} \frac{P(X(t+dt)=\widetilde{x}|X(t)=x)}{dt} \\ \\ q(x,x) = \lim\limits_{dt \to 0} \frac{P(X(t+dt)=\widetilde{x}|X(t)=x)-1}{dt} \end{cases}$$

The function q is called **intensity matrix** of the process.

The element $q(x, \widetilde{x})$ is referred to as the rate at which the process in state x tends to change into $\widetilde{x}$

Background (recall): limiting distribution

Definition

The **limiting distribution** P of a continuous-time Markov chain $\{X(t), t \in \mathcal{T}\}$ is defined as

$$P_{\widetilde{x}} = \lim_{t \to \infty} P(X(t_j) = \widetilde{x} | X(t_i) = x)$$

Therefore, the limiting distribution of $\{X(t), t \in \mathcal{T}\}$ is the distribution that describes the probability of jumping from x to $\widetilde{x}$ in the long run behavior of the process .

$P_{\widetilde{x}}$ is also the stationary distribution of the process

Background (recall): irreducible aperiodic Markov chain and limiting distribution

Definition

A continuous-time Markov chain is **irreducible** if there is a path between any states x and $\widetilde{x}$

A continuous-time Markov chain is **aperiodic** greatest common divisor of the length of all cycles equals one.

Theorem

If $\{X(t), t \in \mathcal{T}\}$ is an irreducible and aperiodic continuous-time Markov chain and the detailed balance condition holds

$$P_{\widetilde{x}} \cdot q(\widetilde{x}, x) = P_x \cdot q(x, \widetilde{x})$$

then $P_{\widetilde{x}}$ is the unique limiting (stationary) distribution of $\{X(t), t \in \mathcal{T}\}$

ERGMs and SAOMs

Let us now consider a particular SAOM:

- objective function for each actor i

$$f_i(\beta, x(i \rightsquigarrow j)) = \sum_{i=1}^{K} \beta_k s_k(x(i \rightsquigarrow j) = \beta' s(x(i \rightsquigarrow j))$$

- rate parameter for actor i

$$\lambda_i = \sum_{h=1}^{n} exp\left(\beta' s(x(i \rightsquigarrow h))\right)$$

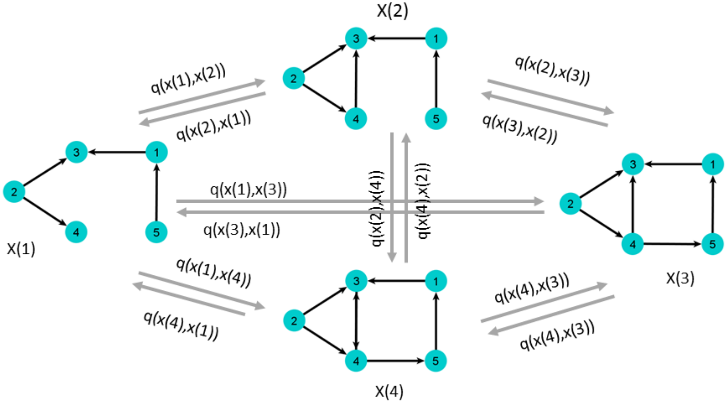
i.e., actors for whom changed relations have a higher value, will indeed change their relation more quickly.

ERGMs and SAOMs

The rate and the objective functions define a continuous-time Markov chain on the set $\mathcal{X}$.

The associated intensity matrix q of the process is:

$$\begin{aligned} q(x, x(i \rightsquigarrow j)) &= \lim_{dt \to 0} \frac{P(X(t+dt) = x(i \rightsquigarrow j) | X(t) = x)}{dt} \\ &= \lambda_i p_{ij} = exp(\beta' s(x(i \rightsquigarrow j)) \end{aligned}$$



Computing the limiting distribution of SAOMs

We can prove that ERGMs

$$P(X = x) = \frac{exp\left(\sum_{i=1}^{K} \beta_k s_k(x)\right)}{\kappa(\theta)} = \frac{exp(\beta' s(x))}{\kappa(\theta)}$$

are the unique stationary distribution of the SAOM defined before

Proof

1. Existence of a unique invariant distribution

- q is irriducible: each network configuration can be reached from any other network configuration in a finite number of steps
- q is aperiodic: at each time point t an actor i has the opportunity not to change anything and, thus, the period of each state is equal to 1

Proof (continue)

2. ERGMs are the stationary distribution of Q

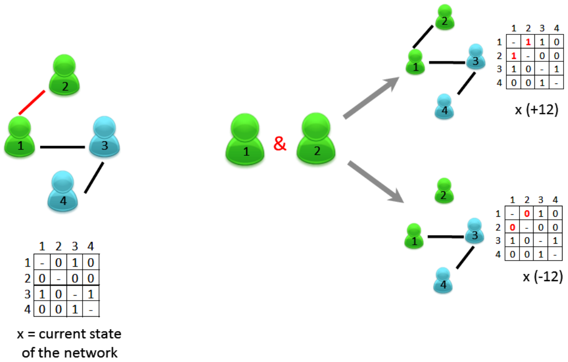
In fact, given two states x and $x(i \rightsquigarrow j)$ of $\{X(t), t \in \mathcal{T}\}$ the balance equation holds when ERGMs is the stationary distribution:

$$
\begin{aligned}
P_{x(i \rightsquigarrow j)} \cdot q(x(i \rightsquigarrow j), x) &= \frac{exp(\beta' s(x(i \rightsquigarrow j))}{\kappa(\theta)} \cdot exp(\beta' s(x)) \\
&= \frac{exp(\beta' s(x))}{\kappa(\theta)} \cdot exp(\beta' s(x(i \rightsquigarrow j)) \\
&= P_x \cdot q(x, x(i \rightsquigarrow j))
\end{aligned}
$$

SAOMs for non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

Joint decision: the decision is made on the payoff deriving from the tie



Actor 1 and 2 choose their action with probability

$$p(x^{+12}) = \frac{exp(f_{12}(x^{+12})}{exp(f_{12}(x^{+12})) + exp(f_{12}(x^{-12}))}$$

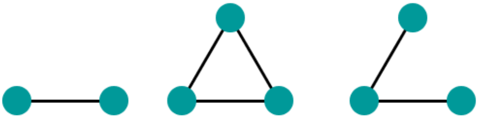
SAOMs for non-directed relations - Tie-based approach

We assume that

- each dyad (i, j) can be selected with the same rate λ
- the objective function is:

$$f_{ij}(\beta, x) = \sum_{i=1}^{k} \beta_k s_{ijk}(x) = \beta' s_{ij} x$$

where $s_{ijk}(x)$ are statistics such as



edges triangles 2-stars ...

but considered from the point of view of each pair (i, j) instead of the point view of a certain actor.

SAOMs for non-directed relations - Tie-based approach

Assuming that at each time point only one pair (i, j) can be selected, the rate function λ and the objective function $f_{ij}(\beta, x)$ define a continuous time Markov-chain with intensity matrix Q:

$$q(x, x^{+ij}) = \lambda p(x^{+ij}) = \lambda \frac{exp(\beta' s_{ij}(x^{+ij})}{exp(\beta' s_{ij}(x^{+ij})) + exp(\beta' s_{ij}(x^{-ij}))}$$

$$q(x, x^{-ij}) = \lambda p(x^{-ij}) = \lambda \frac{exp(\beta' s_{ij}(x^{-ij})}{exp(\beta' s_{ij}(x^{+ij})) + exp(\beta' s_{ij}(x^{-ij}))}$$

The limiting distribution of q is again ERGMs

Computing the limiting distribution of tie-based SAOMs

Proof

1. Existence of a unique invariant distribution

- q is irriducible: each network configuration can be reached from any other network configuration in a finite number of steps
- q is aperiodic: at each time point t a pair (i, j) has the opportunity not to change anything and, thus, the period of each state is equal to 1

Computing the limiting distribution of SAOMs

Proof (continue)

2. ERGMs are the stationary distribution of Q

In fact, given the two states x^{-ij} and x^{+ij} of $\{X(t), t \in \mathcal{T}\}$ the balance equation holds when ERGMs is the stationary distribution:

$$
\begin{aligned}
P_{x^{-ij}} q(x^{-ij}, x^{+ij}) &= \frac{e^{\beta' s(x^{-ij})}}{\kappa(\theta)} \cdot \lambda \cdot \frac{e^{\beta' s_{ij}(x^{+ij})}}{e^{\beta' s_{ij}(x^{+ij})} + e^{\beta' s_{ij}(x^{-ij})}} \\
&= \frac{e^{\beta' s(x^{-ij}) - \beta' s(x^{+ij}) + \beta' s(x^{+ij})}}{\kappa(\theta)} \cdot \frac{\lambda}{1 + e^{(}\beta' s_{ij}(x^{-ij}) - \beta' s_{ij}(x^{+ij}))} \\
&= \frac{e^{\beta' s(x^{+ij})}}{\kappa(\theta)} \cdot \lambda \cdot \frac{e^{\beta' s(x^{-ij}) - \beta' s(x^{+ij})}}{1 + e^{\beta' s_{ij}(x^{-ij}) - \beta' s_{ij}(x^{+ij})}} \\
&\underset{(*)}{=} \frac{e^{\beta' s(x^{+ij})}}{\kappa(\theta)} \cdot \lambda \cdot \frac{e^{\beta' s_{ij}(x^{-ij})}}{e^{\beta' s_{ij}(x^{+ij})} + e^{\beta' s_{ij}(x^{-ij})}} \\
&= P_{x^{+ij}} \cdot q(x^{+ij}, x^{-ij})
\end{aligned}
$$

$(*)$ $\beta' s(x^{-ij}) - \beta' s(x^{+ij})$