Analysing the dynamics
of non-directed networks

Tom A.B. Snijders

University of Groningen
University of Oxford

July 2022

Overview

Non-directed networks ... ?

Most networks occurring in the social sciences are directed.

When non-directed networks arise, they often were constructed as two-mode projections from directed networks.
In such cases it often is advisable to analyze the original data as a two-mode network.

Some networks are non-directed because of how they are collected; e.g., being observed together.
Or the directionality can be omitted from the analysis on purpose.

The Stochastic Actor-oriented Model is designed specifically for directed networks, where the actor has 'control' over outgoing ties.

For non-directed networks, special arrangements have to be made to apply the basic ideas of the Stochastic Actor-oriented Model.

Statistical inference for network dynamics

The method explained here is described in
Snijders, Tom A. B., and Pickup, Mark. 2016.
Stochastic Actor-Oriented Models for Network Dynamics. Pages 221–247 of:
Jennifer Nicoll Victor, Mark Lubell, and Alexander H. Montgomery (eds),
Oxford Handbook of Political Networks. Oxford: Oxford University Press.

We consider the regular type of *network panel data*:
2 or more repeated observations on a network between a fixed set of *n* actors
with ties referring to a <u>state</u> of the relationship between the two actors;
but here ties are non-directed = two-sided.

E.g.: mutual friendship; being a regular sex partner;
collaboration; strategic alliance.

The relation is denoted by the graph / adjacency matrix X,
where tie variable $X_{ij} = X_{ji}$ indicates by values 1 and 0,
respectively, that actors i and j are tied / are not tied.
A specific feature of non-directed = two-sided links
is that it is natural to assume that for the existence of the link,
both actors involved need to agree
(cf. Myerson 1991, Jackson & Wolinsky 1996);
other possibilities may also be considered.

For *non-directed* networks, effects are simpler than for directed networks,
but the dynamics is more complicated to model
because two actors are involved in tie creation / break up.

Basic assumptions

The basic approach of the Stochastic Actor-oriented Model is followed:

1. Between the observation moments, time runs on continuously.
   Changes can be made (unobserved) at any moment t.

2. The network $X(t)$ is the outcome of a Markov process.

3. At each single moment at most one tie variable may change ('ministep').

4. Initial state is given and not modelled as such.

These assumptions were proposed by Holland & Leinhardt, 1977.

Further assumption 2: two-step process:

1. Opportunity for changing one tie variable X_{ij} ;
   these opportunities occur continuously between observations.
   Rate functions express rate of change.

2. Tie probabilities depend on $\left\{ \begin{array}{c} \text{actor } i \\ \text{actor } j \end{array} \right\}$ in some interdependence.

   Objective functions define these probabilities.

1. Opportunities for change

This first model component specifies the stochastic time moments
where a tie variable *could* be changed.
Two options:

[1]. *One-sided initiative & proposal* :
   one actor is chosen – denoted i –
   for whom one tie variable X_{ij} may change,
   where j is to be proposed by i.

[2]. *Two-sided opportunity*:
   a pair of actors meet – denoted (i, j) –
   who may change their tie variable X_{ij}.

The moments where this happens constitute a stochastic process in continuous time:

[1]. for actor i, opportunities occur at a rate λ_i

[2]. for pairs (i, j), meetings occur at a rate λ_{ij} , e.g. $\lambda_{ij} = \lambda_i \lambda_j$.

('At a rate r means that in short time intervals of length dt, the probability of occurrence is approximately $r \times dt$.)

The rate functions λ_i and λ_{ij} may be constant, but can also depend on covariates and network position (degree, etc.)

Parameter interpretation rates of change:

When there is an opportunity for change,
it is permitted that nothing changes.

In Models [1], where the initiative is one-sided,
the rate function is an upper limit to
expected number of changes *per actor* per unit of time.

In Models [2], the rate function is an upper limit to
expected number of changes *per dyad* per unit of time.

('Upper limit' because some ministeps will lead to no change.)

2. Decisions about changing ties

When there is an opportunity for change, actors decide on
changes in their ties depending on preferences and constraints,
all subsumed in one *objective function* $f_i(\beta, x)$
(i is the actor, x is the state of the network)
plus unknown (random) influences.

β represents the unknown parameters
that will have to be estimated statistically.

Probabilities of change depend on the objective function.

Different ways for combining actors' objectives

- D Unilateral imposition of a tie (Dictatorial, disjunctive).
- M Mutual agreement required for a tie to exist (conjunctive).
- C Gain for one may outweigh loss for the other (Compensatory).

This is to be combined with [1]: unilateral initiative; [2]: two-sided opportunity.

Combination [1]-C (one-sided proposal & compensating objectives)
is possible but not implemented in RSiena.

Five modeltypes: 1D, 1M

[1 D.] Forcing model:
one actor *i* takes the initiative,
chooses the best possible change $x_{ij} \Rightarrow 1 - x_{ij}$ (or none)
and unilaterally imposes that this change is made.

[1 M.] Unilateral initiative and reciprocal confirmation:
one actor *i* takes the initiative,
chooses the best possible change $x_{ij} \Rightarrow 1 - x_{ij}$ (or none);
if this is the dissolution of a tie, the change is carried out,
otherwise the new tie is proposed to *j*,
if this actor agrees then the change is carried out,
otherwise nothing happens.

Five modeltypes: 2D, 2M, 2C

[2 D.] Pairwise disjunctive (forcing) model:
actors i and j meet and reconsider their tie variable X_{ij};
if at least one wishes a tie, then they set $X_{ij} = 1$, else $X_{ij} = 0$.

[2 M.] Pairwise conjunctive model:
actors i and j meet and reconsider their tie variable X_{ij};
if both wish a tie, then they set $X_{ij} = 1$, else $X_{ij} = 0$.

[2 C.] Pairwise compensatory (additive) model:
actors i and j meet and reconsider their tie variable X_{ij};
on the basis of their summed objective function $f_i(\beta, x) + f_j(\beta, x)$
they decide on the new value of the tie variable.

Decisions by actor *i*

In models [1],
actor *i* gets the initiative and must choose which tie variable to change;
this is the usual multinomial choice
in the ministep in the Stochastic Actor-oriented Model.

Binary decisions

In Models [2] there are binary (yes/no) decisions to be taken;
also for the agreement by the second actor j in model [1M].
These are based on the comparison by the actor
of the network $x^+(i,j)$ <u>with</u> tie $i - j$, and the network $x^-(i,j)$ <u>without</u> this tie.

This is a binary choice.
The same objective function is used as for the multinomial choice.

This means that in model [1M], the same objective function is used
for a multinomial choice and a binary choice.
The distinction is made by using a different constant term
(defining the balance between tie creation and tie deletion).

This is done by an *offset parameter* β_{offset}.

Offset parameter in model [1M]

The *offset parameter* for the binary choice by j in model [1M] is denoted β_{offset}.

The probability that actor j wishes the tie $i - j$ is given by

$$p_{ij}^{(j2)}(\beta, x) = \frac{\exp\left(f_j(\beta, x^+(i, j)) + \beta_{\text{offset}}\right)}{\exp\left(f_j(\beta, x^-(i, j))\right) + \exp\left(f_j(\beta, x^+(i, j)) + \beta_{\text{offset}}\right)}$$

If $\beta_{\text{offset}} = 0$, the usually strongly negative value of the (out)degree parameter
will lead to low probabilities of acceptance of the proposal by j.
Many ministeps will be needed, and estimation may be very slow.

A positive value for β_{offset} will lead to higher acceptance probabilities.
Reasonable values are about 1, or 2 for larger sparse networks.

In the Method of Moments, the *offset parameter* β_{offset} is collinear with the degree
parameter[1]. Therefore it cannot be estimated.
However, the precise value usually does not matter much.

---

[1] I.e., plausible estimation statistics are collinear.

Effects for non-directed networks

In non-directed networks,
there is no distinction between sender and receiver of the tie.

RSiena therefore contains several effects for which this distinction is dropped,
obtained by adding the corresponding effects for sender and for receiver.

* degree activity plus popularity effect (degPlus).

* covariate effect (egoPlusAltX).

* transitive triads effect (transTriads).

See the manual for some further versions.

This model definition, with components

$$
\begin{cases}
\textit{[1] / [2].} \text{ individual – dyadic initiative, with rate function } \lambda_i - \lambda_{ij} \\
\textit{D / M / C.} \text{ disjunctive – conjunctive – compensatory decisions} \\
\qquad \text{based on objective function } f_i(x)
\end{cases}
$$

yields (given that λ_i or λ_{ij} and $f_i(x)$, $g_i(x)$ are specified)
a model for the network dynamics that can be simulated.

Parameters can be estimated using Method of Moments estimators
(i.e.: equate expected statistics to observed statistics) implemented by
stochastic approximation implemented in RSiena.

Commands to use in RSiena

RSiena will automatically detect that the network is non-directed.

Some special parameters have to be used in `sienaAlgorithmCreate`;
in the first place `modelType`; values are:

2: dictatorial forcing (1.D) (*default*),

3: Initiative with reciprocal confirmation (1.M) (*often the most attractive*),

4: Pairwise dictatorial forcing (2.D),

5: Pairwise mutual model (2.M),

6: Pairwise joint model (2.C).

Second, for Model [1M], `OffSet`: usually from 0.5 to 3,
roughly proportional to $\log(n)$, results not very sensitive.

Recommendations for Model [1M]

Sometimes the start of the estimation leads to divergence
because of a too low value of the `Offset` parameter.

Advice: start with `Offset = 2` or `3`, after that
estimate again with a reduced value of `Offset`, e.g., `1`, and use `prevAns`.

An alternative is to start with Model [1D],
and then with `prevAns` move to Model [1M] with `Offset=1`.

Estimation with a too low value of `Offset` may take very long,
especially if parameters are still far from their appropriate values!

`Offset` is a 'named vector':
In `sienaAlgorithmCreate` use `Offset=c('name'=2)`,
where `name` is the name of the dependent network.

Parameter differences between models

In the first place, the five models differ in the rate and degree parameters.

The rate parameters for the dyadic models are lower, because they refer to the rate at which <u>dyads</u> rather than <u>actors</u> change their ties.

The degree parameters are quite different, because the tie creation–deletion balance is modeled quite differently.

For the other parameters, in many cases the ratio of parameter estimate to standard error is pretty similar between the five models.

Parameters for model [2M] will be larger than for [2D] and [2C], because both actors need to agree to the tie.

For some illustrations of the differences, see the older set of slides https://www.stats.ox.ac.uk/~snijders/siena/TwoSided_sa.pdf ; but the Offset parameter was not yet developed there, so implicitly it is 0.

Example: grooming between vervet monkeys

Christèle Borgeaud, Sebastian Sosa, Cédric Sueur, Redouan Bshary, and Erica van de Waal (2016). Intergroup Variation of Social Relationships in Wild Vervet Monkeys: A Dynamic Network Approach.

Frontiers in Psychology 7, 915.

Three groups of wild vervet monkeys followed over 2 years, with 8 observation periods, which are aggregates over 3 months.

Here:
results for one group ('AK'),
group size varying 26–33,
47 different individuals.

Network: grooming and proximity.

Model specification following the ideas of Borgeaud et al.,
but with my own experimentation, and checking time heterogeneities.

During preliminary modeling,
sienaTimeTest showed large time heterogeneities;
taking out wave 3 gave a large improvement, but heterogeneity remained.

Results with and without wave 3, with and without time dummies,
are very similar.

Results given are for waves 1-2-4-5-6-7-8.

Descriptives

Wave	t_1	t_2	t_4	t_5	t_6	t_7	t_8
Number of individuals	26	29	31	30	26	28	33
Average degree	4.3	3.5	3.4	2.3	2.8	2.8	4.3
Jaccard similarity with previous wave		.29	.17	.23	.19	.17	.14

Covariates

⇒ *Matriline*: 6 kinship groups

⇒ *Hierarchical level* values 1–9 (de Vries I&SI method)

Estimation results

modelType=3, 'Initiative with reciprocal confirmation (1.M)', Offset $\beta_{\text{offset}} = 1$.

Effect	shortName	par.	(s.e.)
effect sex.M on rate	RateX	1.057^{***}	(0.300)
degree (density)	density	-1.091^{***}	(0.126)
GWESP (transitivity)	gwesp	0.339^{**}	(0.115)
degree	degPlus	0.007	(0.010)
same Matriline	sameX	0.336^{**}	(0.129)
sex M	egoPlusAltX	0.177^{**}	(0.055)
same sex	sameX	-0.030	(0.071)
Hierarchy level	egoPlusAltX	0.025^{*}	(0.012)
same Hierarchy level	sameX	0.370^{**}	(0.128)
same sex $\times$ Hierarchy level	unspInt	-0.040^{*}	(0.018)

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

convergence t ratios all < 0.06, overall maximum convergence ratio 0.12.

Interpretation:

- males change grooming partners more frequently;

- evidence for transitivity;

- strong effect of same matriline;

- strong homophily effect ('same') of hierarchy;

- hierarchy positively associated with grooming;

- positive effect of male sex;

- negative interaction same sex $\times$ hierarchy:
  high hierarchy values groom with different-sex partners,
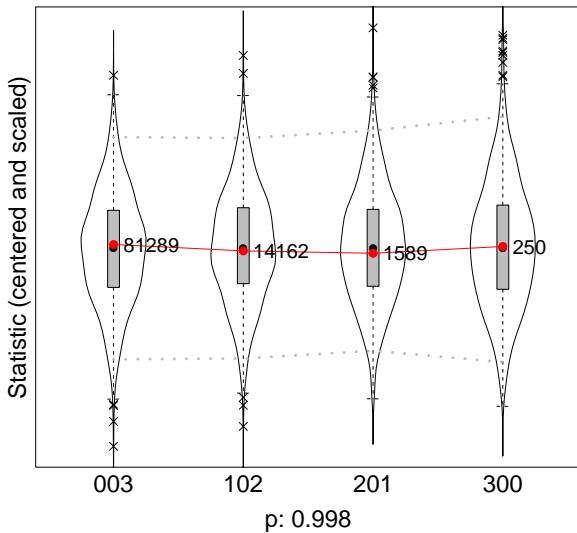 low hierarchy with same-sex.

Goodness of fit

For non-directed networks, `sienaGOF` can be used to check goodness of fit.
No distinction between in- and out-degrees.

Triad census has only four possibilities (number of ties in the triad)
and is easily fitted well.

Goodness of Fit of TriadCensus

Goodness of fit
for triad census

Goodness of fit: clique census

An important network statistic diagnosing the dyadic nature of ties
as opposed to ties following a pattern of connected groups
is the *clique census*.

A clique in a graph is a maximal totally connected subgraph:
a set of nodes that are all mutually connected,
such that for every node outside of the set,
at least one connection with the set is lacking.

If the non-directed network is a combination of many groups of size ≥ 4,
e.g., a combination of one-mode projections of two-mode networks,
this will be shown by an overabundance of cliques of size ≥ 4.

The clique census is an important diagnostic for non-directed networks.
Available at the help page for `sienaGOF-auxiliary`.

In R you need quotes to get to this help page: **?'sienaGOF-auxiliary'**

Goodness of fit: clique census (2)

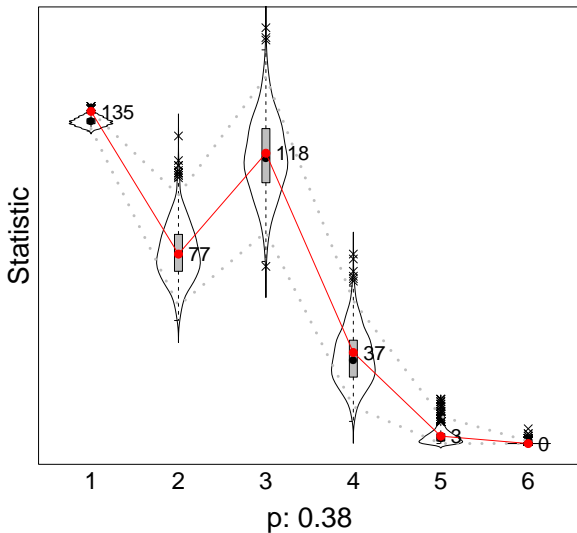Observed clique sizes range from 1 (isolated individuals) to 5.

Totals over waves 2, 4–8:

Clique size	1	2	3	4	5
Observed frequency	135	77	118	37	3

The following page shows that the simulations
reproduce the clique frequencies quite decently.

Goodness of Fit of CliqueCensus

Goodness of fit
for clique census

Summary / Discussion

The fact that two-sided decisions must be modeled
leads to a variety of different models.
Choosing between them might be difficult
if there is a lack of convincing theory, which usually is the case.

modelType=3, 'Initiative with reciprocal confirmation (1.M)',
is often preferable (but not the default...).
For the Offset in sienaAlgorithmCreate, use values such as 1 or 2.

The two-step model of initiative with reciprocal confirmation
can also be used for two-mode networks;
see script TwoModeAsSymmetricOneMode_Siena.R .