Siena Advanced Users' Meeting, 2017

Tom A.B. Snijders

University of Oxford
University of Groningen

AdSUM, February 2017

Overview

This is a collection of slides, *some of which* will be presented at the advanced RSiena meeting.

1. Where to look?

Siena is an evolving endeavour, which may be hard to follow.

1. Where to look?

Siena is an evolving endeavour, which may be hard to follow.

- The CRAN version is out of date.
  Consult the Siena website
  `http://www.stats.ox.ac.uk/~snijders/siena/`
  ('*Downloads*' page) for latest versions.
  Normally, this is the R-Forge version.
- Literature: the 2010 tutorial in *Social Networks*;
  the manual (at the website, frequently updated);
  the scripts at the website;
  in 2017/8, there will be books (we hope).
- The website notes important matters at the '*News*' page:
  incompatibilities, bugs, new developments, papers.

Where to look? (2)

- Follow the Siena/Stocnet discussion list!
  Announcements of new versions, bugs, etc.
- Website '*News*' page, and Appendix B in the manual, give description of changes in the new versions.
- Website '*Literature*' page has a section
  'Presentations (teaching material)'
  including (e.g.) these slides.
- Siena_algorithms.pdf now is at the Siena website (partial explanation of algorithms and code).
- The available effects of 'myeff' are given by effectsDocumentation(myeff).

New convergence criterion

The earlier convergence criterion is tmax,
the absolute maximum of the *t*-ratios for convergence,
considering simultaneously all parameters in the model.

It has appeared that for some models
(e.g., with non-centered actor covariates)
the usual criterion

$$\text{tmax} \leq 0.10$$

is not sufficient.

Therefore, the overall maximum convergence ratio
(included as tconv.max in *sienaFit* objects since some time)
gets a new importance.

Overall maximum convergence ratio

This is defined as the maximum t-ratio for convergence for any linear combination of the parameters,

$$\text{tconv.max} = \max_b \left\{ \frac{b'\big(\bar{s}_j - s^{\text{obs}}\big)}{\sqrt{b'\Sigma b}} \right\} .$$

This is equal to (see *Siena_algorithms.pdf*)

$$\max_c \left\{ \frac{c'\Sigma^{-1/2}\big(\bar{s}_j - s^{\text{obs}}\big)}{\sqrt{c'c}} \right\} = \sqrt{\big(\bar{s}_j - s^{\text{obs}}\big)'\Sigma^{-1}\big(\bar{s}_j - s^{\text{obs}}\big)} .$$

The definition implies that

$$\text{tconv.max} \geq \text{tmax} .$$

Studies comparing results of siena07() with the 'true estimate' (robust mean of many estimations) show:

1. Distance from true estimate is much better indicated by tconv.max than by tmax.
2. When tconv.max exceeds 0.30, distances from the true value are too large.

New criterion

$$\text{tmax} \le 0.10 \ \underline{\text{and}} \ \text{tconv.max} \le 0.25 \, .$$

(Sometimes these values are hard to attain, and tconv.max between 0.25 and 0.30 may also be acceptable)

What to do if convergence is hard to attain

To understand what to do against convergence difficulties, recall the structure of the estimation algorithm with 3 phases:

1. brief phase for preliminary estimation of sensitivity of estimation statistics Z to parameters θ;
2. estimation phase with Robbins-Monro updates for θ, consisting of *nsub* subphases (usually 4) with decreasing step sizes, determined by *firstg*;
3. final phase with *n3* runs, θ constant at estimated value $\hat{\theta}$; this phase is for checking that

   $$\mathrm{E}_{\hat{\theta}}\{Z\} - z \qquad \text{is small}$$

   and for estimating standard errors.
   These are the ‘deviations from targets’.

Algorithm is set in `sienaAlgorithmCreate()`,
parameters are estimated by `siena07()` or `sienacpp()`.

If estimation diverges right away,
check data and model specification;
perhaps use a simpler model.

If estimation still diverges right away, either:

- ⇒ estimate a simpler model, and use the result for *prevAns*;
- ⇒ use a smaller value for *firstg*;
  default is 0.2, suggestions for smaller are 0.01 or 0.001.
  Note that this implies the algorithm moves more slowly,
  and when the estimation is 'on track', it is better
  to continue further estimation runs with *prevAns*
  with a larger value of *firstg* (e.g., the default).

The normal procedure is to repeat estimation,
using the *prevAns* parameter in *siena07*,
until *tconv.max* ≤ 0.25.

For complicated models, this may be unsuccessful, and *tconv.max* from some moment does not systematically decrease any more.

It turns out that *tconv.max* is determined by the length of the last subphase of Phase 2 (& random noise)

If a low value of tconv.max is not easily achieved, for getting better convergence in continued `siena07` runs with *prevAns*:

- ⇒ from the estimation runs where *tconv.max* is 0.3 or so, use algorithm settings with *nsub*=1, *n2start*=‘large’, where ‘large’ is in the range of 1,000 to 10,000, and the ‘regular’ (not-large) value is $40 \times (p + 7)$, with $p =$ number of parameters; *n3* large (e.g., 2,000 or 5,000); *firstg* small (e.g., 0.01).
- ⇒ If you expect it coming right away, instead of this, you can use more subphases: *nsub = 4* or 5 (but further with default settings).
- ⇒ If *tconv.max* still too big, further increase *n2start*.

A related issue, different but sometimes occurring in the same data-model combinations because of their complexity, is that sometimes standard errors are instable and sometimes strongly over-estimated.

This may disappear when re-estimating the model with a sufficiently long Phase 3 (i.e., high *n3*).

Ongoing work by Nynke Niezink.

3. Causality?

Network data are often observational, and relations are crucial for how social actors try to attain their goals.
Therefore, networks in real life are highly endogenous.

Attaining causal conclusions about network effects from non-experimental studies is hard, because if ties are changed, actors will try something else that is similarly helpful for what they try to attain.

Causality in observational research, certainly for networks,
is a Holy Grail: a lofty and important aim,
which we should not expect to attain;
cf. Shalizi & Thomas (2011):
selection and influence are generically confounded.

D.R. Cox / R.A. Fisher about causality:
Make your theories elaborate,
construct explanations at a deeper level.

P. Hedström & P. Ylikoski: *causal mechanisms*.

Network approaches themselves are a deeper level
than traditional quantitative social science approaches,
representing interaction processes,
and in this sense may help in coming closer to causal insights.

Stochastic Actor-oriented Modeling approach does not lead to
causal conclusions in the Holland-Rubin counterfactual sense;
it leads to conclusions about time sequentiality.

Process approach

We should realize that one of the ways in which network research differs from much traditional social science is its framing in terms of dependencies between actors and in terms of representing detailed processes, contrasting with variables defined for isolated actors affecting each other.

....

We should be aware that this may be a large step for many colleagues, supervisors, reviewers, and sometimes leads to confusion and misunderstanding.

Process approach

We should realize that one of the ways in which network research differs from much traditional social science is its framing in terms of dependencies between actors and in terms of representing detailed processes, contrasting with variables defined for isolated actors affecting each other.

variable thinking?

We should be aware that this may be a large step for many colleagues, supervisors, reviewers, and sometimes leads to confusion and misunderstanding.

Process approach

We should realize that one of the ways in which network research differs from much traditional social science is its framing in terms of dependencies between actors and in terms of representing detailed processes, contrasting with variables defined for isolated actors affecting each other.

~~variable thinking~~ process thinking

We should be aware that this may be a large step for many colleagues, supervisors, reviewers, and sometimes leads to confusion and misunderstanding.

Network analysis requires new methodologies:

- We are used to thinking in terms of variables,
  defined for isolated actors, affecting each other;
  or perhaps variables defined for isolated dyads,
  or variables for nested levels;
- We are accustomed to basing models on independence;
- Thinking in terms of processes in networks,
  and dependence between actors, is quite different;
  we are only starting to understand how to specify
  dependence. This implies a larger place for explorative
  parts in theory-guided research.
- Without independence assumptions can rely less on
  mathematical theorems supporting statistical methods.
  We can/should make our methods reproducible.

4. sienacpp()

RSiena has two rooms:

1. front office: user interface in R
2. back office: simulations going on in C++

In siena07(), only the simulations are done in C++;
the further calculations for the Robbins-Monro estimation algorithm are done in R.

Starting from version 1.1-290 (early 2016), RSienaTest contains sienacpp() which produces the same as siena07(), but with all calculations in C++.
(Programmed by Felix Schönenberger.)

(Some options are not yet included, e.g., multigroup data.)

sienacpp() has a small efficiency advantage,
which is relatively important only for
small data sets / small amounts of total change.

Parallellization options may be different.

5. How to specify the model?

This depends of course on the purpose of the research, theoretical considerations, empirical knowledge...

But the following may be a guideline for specifying the network model:

1. Outdegree effect: always.
2. Reciprocity effect: almost always.
3. A triadic effect representing network closure. gwesp, transitive triplets, and/or transitive ties.
4. transitive reciprocated triplets and/or three-cycles (see Block, *Network Science*, 2015).

Interlude: GWESP effect

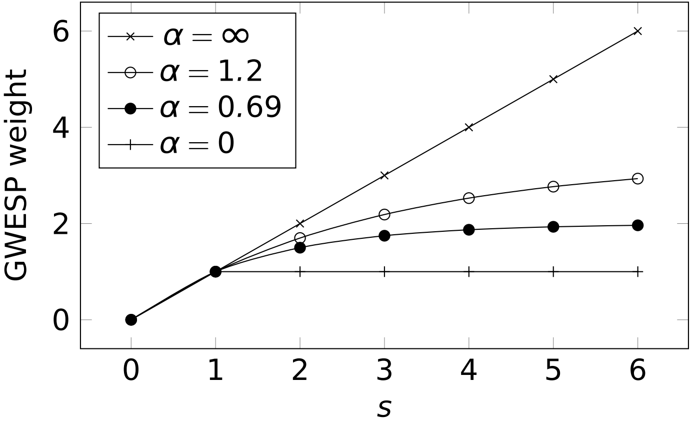
GWESP (*geometrically weighted edgewise shared partners*) (cf. ERGM!!) is intermediate between transTrip and transTies.

$$\text{GWESP}(i, \alpha) = \sum_j x_{ij}\, e^{\alpha} \left\{ 1 - \left(1 - e^{-\alpha} \right)^{\sum_h x_{ih} x_{hj}} \right\} .$$

for $\alpha \geq 0$ (effect parameter = $100 \times \alpha$).

Default $\alpha = \log(2)$, parameter = 69.

Interlude ... GWESP ......



Weight of tie $i \rightarrow j$ for $s = \sum_h x_{ih} x_{hj}$ two-paths.

Interlude ...... GWESP ...

The implementation of GWESP is an *elementary* effect:

For creation of a new tie,
only its role as $i \rightarrow j$ in the formula is counted,
not its role as $i \rightarrow h$.

Therefore it can be interacted with all dyadic effects.

GWESP sometimes yields better fit than transTrip or transTies.

Interlude ...... GWESP ...

The implementation of GWESP is an *elementary* effect:

For creation of a new tie,
only its role as $i \rightarrow j$ in the formula is counted,
not its role as $i \rightarrow h$.

Therefore it can be interacted with all dyadic effects.

GWESP sometimes yields better fit than transTrip or transTies.

The GWESP effect exists in many directions:
`gwespFF, gwespBB, gwespFB, gwespBF, gwespRR`
for F = Forward, B = Backward, R = Reciprocal,
and also for multivariate networks: `gwespFFMix` etc.

How to specify the model? *(continued)*

5. Degree-related effects:
   indegree-popularity ('Matthew effect'), outdegree-activity, outdegree-popularity and/or indegree-activity
   (raw or sqrt versions depending on goodness of fit; for high average degrees, preference for sqrt).
6. Think about what are important covariates!
   For actor covariates: see presentation by Tom Snijders, *Specification of Homophily in Actor-oriented Network Models*:
   for numerical actor variables there may be a combination of tendencies of homophily, aspiration, and social norm;
   use 5 effects:
   ego, alter, ego × alter, ego-squared, alter-squared.

How to specify the model? *(further continued)*

7. Use information about dyadic contact opportunities (same classroom, task dependence, distances, etc.)
8. If there is a strong center-periphery structure, and/or a strong dispersion in the outdegrees, then a dependence of the rate function e.g. on the log-outdegree (*outRateLog*) may be advisable.

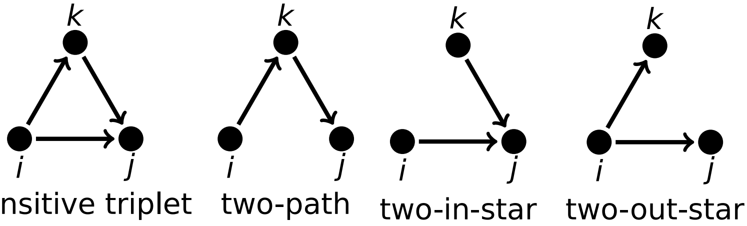
A large set of effects is available in RSiena,
growing over the years because of researchers' requests.

Model specification: hierarchy requirements

There are hierarchy principles somewhat like in regression analysis:
simpler configurations should be used as controls
for complicated configurations.

This leads to heavy controls for multiple network co-evolution and complicated multi-node effects.

Hierarchy: example



The transitive triplet (left) includes three subgraphs (right); actor i can create a transitive triplet by closing $i \rightarrow j$ or $i \rightarrow k$; therefore, to properly test transitivity, the two-path and two-in-star configurations should be included in the model. These correspond to the outdegree-popularity and indegree-popularity effects.

How to specify the model? *(even further continued)*

In addition to allowing you to answer your research questions, the model also should have a good fit to the data.

The fit can be checked, but always incompletely, by using `sienaTimeTest()` and `sienaGOF()`.

Note that difficulties in obtaining convergence of the estimation procedure may be a sign of model misspecification or overspecification.

(The converse is not true!!!)

Assumption checking

Two functions are available in RSiena
for checking model assumptions:

1. `sienaTimeTest()`
   for testing time heterogeneity
   (meaningful only if there are 3 or more waves);

Assumption checking

Two functions are available in RSiena
for checking model assumptions:

1. `sienaTimeTest()`
   for testing time heterogeneity
   (meaningful only if there are 3 or more waves);
2. `sienaGOF()`
   for checking that the RSiena model reproduces sufficiently
   the characteristics of the observed networks.

Both were developed by Josh Lospinoso (Oxford).

sienaTimeTest

For M waves there are $M-1$ periods.

The assumption
that parameters are constant in the $M-1$ periods
is tested by `sienaTimeTest()`.

The `summary()` method also produces effect-wise
and period-wise tests.

See `RscriptSienaTimeTest.r`

sienaTimeTest

For M waves there are $M-1$ periods.

The assumption
that parameters are constant in the $M-1$ periods
is tested by `sienaTimeTest()`.

The `summary()` method also produces effect-wise
and period-wise tests.

See `RscriptSienaTimeTest.r`

Can be used also to check homogeneity between groups
for multi-group projects!

The associated function `includeTimeDummy()` can be used to interact the effects specified by time dummies, representing time heterogeneity.

An alternative for this purpose is to define time variables
(dummies or trend or other time-dependent variables;
defined as changing actor covariates,
constant across actors and changing across waves)
and add those to the data set,
and then specify interactions
between the other effects and these time variables.

This is a bit more work but also more flexible and clearer.

sienaGOF()

The goodness of fit of a model
(does it reproduce the data well enough?)
can be tested by the function `sienaGOF()`.

This requires that `siena07()` was run
with *returnDeps* = TRUE.

This option returns the simulated data sets in Phase 3
as part of the `sienaFIT` object produced by `siena07()`.

(from the help page ...:)

This is done by simulations of auxiliary statistics,
different from the statistics used for parameter estimation.
The fit is good if the average values
of the auxiliary statistics over many simulation runs
are close to the values observed in the data.

A Monte Carlo test based on the Mahalanobis distance
is used to calculate *p*-values.

This is a case where you wish the *p*-values to be *large* enough!

A `plot()` method can be used to diagnose poor fit.

The auxiliary statistics must be given explicitly
in the call of `sienaGOF()`.

Some basic auxiliary statistics are available directly:

`OutdegreeDistribution()`

`IndegreeDistribution()`

`BehaviorDistribution()`

`CliqueCensus()` (useful for nondirected networks!) ;

and the user can also create custom functions.

The help page `sienaGOF-auxiliary` contains
some additional functions using packages `igraph` and `sna`.

Sketch of the use of sienaGOF

See `?sienaGOF` and the script `sienaGOF_new.R`

The basic operation is as follows:

```
results1 <- siena07(myalg, data=mydata, effects=myeff,
             returnDeps=TRUE)
gof1.od <- sienaGOF(results1, verbose=TRUE,
             varName="friendship", OutdegreeDistribution,
             cumulative=TRUE, levls=0:10)
gof1.od
plot(gof1.od)
```

You can adapt the parameters `levls` and `cumulative`.
For `levls` this is important!

Auxiliary functions

Some auxiliary functions are available within RSiena,
('out of the box'),
some are listed on the help page for `"sienaGOF-auxiliary"`,
such as `TriadCensus()` and `GeodesicDistribution()`,
and others can be made by yourself (...)
or in future by others (!!!).

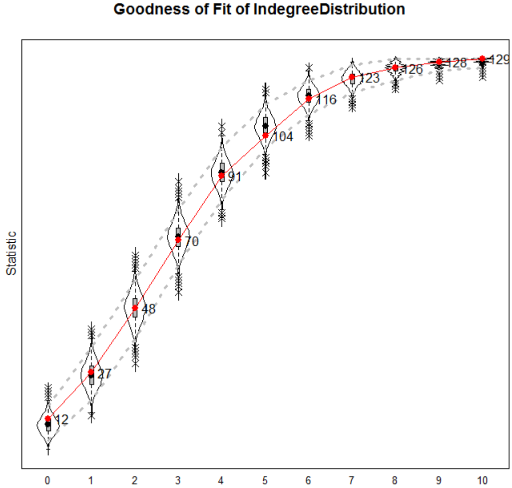
If you wish to use `TriadCensus()` and
`GeodesicDistribution()`, you have to take these,
for the latter along with `igraphNetworkExtraction()`,
from the `sienaGOF-auxiliary` help page and give them to R.

What is available now is not meant to be complete!

Goodness of fit: indegree distribution

Example of Goodness of Fit plot, indegree distribution

`levls=0:10` to cover all observed outdegrees



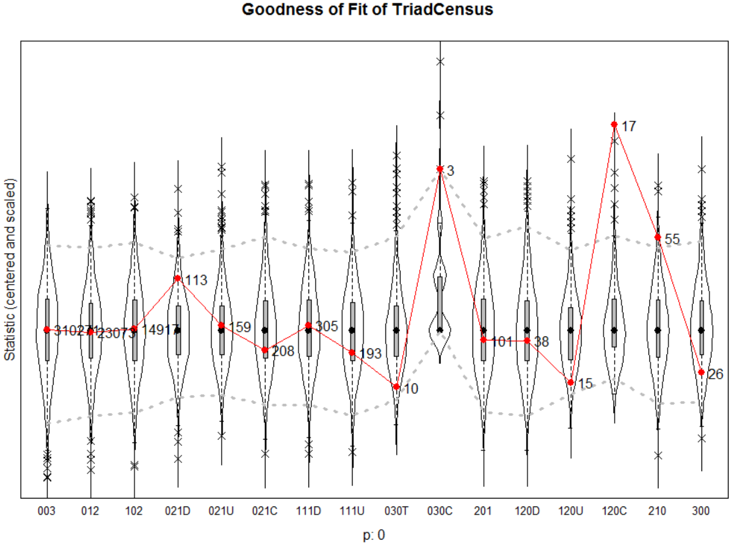
For more variable degrees, larger categories can be used, e.g.

```
gof1.od <- sienaGOF(..., IndegreeDistribution,
        cumulative=TRUE, levls=c(0:5, 10*(1:4)))
```

gives 0, 1, 2, 3, 4, 5, 10, 20, 30, 40 – cumulatively!

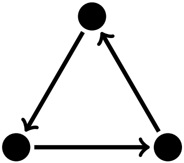
Example goodness of fit: triad census

Example of poor fit, triad census

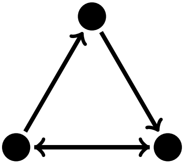


The triads that are not well represented are 030C, 120C, 210.

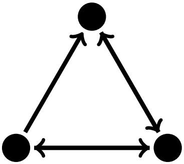
Look them up!



030C



120C



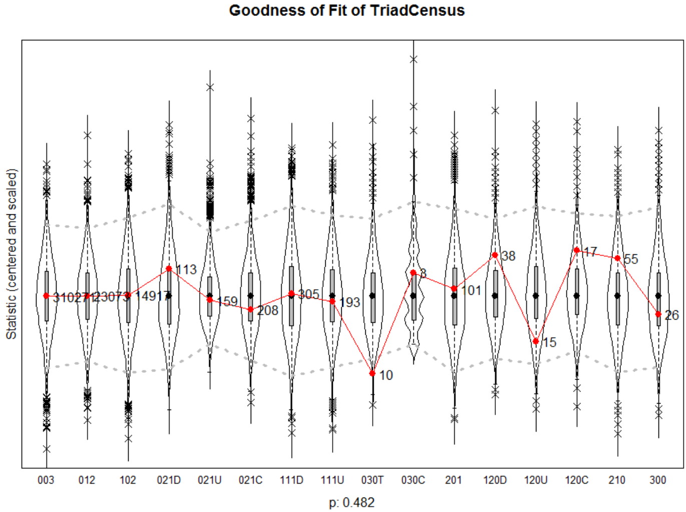
210

All of these triads are fewer in the simulations than in the data.

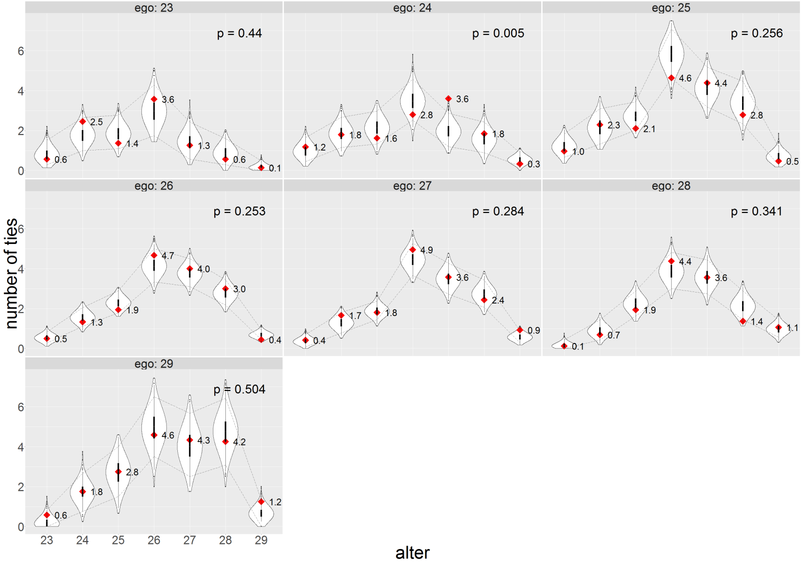
They all contain a 3-cycle (or two).
This suggests that the frequency of 3-cycles is not represented well.

Example goodness of fit: triad census(2)

Example of good fit, triad census



For the combination of senders' and receivers' values of monadic actor covariates, a new set of auxiliary statistics will be added, enabling the following kind of plots (next page).



Overall Mahalanobis combination $p = 0.045$.

How good a fit is required?

Since recently we have been moving to a new standard for publications using Siena, where the fit for the degree and behavior distributions should be adequate.

Of course it is also advisable to consider goodness of fit for the triad census and the geodesic distribution.

It may not always be possible to achieve a fit with $p > 0.05$ for the Mahalanobis combination of all statistics under consideration.

The traditional standards of 'significance' do not necessarily apply to p-values for goodness of fit assessment.

In my experience it usually is possible, to have the data within the confidence band of *plot.sienaGOF*.

If the fit for the degree distribution is poor,
one may consider the degree-related effects listed in the manual.

In all cases,
it is preferable to reflect on the processes under study
and consider additional covariate or structural effects,
interactions, non-linear transformations of covariates,
differences creation – maintenance, etc.:
finding out about misspecification in theoretically meaningful ways is a main road to scientific progress.

Further remarks

See the help pages for further information,
and Sections 5.11 and 8.6 of the manual.

Also see the scripts on the Siena scripts webpage.

First test time homogeneity, then goodness of fit.

Goodness of fit testing can be time consuming;
you may explore it with a Phase 3 of reduced length.

Testing of time homogeneity and goodness of fit
is getting more and more important.

Improving fit in this way
can led to theoretically interesting new insights!

Specification; effects

1. Structural equivalence: Jaccard distances
2. Weighted degree effects
3. Variations of influence effects
4. Distance-two effects
5. Elementary effects
6. Note: influence in one-mode and two-mode networks
7. Miscellaneous

Other new effects are also treated in earlier ‘Advanced Users’ presentations – see the website.

New effects (1): Structural equivalence

A good way of expressing structural equivalence,
i.e., being connected to the same others,
is the *Jaccard similarity* between rows, or columns:

$$J_{\text{out}}(i,j) = \frac{\sum_h x_{ih}\, x_{jh}}{x_{i+} + x_{j+} - \sum_h x_{ih}\, x_{jh}}$$

$$J_{\text{in}}(i,j) = \frac{\sum_h x_{hi}\, x_{hj}}{x_{+i} + x_{+j} - \sum_h x_{hi}\, x_{hj}}$$

Based on these (by summing over the outgoing ties of i),
the effects Jout and Jin are defined.

For multivariate networks: JoutMix, JinMix.

New effects (2): weighted degree effects

Degrees weighted by covariate: inPopX, outPopX, inActX, outActX
useful especially for non-centered X
(version 1.1-306)

New effects (2): Influence

The triple avSim – totSim – avAlt
now is a quartet with a 2 × 2 structure:
{ sim , alt } × { av, tot }

totAlt was implemented for regular influence effects,
influence from reciprocated alters, and
influence from other covariates (non-dependent / exogenous).

totAlt and totSim may need controlling for outdeg.

New effects:

1. totAlt (next to avAlt, totSim, avSim)
2. totRecAlt (next to avRecAlt)

New effects (3): influence weighted by popularity

influence weighted by indegrees: avAltPop, totAltPop
(version 1.1-306)

Reminder of effects (3): Influence from covariates

Influence on a given behavior could also come from another attribute of the alters
(e.g., effect of work attitude of friends on performance).

monadic: `avXAlt, totXAlt`;
dyadic: `avWAlt, totWAlt`;

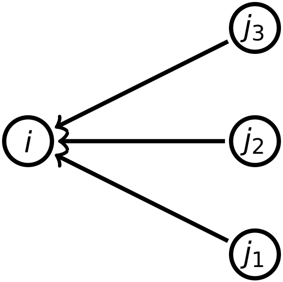
don't confuse with av/totAltW: av/totAlt weighted by W.

tot*Alt may need controlling for outdeg.

Incoming influence effects

The effects avAlt – totAlt – avXAlt – totXAlt
now also have analogues for influence from incoming ties:

3. avInAlt
4. totInAlt
5. avXInAlt
6. totXInAlt



i is influenced by incoming ties $j_1 - j_3$

totInAlt and totXInAlt may need controlling for outdeg.

Extreme influence effects

7. maxAlt
8. minAlt

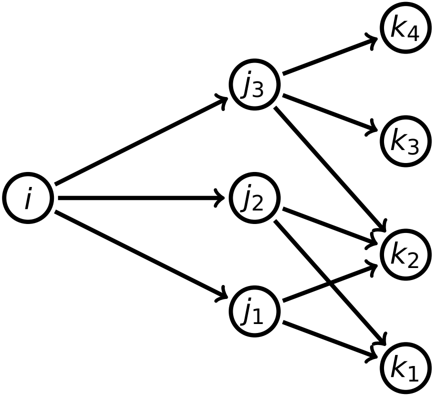
New effects (4): Distance-two

There now is the possibility to express influence at distance 2.

With the distinction average/total this leads to 4 possibilities: average vs. total at step 1 or step 2.

9. avAltDist2
10. totAltDist2
11. avTAltDist2
12. totAAltDist2

i is influenced by the average/total of the alter averages/totals of $j_1 - j_3$



New effects (5)

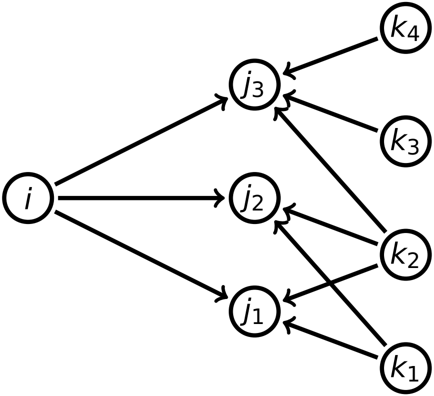
The same for distance-2 averages and totals of covariates:

13. avXAltDist2
14. totXAltDist2
15. avTXAltDist2
16. totAXAltDist2

New effects (6): outgoing - incoming

The same for distance-2 averages and totals where the second step is for incoming ties:

17. avInAltDist2
18. totInAltDist2
19. avTInAltDist2
20. totAInAltDist2
21. avXInAltDist
22. totXInAltDist2
23. avTXInAltDist2
24. totAXInAltDist2

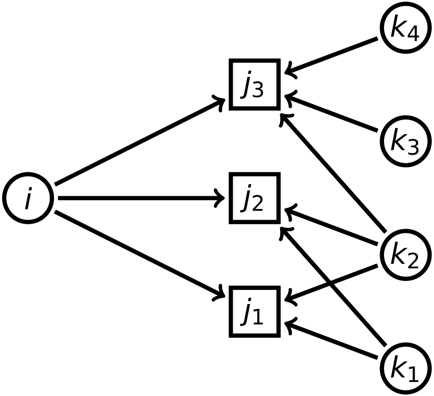


i is influenced by the incoming alter averages of $j_1 - j_3$.

Also ‘sim’ versions (simEgoInDist2 etc.)

New effects (6a)

The *InAltDist2 effects are
also available
for two-mode networks.



This means that it is now possible to model influence
from those out-alters
who have the same affiliations as the focal actor.

Structural equivalence again

These distance-two outgoing–incoming effects
can be regarded as representing influence from
actors who are structurally equivalent (w.r.t. outgoing ties).

An alternative would be to use Jaccard measures (cf. Jin, Jout)
for defining influence effects.
This is still for future consideration.

Elementary effects

SAOM effects have been framed in the triple

1. evaluation
2. maintenance/endowment
3. creation

effects.

If the parameters for a creation and corresponding maintenance effect are the same, then it can be represented just as well by an evaluation effect.

These kinds of effects differ in how they contribute
to the probability of a particular choice in the ministep.

The contributions to probabilities are based on
evaluation function f^{ev}
maintenance function f^{mt}
creation function f^{cr} .

Evaluation function plays a role for any step;
creation function only for upward change;
maintenance function only against downward change.

The definition is on the following page.

The probability that, given a current network x and actor i making the ministep, the network changes to $x^{\pm ij}$, is

$$\frac{\exp\Big(u_i\big(x, x^{\pm ij}\big)\Big)}{1 + \sum_{h \neq i} \exp\Big(u_i\big(x, x^{\pm ih}\big)\Big)}$$

where the objective function is

$$\begin{aligned} u_i\big(x, x^*\big) = f_i^{\text{ev}}(x^*) - f_i^{\text{ev}}(x) &+ \Delta^+\big(x, x^*\big)\big(f_i^{\text{cr}}(x^*) - f_i^{\text{cr}}(x)\big) \\ &+ \Delta^-\big(x, x^*\big)\big(f_i^{\text{mt}}(x^*) - f_i^{\text{mt}}(x)\big) \end{aligned}$$

and

$$\Delta^+\big(x, x^*\big) = \begin{cases} 1 & \text{if tie is created } (x^* = x^{+ij}) \\ 0 & \text{if tie is dropped, or no change} \end{cases}$$

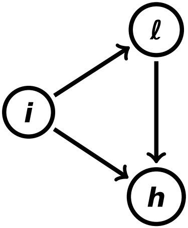
$$\Delta^-\big(x, x^*\big) = \begin{cases} 1 & \text{if tie is dropped } (x^* = x^{-ij}) \\ 0 & \text{if tie is created, or no change.} \end{cases}$$

However, not all probabilities of change can be based on changes in some (evaluation-type) function.

Example : transitive triplets

The transitive triplets effect is defined as

$$s_i(x) = \sum_{j,k} x_{ij}\, x_{ik}\, x_{kj}$$



with change statistic
(change when adding tie $i \rightarrow j$)

$$\delta_{ij}(x) = \sum_k x_{ik}\big(x_{kj} + x_{jk}\big) \, .$$

The first part refers to creating the tie $i \rightarrow j = h$,
the second part to creating the tie $i \rightarrow j = \ell$.

But one could be interested in only transitive closure,
as defined by closing of an open two-path ($i \rightarrow j = h$),
as distinct from creating ties
to those with the same out-choices,
which is a kind of structural equivalence ($i \rightarrow j = \ell$).

This cannot be represented
as a change in an evaluation function.

But one could be interested in only transitive closure,
as defined by closing of an open two-path ($i \rightarrow j = h$),
as distinct from creating ties
to those with the same out-choices,
which is a kind of structural equivalence ($i \rightarrow j = \ell$).

This cannot be represented
as a change in an evaluation function.

Therefore we need a different kind of effect:

But one could be interested in only transitive closure,
as defined by closing of an open two-path ($i \rightarrow j = h$),
as distinct from creating ties
to those with the same out-choices,
which is a kind of structural equivalence ($i \rightarrow j = \ell$).

This cannot be represented
as a change in an evaluation function.

Therefore we need a different kind of effect:

elementary effect

Elementary effect

An elementary effect is a term of the objective function $u_i(x, x^*)$ used to define change probabilities for ministeps, referring to creation and/or maintenance of a tie $i \to j$, without being necessarily a difference $f_i(x^{\pm ij}) - f_i(x)$ of some function f_i (or similar with multiplication by Δ^+ or Δ^-).

Elementary effect

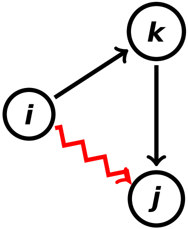
An elementary effect is a term of the objective function $u_i(x, x^*)$ used to define change probabilities for ministeps, referring to creation and/or maintenance of a tie $i \rightarrow j$, without being necessarily a difference $f_i(x^{\pm ij}) - f_i(x)$ of some function f_i (or similar with multiplication by Δ^+ or Δ^-).

Evaluation function is only about the result;
elementary effect can express the detailed process / step that leads to a given configuration.

Example : transTrip1 and transTrip2

transTrip1 (transitive closure)

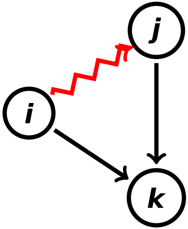
$$s_{ij}(x) = x_{ij} \sum_k x_{ik} x_{kj}$$



transTrip2

(structural equivalence outgoing ties)

$$s_{ij}(x) = x_{ij} \sum_k x_{ik} x_{jk}$$



Elementary effects can lead to the same configuration and therefore have the same target statistic
(such as transTrip1 and transTrip2).

In such cases they cannot be distinguished empirically by estimation by the Method of Moments.

Elementary effects can lead to the same configuration and therefore have the same target statistic
(such as transTrip1 and transTrip2).

In such cases they cannot be distinguished empirically
by estimation by the Method of Moments.

However, they can be be distinguished empirically
by estimation by the Generalized Method of Moments
(under development)
and by likelihood-based methods
(Maximum Likelihood, Bayes).

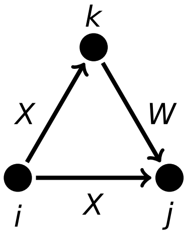
The use of elementary effects can give a more fine-grained representation of the process of network change;
but this will require more data;
like also distinction creation-maintenance requires more data.

Other example of elementary effects

XWX, ‘closure of covariate’.

E.g.: X = bullying, W = defending;
XWX = ’if *k* defends *j* and *i* bullies one of them, then s/he will tend to bully both’.

25. XWX1: like XWX, dependent variable is only one of the XWX ties: $i \rightarrow j$
‘*i* bullies those who are defended by his victims’.
26. XWX2: dependent variable here is $i \rightarrow k$.
‘*i* bullies defenders of his victims’.

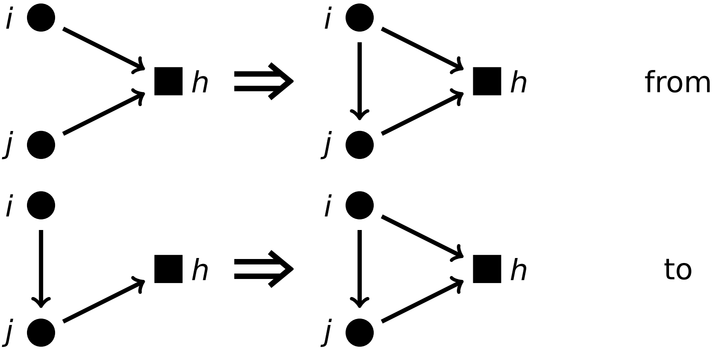


XWX1 and XWX2 are elementary effects.

Still other elementary effects

27. cl.XWX1: like XWX1 but for dependent network.
28. cl.XWX2: like XWX2 but for dependent network.

Influence in one-mode – two-mode co-evolution



Circles (left) are mode-1, squares (right) are mode-2 nodes.

Top: affiliation-based focal closure, effect *from*;

bottom: association-based affiliation closure, effect *to*.

There are a lot of other effects – see the manual! E.g.:

29. reciPop: reciprocal degree popularity
30. reciAct: reciprocal degree activity
31. gwesp.. effects have endowment and creation effects. They also are allowed to interact with other effects (interactionType = "dyadic") , straightforward because implemented as elementary effects.
32. And various others (e.g., interactions between networks and covariates).

6. Co-evolution

Evolution of multiple networks is studied more and more.

Various new effects have been constructed for this purpose: see Section 12.1.2 of the manual.

6. Co-evolution

Evolution of multiple networks is studied more and more.

Various new effects have been constructed for this purpose:
see Section 12.1.2 of the manual.

When a monadic or dyadic variable is regarded
as a control variable,
it still may be advisable to use it as a dependent variable
in the SAOM analysis, rather than as a covariate,
because this will allow the 'control' variable much better
to maintain its correspondence during the simulations
with the focal dependent variables.

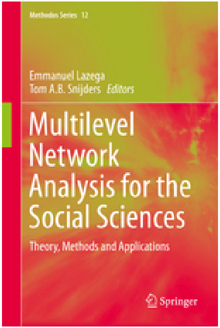
Results using a ‘control network’ as a covariate
will differ quite appreciably from results obtained
while using it as a co-evolving dependent network;
and similarly for monadic variables.

Example: acquaintance or communication
as a control network variable for advice
to study the properties of the ‘purified’ advice relation,
conditional on the condition of acquaintance.

7. Multilevel Analysis of Networks

See MultiMetaSAOM_s.pdf, at website.

Emmanuel Lazega and Tom A.B. Snijders (eds).
Multilevel Network Analysis
for the Social Sciences.
Cham: Springer, 2016.



Special issue of *Social Networks* ‘Multilevel Social Networks’, edited by Alessandro Lomi, Garry Robins, and Mark Tranmer, vol. 44 (January 2016).

Analysis of Multilevel Networks

Multilevel network (Wang, Robins, Pattison, Lazega, 2013):

Network with nodes of several types,
distinguishing between types of ties
according to types of nodes they connect.

Thus, if types of nodes are A, B, C,
distinguish between $A-A$, $B-B$, $C-C$ ties, etc., (*within-type*)
and between $A-B$, $A-C$, etc., ties (*between-type*).

Some may be networks of interest,
others may be fixed constraints,
still others may be non-existent or non-considered.

This generalizes two-mode networks
and multivariate one mode – two mode combinations.

See paper

Tom A.B. Snijders, Alessandro Lomi, and Vanina Torlò (2013).
A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice.
Social Networks, 35, 265-276;

Analysis of longitudinal multilevel networks in RSiena is possible by a trick (thanks to James Hollway).

Consider multilevel network with two node sets, A and B.

There are two one-mode networks internal to A and B,
and two two-mode networks X_1 from A to B; X_2 from B to A.

Specification for RSiena possible by employing
one joint node set $A \cup B$ and two dependent networks:

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} \text{internal } A & 0 \\ 0 & \text{internal } B \end{pmatrix} \\ \text{networks } A,\ B \end{array} \quad \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 0 & \text{two-mode } A \times B \\ \text{two-mode } B \times A & 0 \end{pmatrix} \\ \text{network } X_2 \qquad \text{network } X_1 \end{array}$$

For example:

A a set of organizations, B a set of individuals,
X_2 is a fixed membership relation, X_1 is not there;

networks A and B could be taken apart
in two distinct networks;

if there are only ties between individuals within organizations,
B will be a network of diagonal blocks
and structural zeros between different organizations;

if there are essential differences between individual ties
within organizations or across organizations,
B can be decomposed in two further distinct networks.

For the ‘Analysis of Multilevel Networks’ using RSiena, possibilities exist in principle, as indicated above;

a first example is Snijders, Lomi, Torlò (2013) mentioned above;

see scripts `RscriptSienaTwoModeAsOneMode.R` and `TwoModeAsSymmetricOneMode_Siena.R` on the Siena website

the research program is being continued by James Hollway (Oxford – Zürich – Genève) and by Gennady Zavyalov (Stavanger);

further relevant effects have to be elaborated;

and the field is open!

8. Missing Data in RSiena

The internal treatment of missing tie values in RSiena is simple:

- Impute missing tie variables in wave 1 by 0.
- Impute missing tie variables in later waves by Last Observation Carried Forward.
- Exclude these imputed values from the calculation of the statistics used for estimation in the MoM.

This can be improved if you have more knowledge of the data and also if you are willing to take more effort.

Missing Data: improvements

- ⇒ Sometimes there is enough information to make some imputations,
  based on knowledge of the data,
  with a high degree of confidence.
  If possible, do this!
- ⇒ There was an error in the treatment of missings in *non-centered* monadic covariates
  until and including version 1.1-284.

Missing Data (contd.)

⇒ New option *imputationValues* in *coCovar*, *varCovar* :
these values will be used for imputation of missings
for the simulations,
but (like always happens for missings) are not taken into account for the statistics used for estimation.

Can be used if there are reasonable, not completely reliable values for imputation.

Missing Data (contd. further)

⇒ Papers about treatment of missing data in *Social Networks* by Hipp, Wang, Butts, Jose, Lakon (2015) and Wang, Butts, Hipp, Jose, Lakon (2016) criticize missing data treatment by RSiena; but they disregard the fact that imputed values are not used for the statistics for estimation, only for simulations. Thus the effect of these imputations is only indirect.

⇒ In Wang et al. (2016) it is proposed to do multiple imputations by ERGMs for treating missing data in SAOMS. This might be an improvement of the current defaults, but it disregards the longitudinal dependence!

Intermezzo:
Multiple imputation – how does it work?

Multiple stochastic imputation was developed by Don Rubin.

For a given incomplete data set,
the missing data is imputed independently D times
by drawing from the conditional distribution
of the missing data given the observed data.

This leads to D complete data sets,
that differ only with respect to the imputed values.

Intermezzo:

Multiple imputation – how does it work?

Multiple stochastic imputation was developed by Don Rubin.

For a given incomplete data set,
the missing data is imputed independently *D* times
by drawing from the conditional distribution
of the missing data given the observed data.

This leads to *D* complete data sets,
that differ only with respect to the imputed values.

For each complete data set the desired analysis is executed;
standard errors of parameters are a combination
of the within-data set standard errors,
and the variability of estimates between the data sets.

How to combine the multiple imputations

The parameter of interest is denoted γ.

Suppose that the d'th randomly imputed data set leads to estimates $\hat{\gamma}_d$ and estimated variances W_d ('Within'),

$$W_d = \text{var}\left\{\hat{\gamma}_d \mid \text{ data set } d\right\} .$$

Note that W_d underestimates true uncertainty, because it treats imputed data as real data.

The combined estimate is the average

$$\bar{\gamma}_D = \frac{1}{D} \sum_{d=1}^{D} \hat{\gamma}_d .$$

Combine multiple imputations....

Compute the average within-imputation variance

$$\overline{W}_D = \frac{1}{D} \sum_{d=1}^{D} W_d \ ,$$

and the between-imputation variance

$$B_D = \frac{1}{D-1} \sum_{d=1}^{D} \left(\hat{\gamma}_d - \bar{\gamma}_D \right)^2 \ .$$

Estimated total variability for $\bar{\gamma}_D$ is

$$T_D = \widehat{\text{var}}\left(\bar{\gamma}_D \right) = \overline{W}_D + \frac{D+1}{D} B_D \ , \ s.e.\left(\bar{\gamma}_D \right) = \sqrt{T_D} \ .$$

Another kind of multiple imputation

The ML option in RSiena will give a model-based simulation of the missings in the second wave,
if the first wave has complete data.

This can be used for getting model-based longitudinal imputations:

1. If the first wave has any missings, estimate an ERGM and impute the missings in the first wave using this.
2. Estimate the SAOM parameters provisionally using the default treatment of missing data.

3. For each wave m, $m = 1, \ldots, M-1$:
   given the completed data set for wave m, produce a model-based random draw from the missings in wave $m+1$ from an ML simulation.
   This is not as time-consuming as full ML estimation, because only one simulation is required.
4. Use this complete data set to obtain one estimate $\hat{\gamma}_d$.
5. Repeat this procedure D times and use Rubin's rules for combining the estimates and standard errors.

The main disadvantage is that the future values are not used for the imputations.

This assumes 'missingness at random': i.e., observed data are sufficient for randomly generating missing data.

Example

Waves 2-3-4 of the van de Bunt students data.

Wave 0 is complete, so no ERGM imputation is needed!

Number of missing actors in waves 0–4 are
0; 2; 3; 5; 6, out of 32.

Impute wave 1 – then 2 – then 3 – then 4.

	default		multiple imputation		
Effect	par.	(s.e.)	par.	(s.e.)	m.f.
Rate 1	4.207	(0.640)			
Rate 2	5.063	(0.668)			
outdegree	–1.728***	(0.317)	–1.804***	(0.343)	.16
reciprocity	2.024***	(0.233)	2.100***	(0.260)	.18
trans. trip.	0.324***	(0.048)	0.329***	(0.049)	.12
indeg. - pop.	0.002	(0.038)	0.024	(0.039)	.16
outdeg. - pop.	–0.132***	(0.027)	–0.155***	(0.031)	.11
outdeg. - act.	0.014	(0.014)	0.013	(0.014)	.09
sex alter	0.409*	(0.200)	0.323	(0.204)	.08
sex ego	–0.386†	(0.208)	–0.282	(0.218)	.13
same sex	0.379*	(0.189)	0.362*	(0.193)	.07
program sim.	0.604**	(0.205)	0.687**	(0.213)	.09

par. = estimate; s.e. = standard error; m.f. = missing fraction;
† $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$;

convergence t ratios all < 0.06; overall maximum convergence ratio 0.08.

Note:
in waves 3 and 4 the proportion of missing actors is 0.15;
proportion missing information is of about this size.

Standard errors of the two approaches are similar;
estimates sometimes (3 cases) differ by about half s.e.,
in other cases differ hardly.

Further studies are needed to see
how this procedure performs.

9. Effect sizes ....?

Effect sizes aim for comparability of parameters.

Comparability may be across variables in the same data set, or between models for the same data set, or across data sets.

In linear regression comparability (all types) can be achieved by using standardized predictors, or equivalently by considering

$$\beta_k \,\text{s.d.}(X_k) \qquad \text{or} \qquad \frac{\beta_k \,\text{s.d.}(X_k)}{\text{s.d.}(Y)} \ .$$

For linear regression, there still is an essential difference between such standardized coefficients ($\sim$ one model) and contributions to R^2 ($\sim$ comparison between models).

The straightforward approach to effect sizes
in linear regression models
breaks down in most other generalized linear models.

Several approaches possible for defining effect sizes:

- ⇒ Marginal effects:
  expected change in some outcome variable,
  for a ‘unit’ change in explanatory ‘variable’;
- ⇒ Model-based effects:
  definitions within the model,
  making the effect sizes comparable in some way.

Marginal effects are more directly interpretable,
because they refer directly to observed variables;
they may be complicated because of their dependence
on the values of other variables in the model.

In December 2015 I presented some thoughts
on marginal effects for SAOMs.

Here I present some available techniques
for model-based effects for SAOMs.

Change statistics

Consider a SAOM with evaluation function

$$f_i(\beta, x) = \sum_k \beta_k s_{ki}(x) \ .$$

Define by $\delta_{ij,k}(x)$ the change statistic

$$\delta_{k,ij}(x) \ = \ s_{ki}\big(x^{(+ij)}\big) - s_{ki}\big(x^{(-ij)}\big) \ ,$$

where $x^{(+ij)}$ and $x^{(-ij)}$ are the networks x with and without tie $i \rightarrow j$, respectively.

Change statistics

Consider a SAOM with evaluation function

$$f_i(\beta, x) = \sum_k \beta_k s_{ki}(x) \, .$$

Define by $\delta_{ij,k}(x)$ the change statistic

$$\delta_{k,ij}(x) \; = \; s_{ki}\big(x^{(+ij)}\big) - s_{ki}\big(x^{(-ij)}\big) \, ,$$

where $x^{(+ij)}$ and $x^{(-ij)}$ are the networks x with and without tie $i \rightarrow j$, respectively.

For the model-based effect sizes,
we consider the probabilities in a ministep.

In a ministep, the probability of toggling tie variable x_{ij}, if actor i has the opportunity to make a change, is

$$\pi_{ij}(\beta, x) = \frac{\exp\big(f_i(\beta, x^{(\pm ij)}) - f_i(\beta, x)\big)}{1 + \sum_{h \neq i} \exp\big(f_i(\beta, x^{(\pm ih)}) - f_i(\beta, x)\big)} ,$$

where $x^{(\pm ij)}$ the network in which tie variable x_{ij} is toggled into $1 - x_{ij}$ (and similarly for $x^{(\pm ih)}$).

Note that

$$f_i(\beta, x^{(\pm ij)}) - f_i(\beta, x) = \sum_k \beta_k (1 - 2 x_{ij}) \delta_{ij,k}(x) .$$

The change statistics (with a + or a –) have the role of the explanatory ~~('independent')~~ variables.

For example:

$$\log\left(\frac{\mathrm{P}\{\text{add tie } i \rightarrow j \text{ to } x\}}{\mathrm{P}\{\text{leave network unchanged}\}}\right) = \sum_k \beta_k \, \delta_{ij,\,k}(x) \qquad \text{if } x_{ij} = 0 \ ;$$

$$\log\left(\frac{\mathrm{P}\{\text{drop tie } i \rightarrow j \text{ from } x\}}{\mathrm{P}\{\text{leave network unchanged}\}}\right) = -\sum_k \beta_k \, \delta_{ij,\,k}(x) \quad \text{if } x_{ij} = 1 \ ;$$

and

$$\log\left(\frac{\mathrm{P}\{\text{add tie } i \rightarrow j \text{ to } x\}}{\mathrm{P}\{\text{add tie } i \rightarrow h \text{ to } x\}}\right) = \sum_k \beta_k \big(\delta_{ij,\,k}(x) - \delta_{ih,\,k}(x)\big) \ ,$$

if x does not have the ties $i \rightarrow j$ or $i \rightarrow h$.

Question 1

Two-mode network for Glasgow data for 14 activities:

		daily	weekly	monthly	less
1	I listen to tapes or CDs	**388**	23	5	16
2	I look around in the shops	**65**	290	48	30
3	I read comics, mags or books	**186**	121	65	60
4	I go to sport matches	**30**	**113**	90	200
5	I take part in sports	**218**	117	30	68
6	I hang round in the streets	**216**	64	26	125
7	I play computer games	**157**	109	45	122
8	I spend time on my hobby (e.g. art, an instrument)	**114**	113	36	170
9	I go to something like B.B., Guides or Scouts	**36**	**81**	1	314
10	I go to cinema	**11**	**81**	269	71
11	I go to pop concerts, gigs	**7**	**6**	**92**	326
12	I go to church, mosque or temple	**2**	**52**	**10**	368
13	I look after a pet animal	**197**	25	6	203
14	I go to dance clubs or raves	**15**	**44**	**104**	266
15	I do nothing much (am bored)	37	39	24	331

Number of students participating in each of a list of activities, summed over three waves, for Glasgow data. Bold-faced are frequencies counted as a tie.

Effect	par.	(s.e.)
rate period 1	4.230	(0.268)
rate period 2	4.046	(0.278)
outdegree	–5.891***	(0.660)
outdegree - activity	0.637***	(0.088)
indegree - popularity ($\sqrt{}$)	0.790***	(0.100)
out-in degree assortativity	–0.0184***	(0.0025)
4-cycles	0.0389***	(0.0057)

Estimation results for activity participation of Glasgow students.

Question 1: what about the vastly different parameter sizes?

Some kind of standardization

To compare β_k for a given network x,
consider the formulae

$$\pi_{ij}(\beta, x) = \frac{\exp\Big(f_i(\beta, x^{(\pm ij)}) - f_i(\beta, x)\Big)}{1 + \sum_{h \neq i} \exp\Big(f_i(\beta, x^{(\pm ih)}) - f_i(\beta, x)\Big)} ,$$

$$f_i(\beta, x^{(\pm ij)}) - f_i(\beta, x) = \sum_k \beta_k \, (1 - 2\, x_{ij}) \, \delta_{ij,\,k}(x) \; .$$

This shows that β_k is multiplied by the ‘variable’

$$(1 - 2\, x_{ij}) \, \delta_{ij,\,k}(x) \; .$$

Define

$$\delta_{ii,\,k}(x) = 0 \; .$$

In analogy to linear regression,
we can make the values β_k comparable by considering

$$\sigma^2_{ik}(x) = \text{var}\Big\{\delta_{ij,k}(x)\,(1-2x_{ij}) \;\Big|\; i \text{ fixed }\Big\}$$
$$= \frac{1}{n}\sum_{j=1}^{n}\big(\delta_{ij,k}(x)\,(1-2x_{ij})\big)^2 - \Big(\frac{1}{n}\sum_{j=1}^{n}\delta_{ij,k}(x)\,(1-2x_{ij})\Big)^2 ,$$

the within-actor variance of this 'variable'; and

$$\sigma^2_k(x) = \frac{1}{n}\sum_i \sigma^2_{ik}(x) \, .$$

The product

$$\sigma_k(x)\,\beta_k$$

expresses the parameters β_k, for different k and a given x, on a common scale. The standard deviation is used here as a somewhat arbitrary, but well-known measure of scale.

Answer to Question 1

Effect	$\hat{\beta}_k$	(s.e.)	σ_k	$\sigma_k\,\hat{\beta}_k$
rate period 1	4.227	(0.268)		
rate period 2	4.047	(0.278)		
outdegree	–5.891	(0.660)	0.79	–4.67
outdegree - activity	0.637	(0.088)	5.48	4.33
indegree - popularity ($\surd$)	0.790	(0.100)	6.74	4.29
out-in degree assortativity	–0.0184	(0.0025)	331.54	–6.11
4-cycles	0.0389	(0.0057)	55.08	2.14

$\hat{\beta}_k$: estimates; s.e.: standard errors;
σ_k: mean within-ego standard deviations of change statistics;
$\sigma_k\hat{\beta}_k$: their product.

The order of magnitude of the effects is similar.

Question 2

Effect	par.	(s.e.)
rate (period 1)	14.471	(1.511)
rate (period 2)	11.949	(1.250)
outdegree	–3.419	(0.294)
reciprocity	3.815	(0.315)
GWESP transitive	1.906	(0.125)
GWESP cyclic	0.394	(0.114)
indegree - popularity	–0.008	(0.027)
outdegree - popularity	–0.092	(0.063)
outdegree - activity	0.110	(0.047)
reciprocated degree - activity	–0.279	(0.066)
sex alter	–0.156	(0.091)
sex ego	0.057	(0.106)
same sex	0.591	(0.082)
reciprocity × GWESP transitive	–1.148	(0.167)

How strong are the contributions of the effects to determining the actors' choice probabilities?

Relative Importance

Indlekofer & Brandes (*Network Science*, 2013) proposed a measure for the *Relative Importance* of effects.

This is based on the following approach:
how large is the change in the probability vector π_i
if one of the effects is dropped?

Ingredients of the approach:

- Just replace β_k by 0, do not re-estimate.
- Measure change in probability vector by the ℓ_1 distance, i.e., sum of absolute differences.

Define π_i as vector of probabilities for actor i in next ministep, and $\pi_i^{(-k)}$ as the same if effect k obtains a weight of 0.

Then *importance of effect k for actor i* is

$$\left\|\pi_i - \pi_i^{(-k)}\right\|_1 = \sum_j \left|\pi_{ij} - \pi_{ij}^{(-k)}\right| .$$

Relative importance of effect k for actor i is

$$I_k(X, i) = \frac{\left\|\pi_i - \pi_i^{(-k)}\right\|_1}{\sum_{\ell=1}^{K} \left\|\pi_i - \pi_i^{(-\ell)}\right\|_1} ;$$

Expected relative importance (for random *i*) is

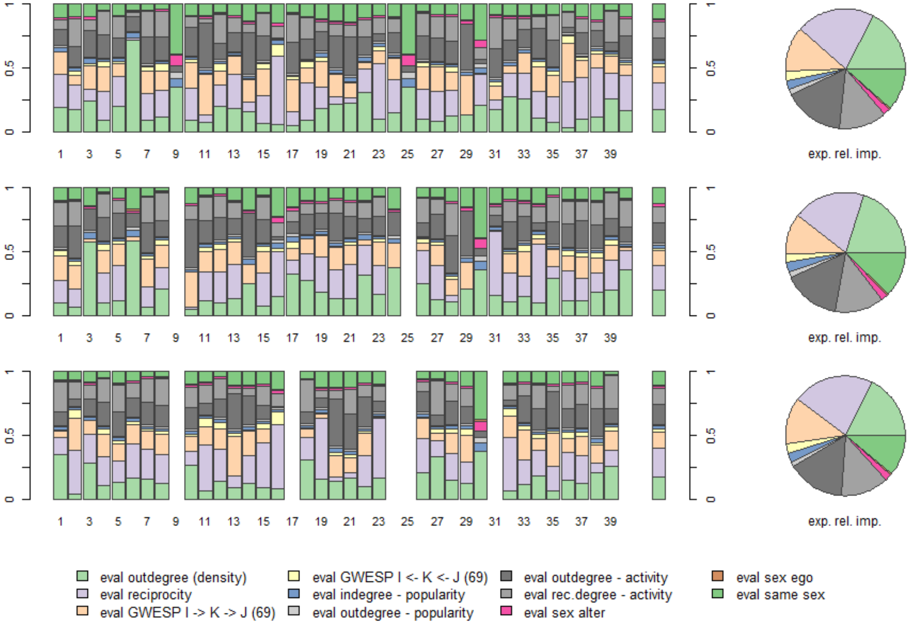
$$\frac{1}{N} \sum_{i=1}^{N} I_k(X, i) .$$

Expected (raw / total) importance can then be defined as

$$\frac{1}{N} \sum_{i=1}^{N} \left\|\pi_i - \pi_i^{(-k)}\right\|_1 .$$

Answer to Question 2

Effect	$\hat{\beta}_k$	(s.e.)	Exp. rel. importance wave 1	wave 2	wave 3
outdegree	-3.5904	(0.2161)	0.1730	0.2005	0.1749
reciprocity	3.5827	(0.2503)	0.2115	0.1952	0.2209
GWESP transitive	1.5569	(0.0997)	0.1225	0.1085	0.1267
GWESP cyclic	0.3290	(0.1178)	0.0252	0.0216	0.0253
indegree - popularity	-0.0274	(0.0247)	0.0245	0.0266	0.0232
outdegree - popularity	-0.0197	(0.0453)	0.0148	0.0155	0.0141
outdegree - activity	0.1626	(0.0323)	0.1626	0.1549	0.1534
reciprocated degree - activity	-0.3915	(0.0560)	0.1283	0.1303	0.1306
sex alter	-0.1458	(0.0886)	0.0179	0.0191	0.0172
sex ego	0.0597	(0.0966)	0.0049	0.0054	0.0047
same sex	0.6515	(0.0782)	0.1147	0.1224	0.1091


0
0.5
1
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39
exp. rel. imp.
eval outdegree (density)
eval reciprocity
eval GWESP I -> K -> J (69)
eval GWESP I <- K <- J (69)
eval indegree - popularity
eval outdegree - popularity
eval outdegree - activity
eval rec.degree - activity
eval sex alter
eval sex ego
eval same sex

Previous page:
Plot of relative importance of effects for first 40 actors (blank = absent) and averaged for all actors (pie-chart).

The graph was produced by

```
plot(RI, actors=1:40, addPieChart = TRUE, legendColumns=4)
```

where RI was the object produced by `sienaRI()`.

Question 3

In the model for network dynamics (Glasgow data), how large are the separate contributions of:

1. dyadic effects (reciprocity)
2. triadic effects (gwesp)
3. degree effects (inPop, outPop, outAct, reciAct)
4. effects of sex (ego, alter, same)?

Entropy-based measures

The uncertainty / variability
in the outcomes of categorical random variables
can be expressed by the *entropy* (Shannon, 1948).

For a probability vector $p = (p_1, \ldots, p_K)$, entropy is defined as

$$H(p) = -\sum_{k=1}^{K} p_k \, {}^2\log(p_k) \ . \qquad (1)$$

Minimum 0 (if one category has $p_k = 1$);
maximum ${}^2\log(K)$ (if $p_k = 1/K$ for all k).

The degree of certainty, or amount of information,
in the outcome of the ministep for actor i can be expressed by

$$R_H(i, \beta, x) = 1 - \frac{H(\pi_i(\beta, x))}{{}^2\log(K)} \ . \qquad (2)$$

For models with constant rate function, this can be averaged:

$$R_H(\beta, x) = \frac{1}{n} \sum_i R_H(i, \beta, x) \; . \tag{3}$$

This is a measure between 0 and 1:
1 if the outcome of the ministep for a given actor is certain;
0 if all actors choose a random change (all probabilities $1/n$).

This measure was proposed in Snijders
(*Mathématiques et Sciences Humaines*, 2004).

For network panel data we may average over waves:

$$R_H(\beta) = \frac{1}{M} \sum_m R_H\big(\beta, x(t_m)\big) \; . \tag{4}$$

Values generally will be low!

Contributions to R_H

For the contribution of an effect, or set of effects,
to the information about the outcome:
estimate the model twice,
giving parameter estimates $\hat{\beta}_1$ for full model
and $\hat{\beta}_2$ for restricted model, and consider the difference

$$R_H(\hat{\beta}_1) - R_H(\hat{\beta}_2) \ . \tag{5}$$

Note that this is not necessarily positive,
because the estimation method does not maximize $R_H(\hat{\beta})$.

Answer to Question 3

Effect	Exp. information		
	wave 1	wave 2	wave 3
Reciprocity only	0.2540	0.2475	0.2326
Reciprocity and triadic	0.2410	0.2596	0.2459
Reciprocity and degree-based	0.2307	0.2613	0.2172
Reciprocity and sex	0.2677	0.2574	0.2467
Rec, triadic, degree-based	0.2571	0.2842	0.2725
Rec, triadic, sex	0.2509	0.2635	0.2599
Rec, degree-based, sex	0.2465	0.2714	0.2342
Full model	0.2596	0.2854	0.2811

I have to check this....

sienaRI

Natalie Indlekofer has contributed the function `sienaRI()`, which assesses the relative importance of effects.
Available since version 1.1-270.

There was also the function `sienaRIDynamics()`, averaging over all changes from one wave to the next, but this had difficulties and was withdrawn (may be revived).

`sienaRI()` was extended by Tom Snijders with measures σ and R_H and further adapted.

10a. Developments in current models

There still is much more to do and explore within the confines of what has already been developed and implemented.

1. The topics mentioned above are open for application / elaboration.
2. Evaluation / creation / maintenance / elementary effects
3. Evaluation / creation / maintenance / effects for behaviour
4. Variants of non-directed models.
5. Comparability of effects across models, data sets
   $\sim$ 'marginal' effects

Developments in current models (contd.)

6. Model selection
7. Importance of GoF for validity of results
8. Extended auxiliary functions for GoF
9. avAlt $\Leftrightarrow$ avSim $\Leftrightarrow$ totAlt $\Leftrightarrow$ totSim
10. Diffusion of innovations – event history analysis
11. Two-mode networks
12. Multivariate (e.g., signed) networks
13. Ordered networks

9b. Hot Issues

- Analysis of Multilevel Networks (see above!)
- Comparison SAOM $\longleftrightarrow$ ERGM (Per Block et al)
- JSiena (Felix Schönenberger)
- Generalized Method of Moments (Viviana Amati)
- Continuous dependent actor variables (Nynke Niezink)
- Settings model (Tom Snijders)
- Marginal effects
- Stable standard errors (Nynke Niezink)
- CUP Books!


The horse race of Siena,
il palio, was originated in year 1581