ANALYSING DYNAMICS OF NON-DIRECTED SOCIAL NETWORKS

Tom A.B. Snijders
University of Oxford

May 2008

Statistical inference for network dynamics

Data:

2 or more repeated observations
on a network between a fixed set of n actors
where ties are non-directed = two-sided,
and refer to a longer-lasting state
of the relationship between the two actors.

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

This may be called *network panel data*.

Purpose of the analysis:

Statistical inference –
fit a plausible model to the repeated network data;
estimate and test parameters that express
the 'rules' governing the dynamics.

This will be done through assuming
a flexible class of stochastic agent-based models
as the processes that represent the network evolution.

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 can also be considered.

In a dynamic model, ‘existence’ is split up
into ‘creation’ and ‘continuation’.

Basic assumptions (Holland & Leinhardt, 1977).

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.

Thus, network dynamics treated
as an *endogenous dynamic process* with an inbuilt inertia.

Current state of the changing network acts as a
dynamic constraint on each actor's behavior.

The change process is decomposed in the smallest possible steps.
No coordination between actors – only reaction to current network.

Further assumption 1: initial state

Nothing is assumed about the initial state.

No equilibrium or stochastic stationarity assumed.
Only the transition probabilities are stationary.

In the agent-based simulation model,
the first observation is taken as the initial condition
$\sim$ the first observation is conditioned upon.

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 ('pay-off') 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:

A. *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.

B. *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:
A. for actor i, opportunities occur at a rate λ_i
B. 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 $r\,dt$.)

The rate functions λ_i 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 A, 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 B, the rate function is an upper limit to
expected number of changes *per dyad* per unit of time.

2. Decisions about changing ties

When there is an opportunity for change,
actors decide on changes in their ties depending on
preferences – costs – 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.

The actors maximize objective function + random disturbance.

Differences in objectives between actors are allowed
only when these can be captured by measured covariates.

There may be asymmetries
between the creation and dissolution of ties:
complement utility function by *endowment function* $g_i(x)$
which contributes only to the value of changes $1 \Rightarrow 0$ (dissolution)
and not to changes $0 \Rightarrow 1$ (creation).

Specification of objective function discussed below.

Different ways for combining actors' objectives

1. Unilateral imposition of a tie (disjunctive).
2. Mutual agreement required for a tie to exist (conjunctive).
3. Gain for one may outweigh loss for the other (compensatory).

This is to be combined with
A: unilateral initiative; B: two-sided opportunity.

Combination A-3 (one-sided proposal & compensating objectives) is possible but less natural.

A1. 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.

A2. 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.

B1. 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 $X_{ij} = 0$.

B2. 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 $X_{ij} = 0$.

B3. 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.

Other possibilities can be thought of:
e.g., agreement necessary to make a change

In models A, when actor i gets the initiative,
she must choose which tie variable to change;
call the possibly changed tie variable x_{ij}.

The hypothetical new network obtained by changing x_{ij}
is denoted by $x(i \rightsquigarrow j)$.
Formally, let $j = i$ denote that nothing changes
(the current situation is the best): $x(i \rightsquigarrow i) = x$.
Actor i chooses the "best" j by maximizing

$$f_i\Big(\beta, x(i \rightsquigarrow j)\Big) \ + \ U_i(t, x, j)\,.$$

$\Uparrow$

random component

If the random component U has a
type 1 extreme value = Gumbel distribution,
the probability that i chooses j is

$$p_{ij}^{(i1)}(\beta, x) = \frac{\exp\big(f(i,j)\big)}{\sum_{h=1}^{n} \exp\big(f(i,h)\big)}$$

where

$$f(i,h) = f_i\big(\beta, x(i \leadsto h)\big) \ .$$

This is the multinomial logit form of a *random utility* model.

The Gumbel distribution has variance $\pi^2/6 = 1.645$ and s.d. 1.28.

Yes/no decisions occur in Models B,
and for the agreement by the second actor in model A2.
These are based on the comparison by the actor
of the network $x^{+}(i,j)$ with tie $i-j$,
and the network $x^{-}(i,j)$ without this tie.
The comparison is made similar as before.

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

$$p_{ij}^{(j2)}(\beta,x) = \frac{\exp\Big(f_j(\beta,x^{+}(i,j))\Big)}{\exp\Big(f_j(\beta,x^{-}(i,j))\Big) + \exp\Big(f_j(\beta,x^{+}(i,j))\Big)}$$

A possible extension of model A2 is to give i the ability to think one step in advance: 'will j accept?',
so that i chooses to change x_{ij}
for the j that maximizes

$$\Big(f_i\big(\beta, x(i \rightsquigarrow j)\big) + U_i(t, x, j) \Big)\, p_{ij}^{(j2)}(\beta, x)$$

or

$$\Big(f_i\big(\beta, x(i \rightsquigarrow j)\big) \Big)\, p_{ij}^{(j2)}(\beta, x) + U_i(t, x, j) \ .$$

Model specification :

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

Covariates can be actor-dependent: v_i
or dyad-dependent: w_{ij} .

Convenient definition of f_i is a weighted sum

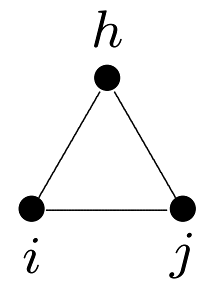
$$f_i(\beta, x) = \sum_{k=1}^{L} \beta_k \, s_{ik}(x) \, ,$$

where weights β_k are statistical parameters
indicating strength of 'effect' $s_{ik}(x)$;
these effects may depend also on v_i and w_{ij} .

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

1. *out-degree effect*, number of friends
   $s_{i1}(x) = x_{i+} = \sum_j x_{ij}$
2. *popularity/activity effect*, sum of degrees of i's friends
   $s_{i2}(x) = \sum_j x_{ij}\, x_{+j} = \sum_j x_{ij} \sum_h x_{hj}$
3. *transitive triads effect*,

   number of transitive patterns in i's ties
   $(i - j,\ j - h,\ i - h)$
   $s_{i3}(x) = \sum_{j,h}\ x_{ij}\, x_{jh}\, x_{ih}$



transitive triad

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

Many other network effects are possible.

Two basic kinds of effect associated with actor covariate v_i :

5. *covariate-related degrees*,
   $s_{i5}(x) = v_i \, x_{i+}$;
   positive effect contributes
   to correlation between degrees and V.

6. *covariate-related similarity*,
   sum of measure of covariate similarity
   between i and his friends, e.g.
   $s_{i6}(x) = \sum_j x_{ij} \Big(1 - |v_i - v_j|\Big)$
   if V has values between 0 and 1;
   positive effect contributes to network autocorrelation of V.

Basic effect for dyadic covariate w_{ij} :

7. *covariate-related preference*,
   values of w_{ij} summed over all others to whom i is tied,
   $s_{i7}(x) = \sum_j x_{ij} \, w_{ij}$;
   positive effect contributes to correlation
   between X_{ij} and W_{ij} .

This model definition, with components

A / B. individual – dyadic initiative, with rate function λ_i

1 / 2 / 3. disjunctive – conjunctive – compensatory decisions
based on objective function $f_i(x)$, endowment function $g_i(x)$

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

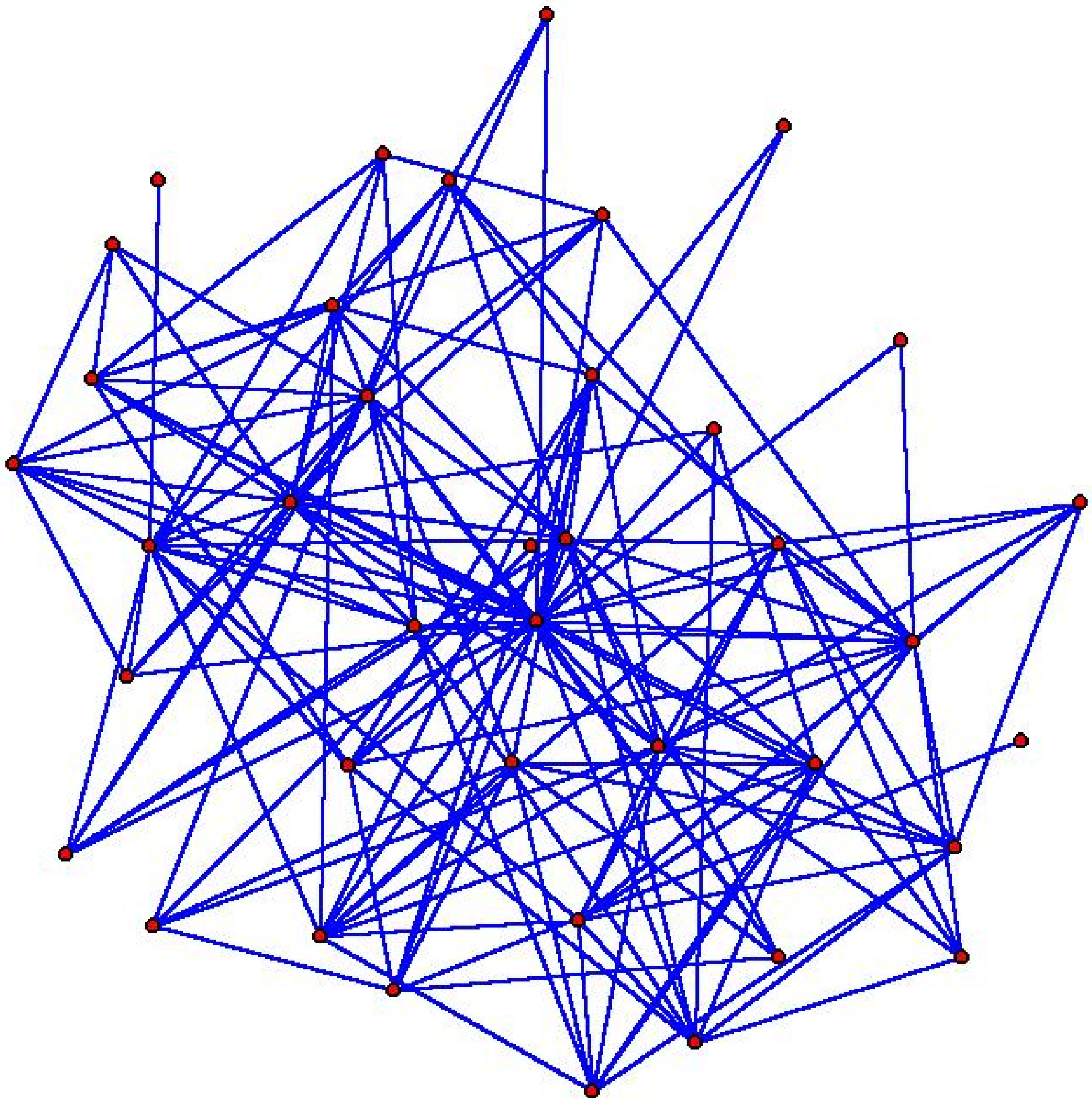
Parameters can be estimated using Method of Moments estimators (i.e.: equate expected statistics to observed statistics) implemented by stochastic approximation as in Snijders (2001).

Example: Assistance Relation in Kapferer's Taylor Shop

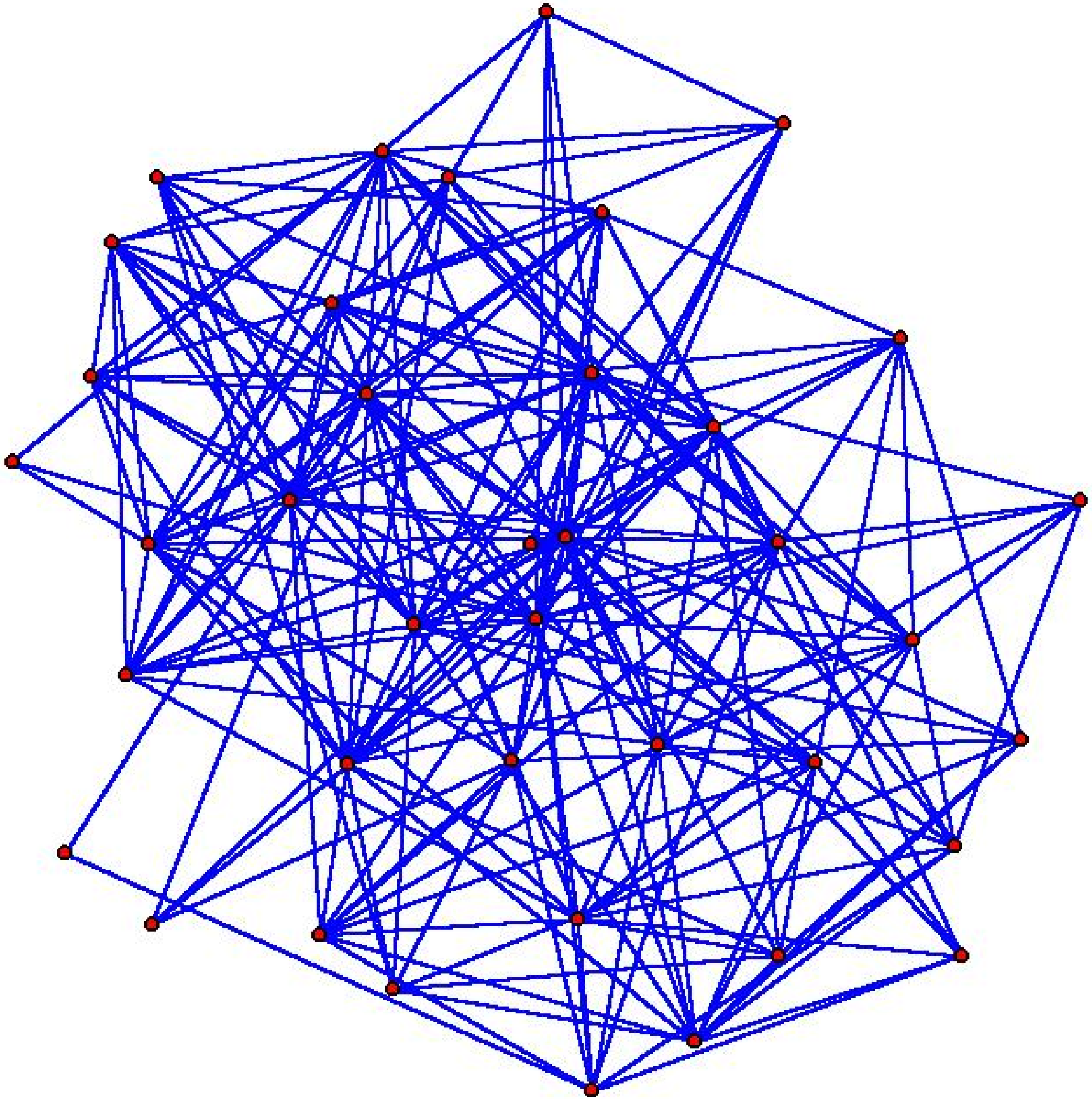
Kapferer (1972):
Interactions in tailor shop in Zambia, period of 10 months, within which there was an abortive strike.

Work- and assistance-related relation.

Dichotomous covariate *status*:
contrast { head tailor (nr 19), cutter (16),
line 1 tailor $(1–3, 5–7, 9, 11–14, 21, 24)$, button machiner $(25–26)$}
to
{ line 3 tailor $(8, 15, 20, 22 – 23, 27 – 28)$, ironer $(29, 33, 39)$,
cotton boy $(30 – 32, 34 – 38)$, line 2 tailor $(4, 10, 17 – 18)$ } .



Assistance network
time 1.
Average degree 8.3



Assistance network
time 2.
Average degree 11.7

Preliminary (trial) analyses yielded the following suggestions:

Rate function:

higher status workers change ties more often:

$$\lambda_i = \lambda_0 \, e^{v_i}$$

Objective function:

status homophily

transitivity, as expressed by preference for transitive triads

No evidence for other components
in objective function, or for endowment function.

(This holds for all model specifications.)

Models with individual initiative

	Model 1 (forcing)		Model 2 (agreement)	
Effect	*par.*	*(s.e.)*	*par.*	*(s.e.)*
Rate	7.60	1.10	12.09	1.51
Status ⇒ rate	1.32	0.39	1.27	0.36
Degree	−1.08	(0.13)	−0.55	0.11
Transitive triads	0.22	(0.03)	0.17	0.03
Status similarity	0.42	(0.11)	0.33	0.09

Models with two-sided initiative

	Model 3 (disjunctive)		Model 4 (conjunctive)		Model 5 (compensatory)	
Effect	*par.*	*(s.e.)*	*par.*	*(s.e.)*	*par.*	*(s.e.)*
Rate	0.88	0.06	0.79	0.06	0.79	0.05
Status ⇒ rate	0.70	0.21	0.71	0.19	0.69	0.20
Degree	−0.85	0.17	−2.03	(0.24)	−1.01	0.16
Transitive triads	0.28	0.06	0.43	(0.08)	0.21	0.07
Status similarity	0.48	0.15	0.73	(0.23)	0.37	0.11

Conclusions:

Results similar in all five models;
higher-status workers change their ties more often;
status-related homophily, transitivity.

Rate functions higher for models with individual initiative
(follows from different definition:
A: rate per actor, B: rate per pair).

Out-degree effect higher when more agreement is required,
because this tends in itself to depress number of ties.

Effects are somewhat different
but transitive triads and status similarity effects
are roughly proportional.

Different model specification:

include equal job titles as additional dyadic covariate

Models with individual initiative

	Model 1 (forcing)		Model 2 (agreement)	
Effect	*par.*	*(s.e.)*	*par.*	*(s.e.)*
Rate	8.04	1.04	13.02	1.67
Status ⇒ rate	1.16	0.38	1.09	0.41
Degree	−1.08	(0.11)	−0.54	0.10
Transitive triads	0.21	(0.03)	0.16	0.03
Status similarity	0.24	(0.13)	0.19	0.09
Same job title	0.46	(0.16)	0.36	0.13

Models with two-sided initiative

	Model 3 (disjunctive)		Model 4 (conjunctive)		Model 5 (compensatory)	
Effect	*par.*	*(s.e.)*	*par.*	*(s.e.)*	*par.*	*(s.e.)*
Rate	0.92	0.06	0.82	0.06	0.81	0.06
Status ⇒ rate	0.58	0.21	0.60	0.21	0.59	0.23
Degree	−0.85	(0.16)	−1.98	0.21	−1.00	0.17
Transitive triads	0.26	(0.06)	0.38	0.07	0.19	0.08
Status similarity	0.28	(0.16)	0.42	0.31	0.21	0.11
Same job title	0.65	(0.22)	0.90	0.25	0.45	0.17

Conclusions from extended model specifications:

Having the same job title is clearly significant,
when controlling for status similarity
(dichotomized status);

status similarity is not quite, or just, significant
when controlling for having same job title,
depending on precise model specification.

Parameter for transitive triads is still significant,
but smaller than in earlier model:
job title represents a bigger portion of the observed transitivity.

Summary / Discussion

This method extends the methodology
for analysing network dynamics of Snijders (*Soc. Meth.* 2001)
to two-sided non-directed relations.

(For directed relations,
a two-sided decision can also be appropriate!)

The methods are included in the program SIENA

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

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.

In the example, conclusions were similar across these models,
but not quite the same.
Similarity of conclusions obviously
saves us some practical complications.

If the timing of tie changes is observed,
then a direct maximum likelihood estimation is possible.