Stochastic actor-based models in network dynamics

Robert Stainforth[1] Stephan Onggo[2]

[1]Maths & Physics, University of Durham; [2]Management Science, University of Lancaster

Introduction

Networks are a collection of interconnected entities whose ties to one another are maintained by actively interacting. Examples include financial deals between companies and groups of people sharing a common relation such as friendship or occupation.

Many of us are part of a network. Indeed, some will be part of more than one. Others will leave networks and join new ones. In doing so people decide about who and what they associate with - they are *actors* choosing to make ties with other actors.

Using longitudinal data [*OECD*] on the migration of people between 13 EU countries between 2000-2009 a set of parameters each describing how a specific aspect of the network evolves may be estimated.

Morph: A popular (clay-)model based actor.

How can a network be constructed?

For a given point in time, t_k, a network of n actors and the ties between them may be represented by a $n \times n$ matrix; A_k.

Ties in a network are directional, known as *out-degrees* or *in-degrees* they are initiated by one actor onto another. The matrix elements; a_{ij} of A_k reflect this:

- For a tie from actor-i to actor-j: $a_{ij} = 1$. (out-degree of i/in-degree of j)
- For a non-tie from actor-i to actor-j: $a_{ij} = 0$. However a tie from actor-j to actor-i may still exist ($a_{ji} = 1$).
- If both $a_{ij} = a_{ji} = 1$ the tie is *reciprocal*.
- $a_{ii} = 0$, $\forall\ i \in [1, n]$, $\mathbb{Z}$.

As an example, consider a network of 4 actors:

$$A_k = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Each row of A_k represents an actor and the ties (if any) it shares with other actors in the network as seen below.

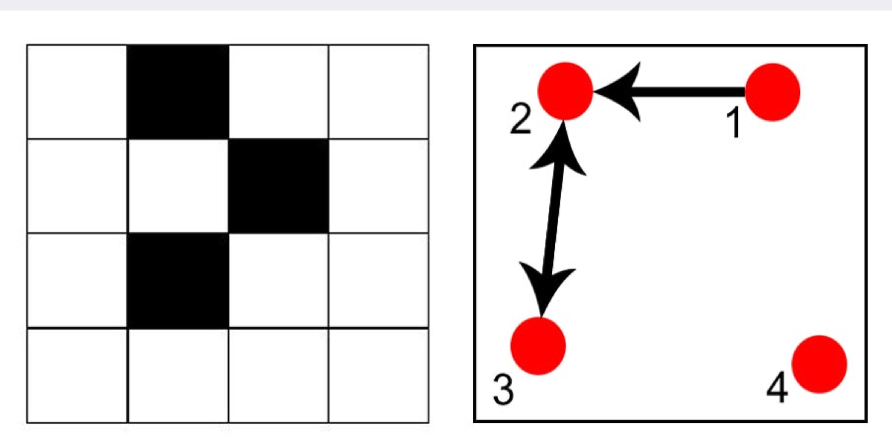

Figure 1: Sociomatrix and Network representations of A_k

How does a network evolve?

Different actors in a network will observe different changes over the same time interval. Consequently one actor is chosen deterministically or at random to play the role of the *ego*; the focal point about which network change is observed.

At an observed time an actor may choose to change one outgoing tie or do nothing. These choices have probabilities determined by the *objective function*, $f_i(\beta, \chi)$. This is a function of the current network, χ, as seen by the ego:

$$f_i(\beta, \chi) = \sum_k \beta_k s_{ki}(\chi) \qquad (1)$$

$s_{ki}(\chi)$ are functions of the network as viewed by actor-i known as *effects*. The β_k terms are the associated weight parameters representing the *tendency* of an actor towards ($\beta_k > 0$) or away ($\beta_k < 0$) from that effect.

If actor-i has the opportunity to make a tie between themselves and one of two other actors, different outcomes of the network can occur, χ_a or χ_b. $f_i(\beta, \chi_b) - f_i(\beta, \chi_a)$ is the *log odds ratio* for choosing between these two outcomes, the ratio of probability between χ_a and χ_b is then:

$$exp(f_i(\beta, \chi_b) - f_i(\beta, \chi_a)) = \frac{exp(f_i(\beta, \chi_b))}{exp(f_i(\beta, \chi_a))} \qquad (2)$$

[*T.A.B. Snijders et al., 2010*]

Effects

Effects are observed patterns in network *change*. The choice of which effects to include in a model (and hence which values of β_k) is completely ad-hoc and is down to study interest and subject area. The effects chosen for this study are:

- *indegree - popularity*: tendency for actors with high in-degrees to attract more in-degrees.
- *outdegree - activity*: tendency for actors with high out-degrees to send out further out-degrees.
- *out-in degree assortativity*: tendency for actors with high out-degrees to preferably tie themselves to actors with high in-degrees. [*R.M. Ripley et al., 2011*]

Study: EU Migration

- Data on the migration of people between 13 EU countries from 2000-2009 was collected (CZ, DK, FI, FR, DE, IE, NL, PO, PT, SK, SE, ES, UK) $\Rightarrow$ ten 13x13 matrices.
- *Dyadic covariates* (number of people) and *constant actor covariates* (average wage) were added to each tie and *actor*($\equiv$country) in the network respectively.
- Effects were included using the R package, *RSiena*. RSiena estimates values of β_k for each effect using a stochastic approximation algorithm.

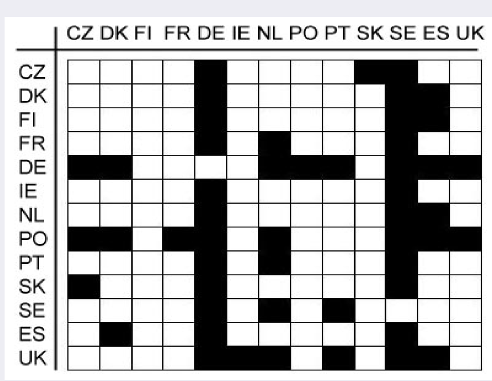

Figure 2: Sociomatrix of the migration network in 2007

Analysis & Results

Due to the ad-hoc nature of the model RSiena returns *p-values* for each effect, p-values vary between zero to one.

- For $p < 0.001$ the effect is significant and should be *included* in the model and hence the *objective function*.
- For $p \gg 0.001$ the effect should be *excluded* from the model.
- For $p \in [0.1, 0.3]$ the effect may be of *some* significance to the model.

The data exhibits a good density of ties between actors over the ten year period. (avg. of 3.443 degrees per country)

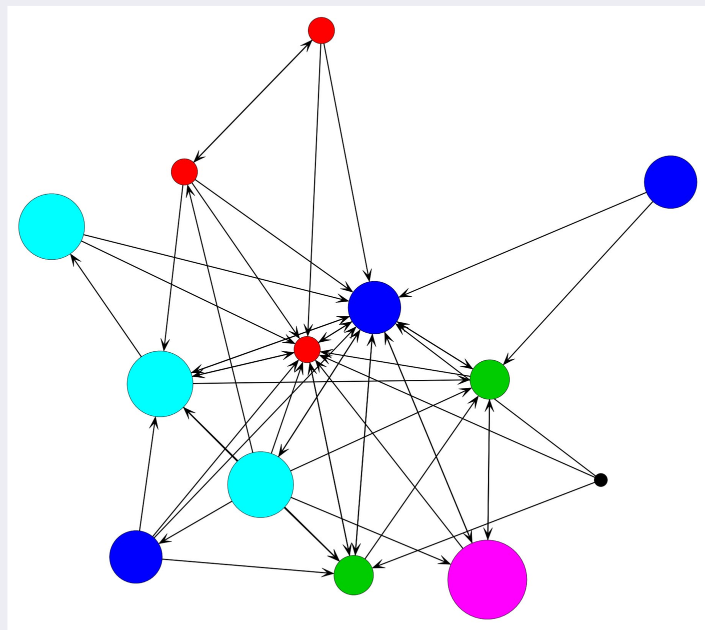

Figure 3: Network of migration between countries in 2006. Different average income brackets of the countries is represented by both the size and color of each node. Larger nodes indicate a larger income.

Effect	Value of β_k	Std. Error	p-value
indegree-pop.	1.3935	0.8520	0.5619
outdegree-actv.	0.6931	0.8824	0.2993
out-in degree-assrt.	-0.5966	0.5969	0.8433

Range of ties 2000-2009: (min, max) = (40,55) = 15

Figure 4: Table of values for effects and associated β_k values

The p-values shown in fig.4 suggest that none of the three studied effects are pivotal in influencing migration. However, an observed p-value of 0.2993 on the outdegree-activity effect suggests that the rate at which people leave a country has a stronger bearing in influencing migration than the popularity of the countries they migrate to:- the indegree popularity effect (p-value = 0.5619).

Conclusion

- All data was collected between the years 2000-2009.
- The data exhibits a good density of ties between actors.
- Too few changes between time intervals to observe *significant* effects.
- Influence of effects can only be inferred.

Improvements to the study would include:

- Data sets representing networks of actors that regularly make/break ties.
- RSiena is under constant development with effects regularly being added. Future effects may better describe migration with more accuracy.

References

OECD Statistics. International migration by nationality (2000-2009).
http://www.oecd-ilibrary.org

T.A.B. Snijders et al. (2010).
Introduction to stochastic actor-based models for network dynamics.

R.M. Ripley et al. (2011).
Manual for SIENA version 4.0 (version May 1, 2011).

http://www.robertstainforth.co.uk r.p.f.stainforth@durham.ac.uk