Using RSiena from within visone

Natalie Indlekofer, Uwe Nagel

June 28, 2010

Longitudinal networks can be statistically analyzed by using RSiena from within visone. Supported is the analysis of one-mode networks with a fixed set of actors. Not supported are two-mode networks, multiple networks, multilevel (multi-group) networks, composition change of actors, and multi-parameter tests.

Contents

1 Loading Longitudinal Network Data

For using RSiena from within ᵛisoₙₑ you need a network collection (cf. Section 3), i.e., a sequence of networks that build a longitudinal network. There are two ways for defining a network collection. First, by loading an RSiena session file[1] and second, by loading data components separately as is usual in ᵛisoₙₑ.

1.1 Loading an RSiena Session File

Please assure first that your R connection works (see Section 6).

By pressing the `open` button in the tool bar or menu, you reach the `open file` dialog. Select as type of file `RSiena session files`, allowed extensions are .txt, .csv, .dat, .prn. Choose your file, and press `ok`.

After setting a project name the data is loaded as specified in the file. Digraph data files are represented as ᵛisoₙₑ networks and linked to a new network collection (3) named by the given project name. All types of actor attributes and dyadic covariates are represented as node attributes (5.1) or dyad attributes (5.2), respectively.

1.2 Loading Network Components Separately

You can also load components of your longitudinal network data separately. We assume that you have loaded all your networks in ᵛisoₙₑ. For analyzing your single network observations as one longitudinal network, it is necessary to put them together to a network collection as described in Section 3. All types of actor attributes and dyadic covariates have to be loaded separately for each network as node attributes (5.1) and dyad attributes (5.2).

2 Analyze Longitudinal Networks with RSiena

The siena functionalities can be accessed by choosing the `task` "siena" on the `analyis` tab.

Here, all actor and dyadic attributes that are available for the active network collection (cf. Section 3) are listed. Specify for all attributes whether they should be regarded for analysis and if, as which type (e.g., constant or changing covariate).

The available types for each attribute depend on the given data. A decimal node attribute (5.1) for example, cannot be used as a siena behavior attribute without conversion into an integer. Further, no node attribute can be used as a changing covariate if a network collection contains only two observations, (see [1, 2]).

---

[1]same format as saved by using `Save to file` in the RSiena GUI or needed for loading files by `Load new session from file`, see [1]

2.1 Model Specification

To specify your model, use the `specify model`-button on the siena tab. A dialog will open, listing all effects that are available for the active network collection and the specified attributes. The table on the right contains the effects that are currently included in the model. Select effects for the analysis by adding or removing them to or from the table. For the estimation process, you can choose whether standard values or current parameter values are used as initial values by ticking the corresponding checkbox. You can also set the number of subphases in phase 2 and the number of iterations in phase 3 (see [1]) with the corresponding sliders.

2.2 Estimation

Start the parameter estimation with the `estimate` button. The estimation progress can be monitored in the Rserve (see 6) window (console, Windows only). When the estimation is finished, results are displayed in the right table of the model specification dialog. For each included effect its estimate, associated standard error and t-statistic (indicating convergence of the estimation process, see [1]) are given. The p-values assume as null hypothesis that corresponding parameter values are 0. They are calculated by the R command `2*pnorm(-abs(parameter estimates/standard errors))`.
Fixing and testing of certain effects is possible by ticking the corresponding checkboxes. Shown p-values of tested effects result from the score-type test as described in [1]. p-values here assume as null hypothesis that parameter values are as given in the specification table. Note, that null hypothesis with parameters set to 0 can be tested by setting the appropriate values in the table to 0.
In the current version it is necessary to confirm value changes in the table by pressing `enter`.
On close of the model specification dialog, ^{v}ison$_e$ asks whether simulation results are to be saved. The resulting file contains the usual RSiena output for estimation results.

2.3 Simulation

The simulation of a number of networks with the current parameter estimates can be started with the `simulate` button. The number of simulations is the number of iterations in phase 3 as set by the user. For each pair of actors a link probability is calculated as average number of simulated networks in which they are linked. Resulting tie probabilities are saved as dyad attribute (see 5.2) with the name `tie probabilities`. Again all results can be saved when the dialog is closed and the resulting file contains the usual RSiena output for estimation and simulation.

3 Network Collections

A network collection is a set of networks with an additional order, e.g., a longitudinal network. A collection can be created and edited with the collection dialog (the second to

last button in the tool bar). Use the `create collection` button to create a collection. Name the collection suitably and add all of the open networks that belong to your longitudinal network. The order of addition defines the order of observations.

To identify nodes that represent the same actor in the different networks, an *identifying* node attribute is needed. This attribute must be present in all included networks and take distinct values for all nodes of each individual network. If no identifying attribute exists, a network collection cannot be defined. Usually, the node attribute `id` can serve as identifying attribute, but using other attributes is of course possible.

A network collection is marked as being `siena compatible` if all included networks contain the same number of nodes. Only network collections that are `siena compatible` in this sense can be analyzed.

All operations are performed on the *active* network collection which can also be set in the network collection dialog.

4 Visualize Longitudinal Networks

By choosing on the `visualization` tab the `category` "layout", then, the `layout` "node layout" and, finally, the `node layout` "dynamic layout", you find functionalities for layouts of longitudinal networks.

5 Attributes

Attributes can be loaded or exported in the *attribute manager* (reachable via a button in the toolbar). For RSiena analysis only node and dyad attributes are used. Note that attributes must be loaded for each network individually.

5.1 Node Attributes

To load node attributes, choose *nodes* in the attribute manager (top) and *import & export* on the left. Using *import* out of the middle options, you can browse for the file that contains your attribute values. Data import supports the csv-format with several options for textframes and cell delimiters, e.g., " " as textframe and ; as cell delimiter.

"id";"gender";"age". . .
"1";"m";"14". . .
"2";"m";"15";. . .
"3";"f";"15";. . .
"4";"m";"16";. . .
⋮

Figure 1: Example for attribute file.

An exemplary node attribute file is shown in Figure 1. Mandatory are:

- a first line containing the names of the attributes

- a column containing an identifying attribute

The identifying attribute selected in the dialog is used to assign the attributes in the file to the nodes of the current network, therefore the name of the identifying attribute must be the same in the network and the file to be imported. Node attributes can be exported in an analogue way.

5.2 Dyad Attributes

To load dyad attributes, choose *dyads* and *import & export* in the attribute manager. After selecting *import*, you can browse for the file that contains your attribute values. The import file must contain a quare matrix with the number of rows and columns corresponding to the number of nodes in the network. Dyad attributes can be exported in an analogue way.

6 Configuration

To use RSiena, $^{\text{vis}}$on$_{\text{e}}$ connects to an instance of R that will finally run the estimations. To achieve this, we use the R-package Rserve, which allows connections via network to instances of R. The usual scenario is that the user installs R and the package Rserve and configures $^{\text{vis}}$on$_{\text{e}}$ to use this R installation. $^{\text{vis}}$on$_{\text{e}}$ will automatically install other needed libraries.
In the following we will assume that the user has a working installation of R (can be downloaded from http://www.r-project.org) and is able to start it.

6.1 Rserve Installation

A first step is the installation of the R-package RServe by running the command

```
install.packages("Rserve", dependencies=T)
```

in R. After successful installation, R can be closed (command `q()`) without saving anything.

6.2 Quick Configuration Guide

The way $^{\text{vis}}$on$_{\text{e}}$ connects to R is configured in the $^{\text{vis}}$on$_{\text{e}}$-options:
file→options..→R - connection

Windows On Windows R is usually started using the provided `Rserve.exe`, which is located in the library directory of R (usually `C:\Users\Peter\Documents\R`, for user "Peter"). Under certain circumstances `Rserve.exe` cannot be started directly from this directory and has to be moved to the `bin`-directory of the R installation (usually `C:\Program Files\R\R-11.1\bin`). You can test this by starting `Rserve.exe`. If it won't open a window that stays open, move `Rserve.exe` to the binary directory of R. Having a working R, we can configure ^{vis}on$_e$(cf. table 6.3):

- uncheck option 1 ("use R in system path")

- give the path to your `Rserve.exe` in option 2 ("path to R/Rserve dir"), omit the file name.

- check option 3 ("use Rserve")

- make sure option 4 is unchecked

- configure option 7 with the path to your R libraries (usually the R directory in your Documents-folder, see above)

Linux If you have an R installed by your package manager with a working copy of Rserve, ignore setup completely. If you want to run another R instance, that is not in the system path, use the appropriate options:

- uncheck option 1

- configure path to your R installation

- if necessary configure library directory in option 7

MacOS Usually no configuration is necessary, if that does not work, try the following command on a system console:

```
R CMD Rserve --vanilla
```

if this does not end with the message

```
Rserv started in daemon mode.
```

you have to start R manually.

6.3 Options - Short Overview

	option	explanation
1	use R in system path (linux/mac)	use the R/Rserve binary found in sysetem path usually only applies on Linux or Apple systems
2	path to R/RServe dir (if not in path)	the directory in which the R or RServe binary can be found (only needed if not in system path)
3	use Rserve (windows)	connect to R using the Rserve.exe binary, the default behaviour on windows systems
4	don't start R, but use	^{v}ison_e won't start its own R instance but use the one configured by points 5 and 6
5	server or ip	name of the server (or its ip-adress) that runs the Rserve to use (see 5)
6	port	the port the Rserve runs on (see 4)
7	library path (NULL uses system path)	path where R stores its libraries

Table 1: ^{v}ison_e options for R configuration

6.4 Using Your Own R Server

If you either want to use an R server on a different computer than the one ^{v}ison_e is running on or have problems with the default configuration, you can start your own server.

Open an R console and execute the following commands:

```
library(Rserve)
 Rserve()
```

In the following we assume that you have started Rserve in the described way. Start ^{v}ison_e, go the configuration menu, tab "R connection" and

- check option 4 ("don't start R but use")

- configure the server name and port in options 5 and 6

If you started R in the way described on the same machine ^{v}ison_e is running on, then the default values should be correct.

References

[1] Ripley, R.M., and Snijders, T.A.B. *Manual for SIENA version 4.0*, 2010. Oxford: University of Oxford, Department of Statistics; Nuffield College. `http://www.stats.ox.ac.uk/siena/`.

[2] Snijders, T.A.B., Steglich, C.E.G., van der Bunt, G.G. *Introduction to Actor-Based Models for Network Dynamics*, 2010. Social Networks, 32, 44-60.