

# Package ‘Rambo’

July 2, 2014

**Type** Package

**Title** The Random Subgraph Model

**Version** 1.1

**Date** 2013-11-13

**Author** Charles Bouveyron, Yacine Jernite, Pierre Latouche, Laetitia Nouedoui

**Maintainer** Pierre Latouche <pierre.latouche@univ-paris1.fr>

**Description** Estimate the parameters, the number of classes and cluster vertices of a random network into groups with homogeneous connection profiles. The clustering is performed for directed graphs with typed edges (edges are assumed to be drawn from multinomial distributions) for which a partition of the vertices is available.

**License** GPL (>= 2.0)

**Depends** stats, sna

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2013-11-13 10:50:05

## R topics documented:

Rambo-package . . . . .	2
plot.rsm . . . . .	3
Regions . . . . .	4
rsm . . . . .	4
summary.rsm . . . . .	6

<b>Index</b>	<b>8</b>
--------------	----------

Rambo-package

*The Random Subgraph Model*

---

**Description**

Partition the vertices of a directed network with typed edges into clusters describing the connection patterns of subgraphs given as inputs. The inference is performed using a variational Bayes EM algorithm and the estimation of the number of clusters is obtained through a variational approximation of the marginal log likelihood.

**Details**

Package: Rambo  
Type: Package  
Version: 1.0  
Date: 2013-06-18  
License: GPL (>=2.0)

**Author(s)**

Yacine Jernite, Laetitia Nouedoui, Charles Bouveyron, Pierre Latouche. Maintainer: Who to complain to <yourfault@somewhere.net>

**References**

Yacine Jernite, Pierre Latouche, Charles Bouveyron, Patrick Rivera, Laurent Jegou and Stephane Lamasse(2012), The Random Subgraph Model for the Analysis of an Ecclesiastical Network in Merovingian Gaul, <http://arxiv.org/abs/1212.5497>

**See Also**

[rsm](#), [summary.rsm](#), [plot.rsm](#)

**Examples**

```
data(Regions)
res <- rsm(Regions$X, Regions$sub, Klist=2:4, nbredo=1, maxit=5)
plot(res)
## summary(res)
```

---

plot.rsm	<i>Plot an object from class rsm</i>
----------	--------------------------------------

---

### Description

Produces plots describing results of an object from class rsm.

### Usage

```
## S3 method for class 'rsm'  
plot(x, frame=0, ...)
```

### Arguments

x	An object of class rsm. Result of function rsm.
frame	An integer with possible values 0, 1, or 2, depending on the kind of graphics requested (see details).
...	Arguments to be passed to the function, such as graphical parameters.

### Details

- if frame=0 plot of the lower bound criterion versus the number of clusters and repartition of the found cluster in the different subgraphs.
- if frame=1 plot of the lower bound criterion versus the number of clusters.
- if frame=2 repartition of found clusters in the different subgraphs.

### See Also

[rsm](#), [summary.rsm](#), [RAMBO](#)

### Examples

```
data(Regions)  
res <- rsm(Regions$X, Regions$sub, Klist=2:4, nbredo=1, maxit=5)  
##plot(res, frame=0)  
## plot(res, frame=1)  
plot(res, frame=2)
```

---

Regions

*A simulated network with subgraphs and typed edges*

---

### Description

The Regions data is a simulated network generated according to the Random Subgraph Model. The network is directed with 30 nodes; Vertices are known to belong to 2 subgraphs. Edges of this network are typed, with 3 kinds of links.

### Usage

```
data(Regions)
```

### Format

A list which contains the discrete adjacency matrix (X) of the network and a vector indicating the subgraph belongings of the nodes (sub).

### References

Yacine Jernite, Pierre Latouche, Charles Bouveyron, Patrick Rivera, Laurent Jegou and Stephane Lamasse(2012), The Random Subgraph Model for the Analysis of an Ecclesiastical Network in Merovingian Gaul, <http://arxiv.org/abs/1212.5497>

### Examples

```
data(Regions)
res <- rsm(Regions$X, Regions$sub, Klist=2:4, nbredo=1, maxit=5)
plot(res)
```

---

rsm

*The Random Subgraph Model*

---

### Description

Partition the vertices of a directed network with typed edges into clusters describing the connection patterns of subgraphs given as inputs. The inference is performed using a variational Bayes EM algorithm and the estimation of the number of clusters is obtained through a variational approximation of the marginal log likelihood.

### Usage

```
rsm(X, sub, Klist, nbredo = 5, disp = F, maxit = 50)
```

**Arguments**

<code>X</code>	an adjacency matrix with oriented and typed edges.
<code>sub</code>	a vector indicating the subgraph in which each vertex belongs.
<code>Klist</code>	a vector indicating the range of possible values for the number of latent clusters; must be at least 2.
<code>nbredo</code>	number of initializations to be made; to avoid local optimum during variational Bayes EM algorithm.
<code>disp</code>	logical; if TRUE, display comments during the estimation.
<code>maxit</code>	maximum number of iterations for the algorithm.

**Details**

`rsm` implements the variational Bayes EM algorithm for the random subgraph model (RSM) as proposed by Jernite et al.(2012). The function takes a directed network with typed edges along with a decomposition of the network into known subgraphs as inputs.

The RSM model assumes that the presence of an edge between two vertices depends on the subgraphs they belong to. If an edge is present, its type is then assumed to be sampled from a multinomial distribution, depending on latent clusters. In practice, the subgraphs are known and given as inputs while the clusters have to be inferred from the data. The clustering of the vertices is performed using a variational Bayes algorithm and the number of clusters is obtained with a model selection criterion which is a variational approximation of the marginal log likelihood.

The algorithm is runned for various values of the number of clusters (`Klist`). For each value of `K` in `Klist`, the algorithm is initialized `nbredo` times.

Assuming that the number of clusters is `K`, `output[[K]]$lower` represents the lower bound of the marginal log likelihood. `output[[K]]$Pi` contains the cluster connectivity. Each element of the  $K \times K$  matrix is a vector of probability of `C` elements. Hence, `output[[K]]$Pi` is an array of size  $K \times K \times C$ .

In order to have a better chance of reaching a global optimum, VBEM is run for several initializations of a kmeans like algorithm (by default, `nbredo` = 5).

**Value**

Returns a list including:

<code>N</code>	Number of vertices of the network
<code>R</code>	Number of subgraphs
<code>C</code>	Number of edges types
<code>Klist</code>	vector indicating the range of possible values for the number of latent cluster; must be greater or equal to 2
<code>X</code>	an adjacency matrix with oriented and typed edges.
<code>K_star</code>	the number of clusters estimated.
<code>gama</code>	matrix of size $R \times R$ containing the probabilities of connection between subgraphs.

output            output list of length(Klist) + 1 items. Each item contains the result of the estimation for a given number of class K. Details of output field:

output[[K]]\$lower            lower bound of the marginal log likelihood.

output[[K]]\$alpha            matrix of size R\*K containing the posterior proportions of each cluster in sub-graphs.

output[[K]]\$Pi            array of size K\*K\*C containing the cluster connectivity. Each element of the K\*K matrix is a vector of C elements.

output[[K]]\$Tau            matrix of posterior probabilities (estimated probabilities for each vertex to belong to the different clusters.

output[[K]]\$Zcol            vector indicating the cluster of each vertex (MAP estimate).

### Author(s)

Yacine Jernite, Laetitia Nouedoui, Charles Bouveyron, Pierre Latouche.

### References

Yacine Jernite, Pierre Latouche, Charles Bouveyron, Patrick Rivera, Laurent Jegou and Stephane Lamasse(2012), The Random Subgraph Model for the Analysis of an Ecclesiastical Network in Merovingian Gaul, <http://arxiv.org/abs/1212.5497>

### See Also

[summary.rsm](#), [plot.rsm](#)

### Examples

```
data(Regions)
res <- rsm(Regions$X, Regions$sub, Klist=2:4, nbredo=1, maxit=5)
plot(res)
summary(res)
```

---

summary.rsm

*Summarizes results of Random Subgraph Model estimates.*

---

### Description

Returns a summary of rsm function results.

### Usage

```
## S3 method for class 'rsm'
summary(object,...)
```

### **Arguments**

<code>object</code>	An object of class <code>rsm</code> . This is assumed to be the result of function <code>rsm</code> .
<code>...</code>	Additional options for the summary function.

### **See Also**

[rsm](#), [plot.rsm](#)

### **Examples**

```
data(Regions)
res <- rsm(Regions$X, Regions$sub, Klist=2:4, nbredo=1, maxit=5)
summary(res)
```

# Index

\*Topic **Regions**

Regions, [4](#)

\*Topic **rsm**

Rambo-package, [2](#)

Regions, [4](#)

rsm, [4](#)

summary.rsm, [6](#)

\*Topic **subgraph**

Regions, [4](#)

rsm, [4](#)

plot.rsm, [2](#), [3](#), [6](#), [7](#)

RAMBO, [3](#)

RAMBO (Rambo-package), [2](#)

Rambo (Rambo-package), [2](#)

Rambo-package, [2](#)

Regions, [4](#)

rsm, [2](#), [3](#), [4](#), [7](#)

summary.rsm, [2](#), [3](#), [6](#), [6](#)