Package ‘NetSAM’

October 8, 2014

Type Package

Title Network Seriation And Modularization

Version 1.4.0

Date 2014-01-23

Author Jing Wang <`jing.wang.2@vanderbilt.edu`>

Maintainer Bing Zhang <`bing.zhang@vanderbilt.edu`>

Description The NetSAM (Network Seriation and Modularization) package takes an edge-list representation of a network as an input,performs network seriation and modularization analysis, and generates as files that can be used as an input for the
one-dimensional network visualization tool NetGestalt
(http://www.netgestalt.org) or other network analysis.

License LGPL

LazyLoad yes

Depends R (>= 2.15.1), methods, igraph (>= 0.6-1), seriation (>= 1.0-6), graph (>= 1.34.0)

Imports methods

Suggests RUnit, BiocGenerics

Collate zzz.R NetSAM.R

biocViews Visualization, Network

R topics documented:

`NetSAM-package` *Network Seriation and Modularization*

Description

The NetSAM (Network Seriation and Modularization) package takes an edge-list representation of a network as an input, performs network seriation and modularization analysis, and generates as files that can be used as an input for the one-dimensional network visualization tool NetGestalt (http://www.netgestalt.org) or other network analysis.

Details

Package:	NetSAM
Type:	Package
Version:	1.3.0
Date:	2014-01-23
License:	LGPL
LazyLoad:	yes

Author(s)

Jing Wang Maintainer: Bing Zhang <bing.zhang@Vanderbilt.edu>

References

NetGestalt: integrating multidimensional omics data over biological networks. Nature Methods 10, 597-598 (2013).

See Also

`NetSAM`

`inputNetwork` *A network example with graphNEL class*

Description

The network is a sub-network of HPRD (Human Protein Reference Database) network with 320 nodes and 769 edges, which is an example to show how to input network with graphNEL class into the package.

Usage

```
data(inputNetwork)
```

Format

graphNEL

`NetSAM` *Network Seriation and Modularization*

Description

The NetSAM function uses random walk distance-based hierarchical clustering to identify the hierarchical modules of the network and then uses the optimal leaf ordering (OLO) method to optimize the one-dimensional ordering of the genes in each module by minimizing the sum of the pair-wise random walk distance of adjacent genes in the ordering.

Usage

```
NetSAM(inputNetwork, outputFileName, minModule = (-1), maxStep = 4, method = "Modularity Cutoff", Modul
```

Arguments

`inputNetwork`	The network under analysis. inputNetwork can be the name of the input network file in the edge-list format (each row represents an edge with two node names separated by a tab or space) or be a data object in R (data object must be graphNEL class or data.frame (or matrix) class with two columns).
`outputFileName`	The name of the output file. If no file path is provided, the output file will be saved to the current path.
`minModule`	The minimum number of nodes for a module (or minimum module size). If the size of a module identified by the function is less than the specified number, the module will not be further partitioned into sub-modules. The default is -1 which means NetSAM will set `minModule` as 5 or 0.3 percent of the number of nodes in the input network, whichever is larger.
`maxStep`	Because NetSAM uses random walk distance-based hierarchical clustering to reveal the hierarchical organization of an input network, it requires a specified length of the random walks. To get the optimal length, the function will test a range of lengths ranging from 2 to `maxStep`. The default is 4.
`method`	To test whether a network under consideration has a non-random internal modular organization, the function provides three options: "Modularity Cutoff", "ZScore" and "Permutation". "Modularity Cutoff" means if the modularity score of the network is above a specified cutoff value, the network will be considered to have internal organization and will be further partitioned. For "ZScore" and "Permutation", the function first uses the edge switching method to generate a given number of random networks with the same number of nodes and an identical degree sequence and calculates the modularity scores for these random

	networks. Then, "ZScore" method will transform the real modularity score to a z score based on the random modularity scores and then transform the z score to a p value assuming a standard normal distribution. The "Permutation" method will compare the real modularity score with the random ones to calculate a p value. Finally, under a specified significance level, the function determines whether the network can be further partitioned. The default is "Modularity Cutoff".
`ModularityThr`	Threshold of modularity score for the "Modularity Cutoff" method. The default is 0.2
`ZRandomNum`	The number of random networks that will be generated for the "ZScore" calculation. The default is 10.
`permuteNum`	The number of random networks that will be generated for the "Permutation" p value calculation. The default is 100.
`pThr`	The significance level for determining whether a network has non-random internal modular organization for the "ZScore" or "Permutation" methods.

Note

Because the seriation step requires pair-wise distance between all nodes, NetSAM is memory consuming. We recommend to use the 64 bit version of R to run the NetSAM. For networks with less than 10,000 nodes, we recommend to use a computer with 8GB memory. For networks with more than 10,000 nodes, a computer with at least 16GB memory is recommended.

Author(s)

Jing Wang

Examples

```
cat("The input network can be a file.\n")
inputNetwork <- system.file("extdata","exampleNetwork.txt",package="NetSAM")
cat("The input network can be also a data object, such as graphNEL object.\n")
data(inputNetwork)
outputFileName <- paste(getwd(),"/NetSAM",sep="")
result <- NetSAM(inputNetwork, outputFileName, minModule = (-1), maxStep = 4, method = "Modularity Cutoff", Modular
```

Index