

Package ‘ergm.graphlets’

June 11, 2015

Type Package

Title ERG Modeling Based on Graphlet Properties

Version 1.0.3

Date 2015-06-10

LinkingTo ergm

Suggests parallel, rlecuyer

Depends network ($\geq 1.7-1$), ergm ($\geq 3.1-0$), statnet.common ($\geq 3.1-0$)

Description Integrates graphlet statistics based model terms for use in exponential-family random graph models ('ergm') as part of the 'statnet' suite of packages.

License GPL (≥ 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2015-06-11 18:31:38

Author Omer Nebil Yaveroglu [aut, cre],
Sean M. Fitzhugh [aut],
Maciej Kurant [aut],
Athina Markopoulou [aut],
Natasza Przulj [aut],
Carter T. Butts [aut]

Maintainer Omer Nebil Yaveroglu <o.yaveroglu10@imperial.ac.uk>

R topics documented:

ergm.graphlets-package	2
emon3	2
graphletCount	3
grorbitCov	4
grorbitDist	5
grorbitFactor	6
spi	7

Index	9
--------------	----------

ergm.graphlets-package

ergm.graphlets: A Package for ERG Modeling Based on Graphlet Statistics

Description

The **ergm.graphlets** package adds four new ERG modeling terms into the **ergm** package. The four modelling terms are:

- 1) graphletCount – Accounts for the number of graphlets in the network
- 2) grorbitCov – Accounts for the link between graphlet degrees of different orbits and the numerical node attributes.
- 3) grorbitFactor – Accounts for the link between graphlet degrees of different orbits and the categorical node attributes.
- 4) grorbitDist – Accounts for the graphlet degree distributions of 2- to 4-node graphlet orbits.

The **ergm.graphlets** package is open-source and released under GPL-2 licence. When publishing results obtained using this package the original authors are to be cited as:

Omer N. Yaveroglu, Sean M. Fitzhugh, Maciej Kurant, Athina Markopoulou, Carter T. Butts, Natasa Przulj. 2013 *ergm.graphlets: A Package for ERG Modeling Based on Graphlet Statistics* <http://CRAN.R-project.org/package=ergm.graphlets>.

We have invested a lot of time and effort in developing the `ergm.graphlets` package. Please cite it in all papers where it is used.

For complete citation information, use `citation(package="ergm.graphlets")`.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, *Journal of Statistical Software* 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

statnet, network, ergm, ergm-terms

emon3

EMON Dataset

Description

The dataset used for illustrating the disaster response network modeling example

Details

The dataset is originally known as EMON dataset in literature. However, we modified the dataset to make it undirected and included some categorical node attributes that were originally provided with the network.

EMON Dataset

EMON dataset The Lake Pomona disaster response network that is used for illustrating the graphlet-Cov terms.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, Journal of Statistical Software 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

statnet, network, ergm, ergm-terms

graphletCount	<i>New Graphlet Count Term for the 'ergm' Package – introduced under 'ergm.graphlets' package</i>
---------------	---

Description

This is a new term for the [ergm](#) package that counts the number of graphlets in the network for a given graphlet type.

Details

The function [ergm](#) is used to fit linear exponential random graph models, in which the probability of a given network, y , on a set of nodes is $\exp\{\theta \cdot g(y)\} / c(\theta)$, where $g(y)$ is a vector of network statistics for y , θ is a parameter vector of the same length and $c(\theta)$ is the normalizing constant for the distribution.

The network statistics $g(y)$ are entered as terms in the function call to [ergm](#). The [ergm](#) package contains a wide range of terms. For the details on the possible `<model terms>`, see [ergm-terms](#).

The [ergm.userterms](#) package provides a template for adding new terms. The terms can be used throughout the [ergm](#) package and behave identically to the supplied terms.

The Graphlet Count Term

`graphletCount(g)` Statistics for the number of times that a graphlet appears in a network can be included in an ERGM by using the `graphletCount` term. The question that the change score function of this term answers is: how do the number of graphlets of type G_i change when an edge is flipped in the network? This term has an optional argument, `g`. `g` is a vector of distinct integers representing the list of graphlets to be evaluated during the estimation of model coefficients. When this argument is not provided, all graphlets are evaluated by default. The term adds one network statistic to the model for each element in `g`. This term is defined for the 30 graphlets with up to 5 nodes. Therefore, `g` accepts values between 0 and 29. Values outside this range are ignored.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, *Journal of Statistical Software* 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

`statnet`, `network`, `ergm`, `ergm-terms`

grorbitCov

New Graphlet Orbit Covariance Term for the 'ergm' Package – introduced under 'ergm.graphlets' package

Description

This is a new term for the `ergm` package that links the graphlet structures in the network for a given graphlet orbit with numeric node attributes.

Details

The function `ergm` is used to fit linear exponential random graph models, in which the probability of a given network, y , on a set of nodes is $\exp\{\theta \cdot g(y)\} / c(\theta)$, where $g(y)$ is a vector of network statistics for y , θ is a parameter vector of the same length and $c(\theta)$ is the normalizing constant for the distribution.

The network statistics $g(y)$ are entered as terms in the function call to `ergm`. The `ergm` package contains a wide range of terms. For the details on the possible `<model terms>`, see `ergm-terms`.

The `ergm.userterms` package provides a template for adding new terms. The terms can be used throughout the `ergm` package and behave identically to the supplied terms.

The Graphlet Orbit Covariance Term

`grorbitCov(attrName, grOrbit)` The correlation between a node's graphlet degree and a numeric attribute value can be included into an ERGM by using the `grorbitCov` term. The question that the change score function of this term answers is: what is the change in covariance between a vector of nodal attributes and nodal orbit degrees when an edge is changed? This term has two arguments: `attrname` and `grorbit`. The `attrname` is a character vector giving the name of a numeric attribute in the network's vertex attribute list. The optional `grorbit` argument is a vector of distinct integers representing the list of graphlet orbits to include into the ERGM model. When `grorbit` is not provided, all graphlet orbits are evaluated by default. The term adds one network statistic to the model for each element in `grorbit`. This term is defined for the 73 orbits corresponding to graphlets with up to 5 nodes. Therefore, `grorbit` accepts values between 0 and 72. Values outside this range are ignored. This term can only be used with undirected networks.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, Journal of Statistical Software 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

`statnet`, `network`, `ergm`, `ergm-terms`

grorbitDist

New Graphlet Degree Distribution term for the 'ergm' Package – introduced under 'ergm.graphlets' package

Description

This is a new term for the `ergm` package that evaluates the graphlet degree distribution of a network for a given graphlet orbit.

Details

The function `ergm` is used to fit linear exponential random graph models, in which the probability of a given network, y , on a set of nodes is $\exp\{\theta \cdot g(y)\} / c(\theta)$, where $g(y)$ is a vector of network statistics for y , θ is a parameter vector of the same length and $c(\theta)$ is the normalizing constant for the distribution.

The network statistics $g(y)$ are entered as terms in the function call to `ergm`. The `ergm` package contains a wide range of terms. For the details on the possible `<model terms>`, see `ergm-terms`.

The `ergm.userterms` package provides a template for adding new terms. The terms can be used throughout the `ergm` package and behave identically to the supplied terms.

The Graphlet Degree Distribution Term

`grorbitDist(grorbit, d)` The graphlet degree distributions of various graphlet orbits can be included into the ERGM by using the `grorbitDist` term. The question that the change score function of this term answers is: how do the number of nodes having graphlet degree n for orbit i change when an edge is flipped? This term has two arguments: `grorbit` and `d`. The `grorbit` argument is a vector of distinct integers representing the list of graphlet orbits to include into the model. The `d` argument is a vector of distinct integers. This term adds one network statistic to the model for each pairwise combination of the arguments in `grorbit` and `d` vectors. The statistic for the combination of (i, j) is equal to the number of nodes in the network that have graphlet degree j for orbit i . This term is defined for the 15 graphlet orbits corresponding to graphlets with up to 4 nodes. Therefore, `grorbit` accepts values between 0 and 14. Graphlets of size 5 are omitted for this term because of the high computational complexity of the change score computation of the term. This term is defined for undirected networks.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, Journal of Statistical Software 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

`statnet`, `network`, `ergm`, `ergm-terms`

grorbitFactor

New Graphlet Orbit Factor Term for the 'ergm' Package – introduced under 'ergm.graphlets' package

Description

This is a new term for the `ergm` package that links the graphlet orbits of a given type with categorical node attributes.

Details

The function `ergm` is used to fit linear exponential random graph models, in which the probability of a given network, y , on a set of nodes is $\exp\{\theta \cdot g(y)\} / c(\theta)$, where $g(y)$ is a vector of network statistics for y , θ is a parameter vector of the same length and $c(\theta)$ is the normalizing constant for the distribution.

The network statistics $g(y)$ are entered as terms in the function call to `ergm`. The `ergm` package contains a wide range of terms. For the details on the possible `<model terms>`, see `ergm-terms`.

The `ergm.userterms` package provides a template for adding new terms. The terms can be used throughout the `ergm` package and behave identically to the supplied terms.

The Graphlet Orbit Factor Term

`grorbitFactor(attrname, grorbit, base)` The `grorbitFactor` term adds a relationship between graphlet degrees and a categorical node attribute into an ERGM. The question that the change score function of this term answers is: what is the change in the total graphlet degree (for a given orbit) for those nodes with a given attribute value, for a particular edge change? This term has three arguments: `attrname`; `grorbit`; and `base`. `attrname` is a character vector giving the name of a categorical attribute in the network's vertex attribute list. The optional `grorbit` argument is a vector of distinct integers representing the list of graphlet orbits to include into the model. When `grorbit` is not provided, all graphlet orbits are evaluated by default. The optional `base` argument is a vector of distinct integers representing the list of categories in `attrname` that are going to be omitted. When this argument is set to 0, all categories are evaluated. When this argument is set to 1, the category having the lowest value (or lexicographically first name) is eliminated. The term sorts all values of the categorical attribute lexicographically and `base` term defines the indexes of the categories to be omitted in this sorted list. For example, if the "fruit" attribute has values "orange", "apple", "banana" and "pear", `grorbitFactor("fruit" , 0 , 2:3)` will ignore the "banana" and "orange" factors and evaluate the "apple" and "pear" factors. When the `base` argument is not provided, the argument is set to 1 by default. The `grorbitFactor` term adds $a \times |\text{grorbit}|$ terms into the model where a represents the number of attribute values that are evaluated in the model and $|\text{grorbit}|$ is the number of graphlet orbits to be evaluated in the model. This term is defined for the 73 graphlet orbits corresponding to graphlets with up to 5 nodes. Therefore, `grorbit` accepts values between 0 and 72. Values outside this range are ignored. This term is defined for undirected networks.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, *Journal of Statistical Software* 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

`statnet`, `network`, `ergm`, `ergm-terms`

 spi

SPI Dataset

Description

The dataset used for illustrating the protein structure network modeling example

Details

The dataset is preprocessed for including the assembly attributes to nodes.

SPI Dataset

SPI dataset The protein structure network of matriptase complex that is used in Milo et al. (2004) paper.

References

Yaveroglu ON, Fitzhugh SM, Kurant M, Markopoulou A, Butts CT, Przulj N (2015). **ergm.graphlets**: A Package for ERG Modeling Based on Graphlet Properties, Journal of Statistical Software 65(12), 1-29, URL <http://www.jstatsoft.org/v65/i12/>.

See Also

statnet, network, ergm, ergm-terms

Index

*Topic **emonData**

emon3, [2](#)

*Topic **models**

ergm.graphlets-package, [2](#)

graphletCount, [3](#)

grorbitCov, [4](#)

grorbitDist, [5](#)

grorbitFactor, [6](#)

*Topic **package**

ergm.graphlets-package, [2](#)

*Topic **spiData**

spi, [7](#)

EMON (emon3), [2](#)

emon3, [2](#)

ergm, [3–6](#)

ergm (ergm.graphlets-package), [2](#)

ergm-terms (ergm.graphlets-package), [2](#)

ergm.graphlets

(ergm.graphlets-package), [2](#)

ergm.graphlets-package, [2](#)

graphletCount, [3](#)

grorbitCov, [4](#)

grorbitDist, [5](#)

grorbitFactor, [6](#)

InitErgmTerm.graphletCount

(graphletCount), [3](#)

InitErgmTerm.grorbitCov (grorbitCov), [4](#)

InitErgmTerm.grorbitDist (grorbitDist),

[5](#)

InitErgmTerm.grorbitFactor

(grorbitFactor), [6](#)

SPI (spi), [7](#)

spi, [7](#)