Package 'orca'

September 20, 2024

Version 1.1-3

Date 2024-09-19

Title Computation of Graphlet Orbit Counts in Sparse Graphs

Description Implements orbit counting using a fast combinatorial approach. Counts orbits of nodes and edges from edge matrix or data frame, or a graph object from the graph package.

License LGPL-3

Depends R (>= 3.1)

Enhances graph

NeedsCompilation yes

Collate orca.R

LazyLoad yes

LazyData yes

Author Tomaz Hocevar [aut, cre],
Janez Demsar [aut]

Maintainer Tomaz Hocevar <tomaz.hocevar@fri.uni-lj.si>

Repository CRAN

Date/Publication 2024-09-20 14:00:06 UTC

Contents

---

karate   *Karate Club network*

Description

The network representing the friendships between members of a university-based karate club, which was originally used to model the fission process with a mathematical model. The network consists of 34 nodes (club members) and 77 edges (friendships).

Usage

```
karate
```

Format

A data frame with 77 observations and 2 columns.

Source

https://networkdata.ics.uci.edu/data.php?id=105

References

W. W. Zachary (1977) An information flow model for conflict and fission in small groups, *Journal of Anthropological Research* 33(4), 452-473.

---

orca   *Orbit counting*

Description

Count the node or edge orbits in 4- or 5- node graphlets for all nodes (edges) in the given graph.

Usage

```
count4(graph)
count5(graph)
ecount4(graph)
ecount5(graph)
```

Arguments

graph   A graph given as a nx2 edge matrix, a data frame with edges or a graph object
                   from the package 'graph'. The nodes in the matrix or data frame are given by
                   integer indices that start with 1.

Value

A numeric matrix or orbit counts. Rows correspond to graph nodes or edges in the same order as on the input, and the columns corresponding to orbits.

Author(s)

Tomaz Hocevar and Janez Demsar

References

Tomaz Hocevar, Janez Demsar (2016): Computation of Graphlet Orbits for Nodes and Edges in Sparse Graphs. Journal of Statistical Software, 71(10), pp. 1-24.

Examples

```
library(orca)

# Load and show the orbit counts for the Karate graph
data("karate")
count4(karate)


## Not run:
# Simple analysis of School Wikipedia network: find the most similar
# nodes with respect to the local network topology
# Requires data from http://www.biolab.si/supp/Rorca/_downloads/schools-wiki.zip

library("FNN")

nodes <- scan("schools-wiki-nodes.txt", what="", sep="\n")
edges <- read.table("schools-wiki-edges.txt")

orbits <- count4(edges)
nn <- get.knn(orbits, k=10)
neighbours <- nn$nn.index
distances <- nn$nn.dist

check <- c("Canada", "Germany", "Isaac Newton", "Albert Einstein",
           "Mahatma Gandhi", "Mahabharata")
node_indices <- match(check, nodes)
for (i in 1:length(check)) {
    cat("\n\n", check[i], ": ", sep="")
    cat(nodes[neighbours[node_indices[i], ]], sep=", ")
    cat("\n")
    cat(round(distances[node_indices[i], ]), sep=", ")
}

## End(Not run)
```

---

petersen	*Petersen graph*

Description

The Petersen graph is a graph with 10 vertices and 15 edges, usually drawn as a pentagram within a pentagon.

Usage

```
petersen
```

Format

A data frame with 15 observations and 2 columns.

References

J. Petersen Sur la théorème de Tait. *L'Intermédiare des Math*. 5, 225-227.

---

usastates	*Contiguous USA Graph*

Description

The nodes of Contiguous USA Graph represent the 49 contiguous states of the USA. Two nodes are connected if there exists at least one driveable road between the corresponding states.

Usage

```
usastates
```

Format

A data frame with 107 observations and 2 columns.

Source

http://www-cs-staff.stanford.edu/~uno/contiguous-usa.dat

References

D. E. Knuth (2008) *The Art of Computer Programming, Volume 4, Fascicle 0: Introduction to Combinatorial Functions and Boolean Functions*, p. 15. Upper Saddle River, NJ: Addison-Wesley.

---

yeast   *Yeast protein-protein interaction network*

Description

The network representing interactions between 2361 proteins.

Usage

```
yeast
```

Format

A data frame with 6646 observations and 2 columns.

Source

http://vlado.fmf.uni-lj.si/pub/networks/data/bio/Yeast/Yeast.htm

References

S. Sun, L. Ling, N. Zhang, G. Li and R. Chen (2003) Topological structure analysis of the protein-protein interaction network in budding yeast, *Nucleic Acids Research*, 31(9), 2443-2450.

Index