Package 'PhenoGeneRanker'

January 21, 2025

Type Package

Title PhenoGeneRanker: A gene and phenotype prioritization tool

Version 1.14.0

Description This package is a gene/phenotype prioritization tool that utilizes multiplex heterogeneous gene phenotype network. PhenoGeneRanker allows multi-layer gene and phenotype networks. It also calculates empirical p-values of gene/phenotype ranking using random stratified sampling of genes/phenotypes based on their connectivity degree in the network. https://dl.acm.org/doi/10.1145/3307339.3342155.

License Creative Commons Attribution 4.0 International License

Imports igraph, Matrix, foreach, doParallel, dplyr, stats, utils, parallel

Suggests knitr, rmarkdown

Encoding UTF-8

VignetteBuilder knitr

RoxygenNote 7.1.1

biocViews BiomedicalInformatics, GenePrediction, GraphAndNetwork, Network, NetworkInference, Pathways, Software, SystemsBiology

git_url https://git.bioconductor.org/packages/PhenoGeneRanker

git_branch RELEASE_3_20

git_last_commit 6e146fc

git_last_commit_date 2024-10-29

Repository Bioconductor 3.20

Date/Publication 2025-01-20

Author Cagatay Dursun [aut, cre]

Maintainer Cagatay Dursun <cagataydursun@gmail.com>

Contents

`CreateWalkMatrix` *Create Walk Matrix*

Description

Generates a Walk Matrix (Transition Matrix) from Gene and Phenotype networks for RWR.

Usage

```
CreateWalkMatrix(
  inputFileName,
  numCores = 1,
  delta = 0.5,
  zeta = 0.5,
  lambda = 0.5
)
```

Arguments

`inputFileName`	The name of the text file that contains the names of gene and phenotype network files. For more information on the file formatting, please refer the vignette.
`numCores`	This is the number of cores used for parallel processing.
`delta`	This is the probability of jumping between gene layers. High delta means RWR is high likely to jump to other layers in gene multiplex network. It has a default value of 0.5.
`zeta`	This is the probability of jumping between phenotype layers. High zeta means RWR is high likely to jump to other layers in phenotype multiplex network. It has a default value of 0.5.
`lambda`	This is the probability of jumping between gene and phenotype multiplex networks. High lambda means RWR is more likely to exploit the bipartite relation. It has a default value of 0.5.

Value

This returns a list containing the walk matrix, a sorted list of gene ids, a sorted list of phenotype ids, the connectivity degree of the genes, the connectivity degree of the phenotypes, the number of gene layers, the number of phenotype layers, the number of genes and the number of phenotypes in the final complex network.

Examples

```
wm <- CreateWalkMatrix('input_file.txt')
customWm <- CreateWalkMatrix('input_file.txt', numCores=1, delta=0.7, zeta=0.7, lambda=0.7)
```

RandomWalkRestart *Random Walk Restart (RWR)*

Description

This method runs the random walk with restart on the provided walk matrix. It returns a data frame including ranked genes and phenotypes, and RWR scores of the genes and phenotypes. If generatePvalue is TRUE then it generates p-values along with the ranks.

Usage

```
RandomWalkRestart(
  walkMatrix,
  geneSeeds,
  phenoSeeds,
  generatePValue = TRUE,
  numCores = 1,
  r = 0.7,
  eta = 0.5,
  tau = NULL,
  phi = NULL,
  S = 1000
)
```

Arguments

walkMatrix	This is the walk matrix generated by the function CreateWalkMatrix.
geneSeeds	This is a vector for storing the ids of the genes that RWR starts its walk. The final ranks show the proximity of the genes/phenotypes to the seed genes.
phenoSeeds	This is a vector for storing the ids of the phenotypes that RWR starts its walk. The final ranks show the proximity of the genes/phenotypes to the seed phenotypes.
generatePValue	If this is TRUE, it will generate the probability values for each of the gene/phenotype rankings. If it is FALSE then the function will only return the ranks of genes/phenotype.
numCores	This is the number of cores used for parallel processing.
r	This parameter controls the global restart probability of RWR, and it has a default value of 0.7.
eta	This parameter controls restarting of RWR either to a gene seed or phenotype seeds, higher eta means utilizing gene seeds more than phenotype seeds, and it has a default value of 0.5.
tau	This is a vector that stores weights for each of the 'gene' layer in the complex gene and phenotype network. Each value of the vector corresponds to the order of the network files in the input file of CreateWalkMatrix function. The weights must sum up to the same number of gene layers. Default value gives equal weight to gene layers.
phi	This is a vector that stores weights for each of the 'phenotype' layer in the complex gene and phenotype network. Each value of the vector corresponds to the order of the network files in the input file of CreateWalkMatrix function. The weights must sum up to the same number of phenotype layers. Default value gives equal weight to phenotype layers.

S This is the number of random samples to be used for p-value calculation It is highly recommended to use S=1000.

Value

If the parameter generatePValue is TRUE, then this function returns a data frame of ranked genes/phenotypes with p-values with three columns; Gene/Phenotype ID, score, p-value. If generatePValue is FALSE, then it returns a data frame of ranked genes/phenotypes with two columns; Gene/Phenotype ID, score.

Examples

```
wm <- CreateWalkMatrix('input_file.txt')
ranksWithoutPVal <- RandomWalkRestart(wm, c('g1', 'g2'), c('p1'), FALSE)
ranksWithPVal <- RandomWalkRestart(wm, c('g1', 'g2'), c(), TRUE, S=10)
ranksWithoutPval <- RandomWalkRestart(wm, c('g1'), c(),
       FALSE, 1, 0.5, 0.6, tau=c(1.5,0.5), phi=c(0.5,1.5))
```

Index