Package: DRaWR (via r-universe)

October 14, 2024

Title Discriminative Random Walk with Restart

Version 1.0.3

Author Charles Blatti [aut, cre]

Maintainer Charles Blatti <blatti@illinois.edu>

Description We present DRaWR, a network-based method for ranking genes
or properties related to a given gene set. Such related genes
or properties are identified from among the nodes of a large,
heterogeneous network of biological information. Our method
involves a random walk with restarts, performed on an initial
network with multiple node and edge types, preserving more of
the original, specific property information than current
methods that operate on homogeneous networks. In this first
stage of our algorithm, we find the properties that are the
most relevant to the given gene set and extract a subnetwork of
the original network, comprising only the relevant properties.
We then rerank genes by their similarity to the given gene set,
based on a second random walk with restarts, performed on the
above subnetwork.

Depends R (>= 3.0)

License GPL-2

LazyData true

Imports Matrix, ROCR, methods

RoxygenNote 5.0.1.9000

NeedsCompilation no

Date/Publication 2022-03-24 14:10:02 UTC

Repository https://cblatti3.r-universe.dev

RemoteUrl https://github.com/cran/DRaWR

RemoteRef HEAD

RemoteSha 01afe80e047b7f6fe5451cf97d1161c924d2b4a0

1

Contents

DRaWR   *DRaWR*

Description

This function runs the DRaWR two stage random walk with restart method.

Usage

```
DRaWR(possetfile = "extdata/sample.setlist", unifile = "extdata/sample.uni",
  networkfile = "extdata/sample.edge", outdir = "output_",
  restarts = c(0.7), nfolds = 1, st2keep = 1, undirected = TRUE,
  unweighted = FALSE, normalize = "type", maxiters = 50, thresh = 1e-04,
  property_types = c("allen_brain_atlas", "chip_binding", "gene_ontology",
  "motif_u5", "pfam_domain", "T1", "T2"), writepreds = 0)
```

Arguments

possetfile	(string): location of file containing location of gene sets to test.
unifile	(string): location of file listing gene universe.
networkfile	(string): location of file containing network contents.
outdir	(string): prefix of location of file to write performance results (optionally prediction results).
restarts	(vector): vector of restart values to test. Default is c(0.7).
nfolds	(int): number of folds for cross validation, 1 is no cross-validation. Default is 4.
st2keep	(int): number of property nodes to keep in second stage for each property type. Default is 50.
undirected	(bool): boolean to make network undirected.
unweighted	(bool): boolean to make network unweighted.
normalize	(string): "type" or "none". Default is 'type'.
maxiters	(int): maximum number of allowable iterations. Default is 50.
thresh	(float): threshold for L1 norm convergence. Default is 0.001.
property_types	(vector): list of possible property types. Default is c("go_curated_evidence", "go_inferred_evidence", "pfam_domain").
writepreds	(boolean): write predictions out to a file. Default is FALSE

Examples

```
DRaWR(possetfile = system.file("extdata", "sample.setlist", package="DRaWR"),
 unifile = system.file("extdata", "sample.uni", package="DRaWR"),
 networkfile = system.file("extdata", "sample.edge", package="DRaWR"),
 outdir = "exampleRun_", restarts = c(.7), nfolds = 1, st2keep = 1,
 undirected = TRUE, unweighted = FALSE, normalize = "type", maxiters = 50,
 thresh = 0.0001, property_types = c("T1", "T2"), writepreds = 0)
```

RWR   *RWR*

Description

This function runs a random walk with restart using two supported matrix representations.

Usage

```
RWR(boolSparceMat, transmat, restart, query, startvec, maxiters, thresh)
```

Arguments

boolSparceMat	(bool): Boolean to indicate sparce Matrix or list matrix.
transmat	(sparce Matrix / list matrix): transition probabilities.
restart	(float): probability of restart.
query	(vector): probability of restarting at all nodes.
startvec	(vector): initial probability of being at any node.
maxiters	(int): maximum number of allowable iterations.
thresh	(float): threshold for L1 norm convergence.

Value

list of 'iter':number of iterations, 'diff': L1 norm of difference, 'vec': converged probability distribution vector.

Examples

```
RWR(boolSparceMat=TRUE, transmat=transmat, restart=.3, query=c(rep(0.1,10),rep(0,5)),
 startvec=rep(1/15,15), maxiters=10, thresh=0.001)
```

---

threeCol2listMat   *threeCol2listMat*

Description

This function takes a three vectors of equal length (source nodes, target nodes, and edge weights) and return the adjacency matrix as a list of vectors.

Usage

```
threeCol2listMat(a = c("a", "b", "c", "c"), b = c("a", "b", "b", "b"),
  v = c(1, 2, 3, 4))
```

Arguments

a	(vector): vector of source node names.
b	(vector): vector of target node names.
v	(vector): vector of edge weights names.

Value

list of vectors matrix representation.

Examples

```
threeCol2listMat(a = c("a","b","c","c"), b = c("a","b","b","b"), v = c(1,2,3,4))
```

---

threeCol2MaxMat   *threeCol2MaxMat*

Description

This function takes a three vectors of equal length (source nodes, target nodes, and edge weights) and return the adjacency matrix as a sparse Matrix.

Usage

```
threeCol2MaxMat(a = c("a", "b", "c", "c"), b = c("a", "b", "b", "b"),
  v = c(1, 2, 3, 4))
```

Arguments

a	(vector): vector of source node names.
b	(vector): vector of target node names.
v	(vector): vector of edge weights names.

Value

sparce Matrix.

Examples

```
threeCol2MaxMat(a = c("a","b","c","c"), b = c("a","b","b","b"), v = c(1,2,3,4))
```

---

transmat   *Sample transition matrix.*

Description

A Matrix containing the normalized transition matrix from the test network

Usage

```
transmat
```

Format

a Matrix containing the normalized transition matrix from the test network.

Index