Package ‘moc.gapbk’

March 7, 2019

Type Package
Title Multi-Objective Clustering Algorithm Guided by a-Priori Biological Knowledge
Version 0.1.0
Date 2019-03-03
Author Jorge Parraga-Alava [aut, cre]
          (<https://orcid.org/0000-0001-8558-9122>),
          Marcio Dorn [aut],
          Mario Inostroza-Ponta [aut]
Maintainer Jorge Parraga-Alava <jorge.parraga@usach.cl>
Description Implements the Multi-Objective Clustering Algorithm Guided by a-Priori Biological Knowledge (MOC-GaPBK) which was proposed by Parraga-Alava, J. et. al. (2018) <doi:10.1186/s13040-018-0178-4>.
Depends R (>= 3.2.5)
License GPL-2
Encoding UTF-8
Imports stats, amap, mco, nsga2R, reshape2, miscTools, matrixStats,
          fields, Rmisc, plyr, foreach, parallel, doParallel, utils,
          doSNOW, doMPI
RoxygenNote 6.1.1
Suggests knitr, rmarkdown
LazyData true
NeedsCompilation no
Repository CRAN
Date/Publication 2019-03-07 17:20:03 UTC

R topics documented:
moc.gabk ................................................. 2

Index 4

1
moc.gabk  

*Perform the Multi-Objective Clustering Algorithm Guided by a-Priori Biological Knowledge (MOC-GaPBK)*

**Description**

This function receives two distance matrices and it performs the MOC-GaPBK.

**Usage**

```r
moc.gabk(dmatrix1, dmatrix2, num_k, generation = 50, pop_size = 10, 
          rat_cross = 0.8, rat_muta = 0.01, tour_size = 2, 
          neighborhood = 0.1, local_search = FALSE, cores = 2)
```

**Arguments**

- `dmatrix1`: A distance matrix. It should have the same dimensions as `dmatrix2`. It is mandatory.
- `dmatrix2`: A distance matrix. It should have the same dimensions as `dmatrix1`. It is mandatory.
- `num_k`: The number k of groups represented by medoids in each individual. It is mandatory.
- `generation`: Number of generations to be performed by MOC-GaPBK. By default 50.
- `rat_cross`: Probability of crossover. By default 0.80.
- `rat_muta`: Probability of mutation. By default 0.01.
- `neighborhood`: Percentage of neighborhood. A real value between 0 and 1. It is computed as `neighborhood*pop_size` to determine the size of neighborhood. By default 0.10.
- `local_search`: A boolean value indicating whether the local searches procedures (PR and PLS) are computed. By default `FALSE`.
- `cores`: Number of cores to be used to compute the local searches procedures. By default 2.

**Details**

MOC-GaPBK is a method proposes by Parraga-Alava, J. et. al. 2018. It carries out the discovery of clusters using NSGA-II algorithm along with Path-Relinking (PR) and Pareto Local Search (PLS) as intensification and diversification strategies, respectively. The algorithm uses as objective functions two versions of the Xie-Beni validity index, i.e., a version for each distance matrix (dmatrix1, dmatrix2). More details about this compute can be found in: <https://doi.org/10.1186/s13040-018-0178-4>. MOC-GaPBK yield a set of the best clustering solutions from a multi-objective point of views.
moc.gabk

Value

- **population**: The population of medoids including the objective functions values and order by Pareto ranking and crowding distance values.
- **matrix.solutions**: A matrix with results of clustering. Each column represents a clustering solution available in Pareto front.
- **clustering**: A list containing named vectors of integers from 1:k representing the cluster to which each object is assigned.

Author(s)

Jorge Parraga-Alava, Marcio Dorn, Mario Inostroza-Ponta

References


Examples

```r
#Generates a data matrix of dimension 50x20
library("amap")
library("moc.gapbk")
x <- matrix(runif(100*20, min = -5, max = 10), nrow=50, ncol=20)

#Compute two distance matrices
dmatrix1<- as.matrix(amap::Dist(x, method = "euclidean"))
dmatrix2<- as.matrix(amap::Dist(x, method = "correlation"))

#Performs MOC-GaPBK with 5 cluster
example<-moc.gabk(dmatrix1, dmatrix2, 5)
example$population
example$matrix.solutions
example$clustering
```
Index

*Topic **Crowding**
moc.gabk, 2

*Topic **GaPBK**
moc.gabk, 2

*Topic **NSGA-II**
moc.gabk, 2

*Topic **Pareto**
moc.gabk, 2

*Topic **Xie-Beni**
moc.gabk, 2

*Topic **distance**
moc.gabk, 2

*Topic **index**
moc.gabk, 2

*Topic **ranking**
moc.gabk, 2

*Topic **validity**
moc.gabk, 2

moc.gabk, 2