



HAL
open science

Partitioning Methods On Dissimilarity Matrices Set

Francisco F. A.T. de Carvalho, Yves Lechevallier

► **To cite this version:**

Francisco F. A.T. de Carvalho, Yves Lechevallier. Partitioning Methods On Dissimilarity Matrices Set. European Conference on Data Analysis, GfKI and SFC, Jul 2013, Luxembourg, Luxembourg. hal-00916906

HAL Id: hal-00916906

<https://inria.hal.science/hal-00916906>

Submitted on 10 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Partitioning Methods On Dissimilarity Matrices Set

Francisco de A. T. de Carvalho¹ & Yves Lechevallier²

¹ CIn/UFPE, Recife-PE, Brésil fatc@cin.ufpe.br

² INRIA, Paris-Rocquencourt - 78153, France Yves.Lechevallier@inria.fr

Abstract. Cluster analysis have been widely used in numerous fields including pattern recognition, data mining and image processing. Their aim is to organize a set of items into clusters such that items within a given cluster have a high degree of similarity. In particular, partitioning clustering models aims to organize a set of items into a pre-defined number of clusters. Our reference clustering model is the k-means like algorithms .

In this presentation, we introduce partitioning clustering models and algorithms that are able to partitioning objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices. The aim is to obtain a collaborative role of the different dissimilarity matrices in order to obtain a final consensus partition. These matrices have been generated using different sets of variables and dissimilarity functions. These methods are designed to furnish a partition and a prototype for each cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fitting between the clusters and their representatives.

Experiments with data sets (synthetic and from UCI machine learning repository) show the usefulness of the proposed methods.

References

- [1] DE CARVALHO, F. A. T. , LECHEVALLIER, Y. and DE MELO, F. M.(2012): Partitioning hard clustering algorithms based on multiple dissimilarity matrices. *Pattern Recognition* 45, 447-464.
- [2] PEDRYCZ, W.(2002): Collaborative fuzzy clustering. *Pattern Recognition Letters* 23, 675-686.
- [3] FRIGUI, H., HWANGA, C. and RHEE, F. C.-H.(2007): Clustering and aggregation of relational data with applications to image database categorization. *Pattern Recognition* 40, 3053-3068.

Keywords

PARTITIONING, MULTIPLE DISSIMILARITY MATRICES