

4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Objective function-based clustering via near-Boolean optimization

Giovanni Rossi

University of Bologna, Italy

Objective function-based clustering is here looked as a maximum-weight set partitioning combinatorial optimization problem, with the instance given by a pseudo-Boolean (set) function assigning real-valued cluster scores (or costs, in case of minimization) to data subsets, while on every partition of data the global objective function takes the value given by the sum over clusters (or blocks) of their individual score. The instance may thus maximally consist of $2n$ reals, where n is the number of data, although in most cases the scores of singletons and pairs also fully determine the scores of larger clusters, in which the pseudo-Boolean function is quadratic. This work proposes to quantify the cluster score of fuzzy data subsets by means of the polynomial MLE (multi-linear extension) of pseudo-Boolean functions, thereby translating the original discrete optimization problem into a continuous framework. After analyzing in these terms, the modularity maximization problem, two further examples of quadratic cluster score functions for graph clustering are proposed, while also analyzing greedy (local and global) search strategies.

Biography

Giovanni Rossi completed his PhD in Economics and joined Computer Science department at University of Bologna. His research interests include "Discrete applied mathematics, combinatorial geometries, lattice and pseudo-Boolean functions, network community structure and graph clustering".

giovanni.rossi6@unibo.it

Notes: