Mach Learn
DOI 10.1007/s10994-013-5350-y

A flexible cluster-oriented alternative clustering algorithm for choosing from the Pareto front of solutions

Duy Tin Truong · Roberto Battiti

Received: 25 May 2012 / Accepted: 3 April 2013

Abstract Supervised alternative clustering is the problem of finding a set of clusterings which are of high quality *and* different from a given *negative* clustering. The task is therefore a clear multi-objective optimization problem. Optimizing two conflicting objectives at the same time requires dealing with trade-offs. Most approaches in the literature optimize these objectives sequentially (one objective after another one) or indirectly (by some heuristic combination of the objectives). Solving a multi-objective optimization problem in these ways can result in solutions which are dominated, and not Pareto-optimal. We develop a direct algorithm, called **COGNAC**, which fully acknowledges the multiple objectives, optimizes them directly and simultaneously, and produces solutions approximating the Pareto front. **COGNAC** performs the recombination operator at the *cluster level* instead of at the object level, as in the traditional genetic algorithms. It can accept arbitrary clustering quality and dissimilarity objectives and provides solutions dominating those obtained by other state-of-the-art algorithms. Based on **COGNAC**, we propose another algorithm called **SGAC** for the sequential generation of alternative clusterings where each newly found alternative clustering is guaranteed to be different from all previous ones. The experimental results on widely used benchmarks demonstrate the advantages of our approach.

Keywords Alternative clustering · Multi-objective optimization · Cluster-oriented recombination · Genetic algorithms

1 Introduction

Given a dataset, traditional clustering algorithms often only provide a single set of clusters, a single view of the dataset. On complex tasks, many different ways of clustering exist,

Editors: Emmanuel Müller, Ira Assent, Stephan Günnemann, Thomas Seidl, and Jennifer Dy.

D.T. Truong (✉) · R. Battiti
Department of Information Engineering and Computer Science, University of Trento, Trento, Italy
e-mail: truong@disi.unitn.it

R. Battiti
e-mail: battiti@disi.unitn.it

Published online: 24 April 2013