Invited Talk

Constraint Programming for Constrained Clustering

Christel Vrain

Université d'Orléans
INSA Centre Val de Loire, LIFO EA 4022
45067 Orléans, France
christel.vrain@univ-orleans.fr

Abstract. Several works have shown the interest of declarative frameworks, such as Constraint Programming, SAT, Integer Linear Programming, for Data Mining [9,10,12,8,14,15]. Relying on Constraint Programming (CP) has several advantages: first, its inherent declarativity allows to easily model constraints on the Data Mining problem at hand, second, CP has the ability either to enumerate all the solutions or to find an optimal solution, in case of an optimization problem. Moreover, it relies on constraint propagation, which often allows to efficiently prune the search space.

I will mainly focus on constrained clustering [16,7]: user constraints are put on the solutions, thus allowing to get a clustering closer to the one expected by the user.

After giving some backgrounds on CP, I will present the seminal work of [9] on CP for itemset mining, its extension to k-pattern set mining under constraints [13] and its application to conceptual clustering.

The talk will then mainly be focused on distance-based constrained clustering. I will show how we have modeled this task in CP [1,4,3], the difficulties we have had to face, the solutions we have developed [2,5].

The interest of relying on CP will be illustrated through several extensions [4,6,11].

The work on distance-based constrained clustering in CP is a joint work with T.B.H. Dao and K.C. Duong.

References


2. Thi-Bich-Hanh Dao, Kanh-Chuong Duong, and Christel Vrain. A Filtering Algorithm for Constrained Clustering with Within-Cluster Sum of Dissimilarities