Title: Ant Colony Optimization with Partial-Complete Searching for Attribute Reduction

Author: Fan Min Zhi-Heng Zhang Ji Dong

PII: S1877-7503(16)30460-4
DOI: http://dx.doi.org/doi:10.1016/j.jocs.2017.05.007
Reference: JOCS 676

To appear in:

Received date: 13-12-2016
Revised date: 7-4-2017
Accepted date: 5-5-2017

Please cite this article as: Fan Min, Zhi-Heng Zhang, Ji Dong, Ant Colony Optimization with Partial-Complete Searching for Attribute Reduction, *<![CDATA[Journal of Computational Science]]>* (2017), http://dx.doi.org/10.1016/j.jocs.2017.05.007

Ant Colony Optimization with Partial-Complete Searching for Attribute Reduction

Fan Min[a], Zhi-Heng Zhang[b], Ji Dong[a]

[a]*School of Computer Science, Southwest Petroleum University, Chengdu 610500, China*
[b]*School of Science, Southwest Petroleum University, Chengdu 610500, China*

Abstract

The time-cost-sensitive attribute reduction problem is more challenging than the classical reduct problem since the optimal solution is sparser. Ant colony optimization (ACO) is an effective approach to this problem. However, the efficiency is unsatisfactory since each ant needs to search for a complete solution. In this paper, we propose a partial-complete searching technique for ACO and design the APC algorithm. Partial searching is undertaken by pioneer ants through selecting only a few attributes to save time, while complete searching is undertaken by harvester ants for complete solutions. Experiments are undertaken on seven real-world and a set of artificial datasets with various settings of costs. Compared with two bio-inspired and two greedy algorithms, APC is more efficient while obtaining the same level of quality metrics. The APC algorithm can be also extended for other combinatorial optimization problems.

Keywords: Ant colony optimization; Attribute reduction; Cost; Heuristic algorithm; Partial-complete searching.

1. Introduction

Attribute reduction, also called feature selection, is a combinatorial optimization problem. The task is to find one or all optimal reducts from the $2^m - 1$ non-empty attribute subsets, where m is the number of attributes. A number of attribute reduction problems exist for different situations. Unsupervised attribute reduction [11, 32, 37] is required for information systems, while supervised attribute reduction [5, 19, 26, 40, 47, 54] is applicable to decision systems. The data type include symbolic [15, 37], numeric [15, 16], fuzzy [27, 32, 55, 57], interval [7, 20, 24, 45], ordered [22, 49] among others. The constraint of attribute reduction include preserving the discernibility [35, 43, 53], the positive region [36], the conditional entropy [6, 40, 44], the class distribution [44, 59], the dominance relation [17, 51], etc. The optimization objective include

*Corresponding author. Tel.: +86 135 4068 5200.
Email: minfanphd@163.com.