

数据库与信息处理

## 基于密度与划分方法的聚类算法设计与实现

孟海东, 宋飞燕, 郝永宽

内蒙古科技大学 网络中心, 内蒙古 包头 014010

收稿日期 修回日期 网络版发布日期 2007-9-9 接受日期

**摘要** 在分析常用聚类算法的特点和适应性基础上提出一种基于密度与划分方法的聚类算法。该算法根据数据对象密度分布状态来自动确定聚类簇密度吸引中心点和聚类簇的初始划分;然后利用划分的方法,根据密度可达定义来寻找密度可达数据对象簇,从而完成数据对象簇的最终聚类。实验证明该算法能够很好地处理具有任意形状和大小的簇,能够有效地屏蔽噪声和离群点的影响和发现孤立点;同时也减小了输入参数对领域知识的依赖性。

**关键词** [数据挖掘](#) [聚类](#) [密度函数](#) [密度可达](#) [划分方法](#)

分类号

## Design and implementation of clustering algorithm based on density and partition method

MENG Hai-dong, SONG Fei-yan, HAO Yong-kuan

Network Center, Inner Mongolia University of Science and Technology, Baotou, Inner Mongolia 014010, China

### Abstract

A clustering algorithm based on density and partitioning method is presented according to the analysis of the strengths and weaknesses of traditional clustering algorithms. The algorithm can automatically locate the dense centers of clusters, and determine initial partitions of the clusters. On the basis of initial partitions of the clusters, density reachable clusters of data objects are found out by using partitioning method, and the final clusters are produced. The experimental results demonstrate that the algorithm can handle clusters of arbitrary shapes and sizes, minimize the influences of noise and deviation of data objects, and locate the outliers. At the same time, the algorithm can minimize the dependency of input numbers on specialist knowledge.

**Key words** [data mining](#) [clustering](#) [density function](#) [density reachable](#) [partitioning method](#)

DOI:

通讯作者 孟海东 [E-mail: pixiu\\_lyf@163.com](mailto:pixiu_lyf@163.com)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(1656KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

#### 服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

#### 相关信息

▶ 本刊中 [包含“数据挖掘”的相关文章](#)

▶ 本文作者相关文章

· [孟海东](#)

· [宋飞燕](#)

· [郝永宽](#)