

研究、探讨

异构属性数据的量子聚类方法研究

李志华, 王士同

江南大学 信息工程学院, 江苏 无锡 214122

收稿日期 2008-4-28 修回日期 2008-6-26 网络版发布日期 接受日期

摘要 研究了异构属性数据的聚类问题。通过挖掘样本中的结构信息, 用加权的Mahalanobis距离来度量异构样本的相异性; 根据分类属性数据的分布与粒子在量子势能场中的分布不平衡的相似性, 重写量子势能公式为距离量子势能的形式, 提出了一种新的异构属性数据量子聚类WMDQC算法。通过进一步集成该算法和AHC算法为WMDQCM聚类方法, 用AHC算法更高效地挖掘样本中有利于聚类的结构线索。实验结果表明, 方法具有比较优势, 显著地改善了聚类性能, 具有一定的实用价值。

关键词 [异构数据](#) [相异性度量](#) [Mahalanobis距离](#) [量子势能](#) [聚类算法](#)

分类号 [TP181](#)

Weighted mahalanobis distance-based quantum clustering approach for heterogeneous data

LI Zhi-hua, WANG Shi-tong

School of Information and Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract

The dissimilarity measure and clustering approach about the heterogeneous dataset are studied, and a Weighted Mahalanobis Distance-based Quantum Clustering (WMDQC) algorithm is presented in this paper. Data often do appear in homogeneous groups, the WMDQC utilizes the structural information to improve the clustering accuracy. Unlike the numeric data, categorical data are often unbalancedly distributed, whose distribution are often unrelated with their distance measure. These characteristics are very similar to the particle world in quantum mechanism, so the WMDQC ascertains the clustering centers by the rewriting quantum potential. Further, a WMDQC-based method WMDQCM is proposed, the WMDQCM mines the structural clue by the agglomerative hierarchical clustering AHC algorithm to construct the weight matrix. By presenting the above to the WMDQC, the final clustering results are obtained. The new WMDQCM exhibits its robustness to initialization and clustering capability to heterogeneous dataset. Experimental results compared with other methods demonstrate that the proposed method has promising performance.

Key words [heterogeneous data](#) [dissimilarity measure](#) [Mahalanobis distance](#) [quantum potential](#) [clustering algorithm](#)

DOI: 10.3778/j.issn.1002-8331.2009.23.018

通讯作者 李志华 ezhli@yahoo.com.cn

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(480KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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“异构数据”的相关文章](#)
- ▶ [本文作者相关文章](#)

- [李志华](#)
- [王士同](#)