

数据库、信号与信息处理

基于K-均值聚类和凝聚聚类的离群点查找方法

曾颖, 罗可, 邹瑞芝

长沙理工大学 计算机与通信工程学院, 长沙 410076

收稿日期 2008-6-2 修回日期 2009-4-22 网络版发布日期 2009-10-10 接受日期

摘要 离群点发现是数据挖掘研究的一个重要方面。根据数据流的特点, 给出了一种基于K-均值聚类和凝聚聚类的离群点发现方法, 先用K-均值聚类对数据流进行处理, 生成中间聚类结果, 然后用凝聚聚类对这些中间结果进行再次选择, 最后找出可能存在的离群点。

关键词 [数据挖掘](#) [离群点](#) [K-均值聚类](#) [凝聚聚类](#)

分类号 [TP311](#)

Outliers detection method based on K-means and agglomerative clustering

ZENG Ying, LUO Ke, ZOU Rui-zhi

College of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410076, China

Abstract

Outliers detection is an important issue in data mining. In this paper, according to the characteristics of data streams, an outliers detection method based on k-means and agglomerative clustering is proposed, which uses k-means clustering to find some intermediate results, and applies agglomerative clustering on each intermediate result to pick out potential outliers.

Key words [data mining](#) [outliers](#) [k-means clustering](#) [agglomerative clustering](#)

DOI: 10.3778/j.issn.1002-8331.2009.29.039

通讯作者 曾颖

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(618KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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](#)

- [曾颖](#)
- [罗可](#)
- [邹瑞芝](#)