

## 基于新的成员选择方法的聚类融合算法

冯中慧<sup>1</sup>,何亮<sup>1</sup>,王栋<sup>1,2</sup>

(1 西安交通大学 电子与信息工程学院, 陕西 西安 710049;

2 洋浦经济开发区经济发展局, 海南 海口 578001)

**摘要:** 聚类融合是聚类分析领域的一个研究热点, 它将针对同一问题产生的多个聚类成员(即聚类结果)融合成一个结果, 以提高聚类算法的鲁棒性和准确性. 在聚类融合过程中, 聚类成员的“质量”非常重要, 一些“质量”比较差的聚类成员会直接影响聚类融合的结果. 在深入研究聚类融合算法基础上, 提出新的加权 JP (Jaccard index-Precision) 聚类成员选择方法和基于该方法的聚类融合算法. 在多个不同数据集上的实验结果表明, 利用这种新的聚类成员选择方法可以有效地改善聚类融合算法结果的准确性和鲁棒性, 性能显著提高.

**关键词:** 聚类融合; 聚类成员选择方法; 聚类精度; 差异度

## Clustering Ensemble Algorithm Based on

## New Members Selection Method

FENG Zhong-hui<sup>1</sup>, HE Liang<sup>1</sup>, WANG Dong<sup>1,2</sup>

(1 School of Electronic and Information Engineering, Xi'an Jiaotong University,

Xi'an 710049, China; 2 Hainan Province Yangpu Economic Development Zone Economic and Development Department, Haikou 578001, China)

**Abstract:** Clustering ensemble is one attractive field of cluster analysis. Focus on improving the robustness and precision of clustering algorithms, clustering ensemble methods combine multiple clustering members (clustering results) of the same issue into one result. During the process of clustering ensemble, the quality of clustering members is very important, some poor clustering members will directly affect the results of clustering ensemble. In this paper, clustering ensemble algorithms are studied, clustering ensemble algorithms based on new clustering members selection method using weighted JP (Jaccard index-Precision) are proposed. Experimental results show that precision and robustness of clustering are both improved on different datasets. The clustering performance is improved significantly.

**Key words:** clustering ensemble; clustering members selection method; clustering precision; diversity

**作者简介:**

冯中慧 女, (1977-), 博士, 讲师. 研究方向为数据挖掘、自然语言处理.

何亮 (通讯作者) 男, (1975-), 博士, 讲师. 研究方向为数据挖掘、机器学习. E-mail: lhe@mail.xjtu.edu.cn.

王栋 男, (1988-), 硕士. 研究方向为数据挖掘.