

学校编码: 10384

分类号 _____ 密级 _____

学号: 23220081153403

UDC _____

厦 门 大 学

硕 士 学 位 论 文

基于改进投影法的模糊聚类算法及有
效性研究

Study of Fuzzy Clustering Algorithm and Validity Based on
Improved Projection Method

张 宏 威

指导教师姓名: 王周敬 教授

专 业 名 称: 系 统 工 程

论文提交日期: 2011 年 月

论文答辩时间: 2011 年 月

学位授予日期: 2011 年 月

答辩委员会主席: _____

评 阅 人: _____

2011 年 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

摘 要

数据挖掘(Data Mining, DM),就是从大量的无序数据中发现隐含的、有效的、有价值的、可理解的模式,进而发现有用的知识,并得到时间的趋向和关联,为用户提供问题求解层次的决策支持能力的过程。聚类作为数据挖掘的一个重要分支,已经被广泛的应用到群体决策、图像分割、模式识别等领域。

现实世界中事物之间的界限往往是模糊的,对事物的分类必然伴随着模糊性,因此产生了模糊聚类分析。在所有的模糊聚类算法中,模糊 c 均值聚类算法的理论最为完善,应用最为广泛。

作为一般模糊集的推广,直觉模糊集和区间直觉模糊集被证明能更加细腻地刻画和描述客观世界的模糊性本质。如何利用传统的模糊聚类算法的思想,去处理属性值为直觉或者区间直觉模糊数的聚类问题,已经成为新的研究热点。

鉴于此,本文主要对如下几个方面进行了研究:

1. 对属性值为数值型数据的聚类问题,提出一种基于改进投影法的模糊聚类算法。算法首先采用改进的投影方法,构造新的目标函数;然后根据拉格朗日乘数法推导出聚类的迭代公式;最后进行实验对比,讨论新算法性能。

2. 提出基于投影法与改进投影法的模糊聚类有效性函数,并采用不同的样本数据以及聚类算法实验,和基于距离及基于模糊划分的聚类有效性函数进行实验对比,说明新有效性函数的合理性和可行性。

3. 对属性值为直觉和区间直觉模糊数的、属性权重完全未知或者以一般模糊数给出的属性权重的聚类问题进行研究。提出了基于直觉模糊 c 均值聚类算法的有效性函数,和基于区间直觉模糊集的 c 均值聚类算法及有效性函数。并通过实验分析算法的可行性和有效性函数的性能。

关键词: 投影法; 模糊聚类; 直觉模糊集; 区间直觉模糊集; 有效性函数

厦门大学博硕士学位论文摘要库

ABSTRACT

Data Mining, It is a process to find implicit, valid, valuable and understandable model from a large number of disordered, and then discover useful knowledge, obtain time trends and associations, provide users with decision support capabilities of problem-solving level. As an important branch of data mining, clustering has been widely applied to group decision making, image segmentation and pattern recognition etc.

In the real word things often have fuzzy boundaries, so the classification of things must be accompanied by fuzziness. In all of the fuzzy clustering algorithms, fuzzy c means clustering algorithm has the most perfect theory and the most widely used.

As a general extension of fuzzy sets, intuitionistic fuzzy sets and interval-valued intuitionistic fuzzy sets are shown to depict and describe ambiguity of the objective nature of the world more delicate. How to use the traditional idea of fuzzy clustering algorithm, to deal with fuzzy clustering problem that the property value of intuitionistic or interval-valued intuitionistic fuzzy numbers, has become a new hotspot.

Views of these situations, following aspects were studied:

1. A clustering algorithm based on improved projection is been proposed to solve the clustering problem that attribute value is numerical data. Firstly, adopt the modified projection method to construct the new objective function; and then the iterative formula are derived by Lagrange multiplier method; finally, the performance of the new algorithm are been discussed by experiment.

2. Propose fuzzy clustering validity function based on projection and function based on improved projection, and use different samples and clustering algorithms to compare the effectiveness of the functions with functions based on distance and function based on fuzzy partition, indicate the rationality and feasibility of these function.

3. Study the clustering issues that the property value is intuitionistic number or

interval-valued intuitionistic number, the weight of property is unknown or fuzzy number. Propose a validity function based on intuitionistic fuzzy c means clustering algorithm, interval-valued intuitionistic fuzzy c means clustering algorithm and validity function. At last the effectiveness of the algorithm and functions are analyzed by experiments.

Key words: Projection method, Fuzzy clustering, Intuitionistic fuzzy sets, Interval-valued intuitionistic fuzzy sets, Validity function

厦门大学博硕士学位论文摘要库

目 录

第一章 绪 论	1
1.1 研究的意义	1
1.2 国内外研究现状	2
1.3 论文研究工作及组织结构	4
第二章 聚类算法基本知识介绍	7
2.1 聚类分析概述	7
2.2 各种聚类方法介绍	9
2.2.1 分层聚类法	10
2.2.2 分割聚类法	11
2.2.3 基于密度的聚类法	11
2.2.4 基于网格的聚类法	12
2.2.5 基于模型的聚类法	12
2.2.6 基于约束的聚类法	12
2.2.7 其他新发展的聚类法	13
2.3 模糊 C 均值聚类算法	14
2.3.1 FCM 算法基本步骤	15
2.3.2 模糊聚类有效性函数	17
2.4 本章小结	18
第三章 基于改进投影法的模糊聚类算法	19
3.1 投影法及改进投影法原理	19
3.1.1 投影法基本原理	19

3.1.2 改进投影法基本原理	20
3.2 基于改进投影法的聚类算法 (PFCM)	22
3.2.1 PFCM 算法描述及实现	22
3.2.2 算法分析	25
3.3 仿真实验	25
3.4 本章小结	30
第四章 基于投影法和改进投影法的模糊聚类有效性函数	31
4.1 基于投影法的有效性指标	31
4.2 基于改进投影法的有效性指标	32
4.3 实验仿真分析	33
4.4 本章小结	40
第五章 直觉和区间直觉模糊聚类算法及有效性分析	41
5.1 直觉和区间直觉模糊集理论及应用	41
5.2 直觉和区间直觉模糊集定义及运算法则	42
5.2.1 直觉模糊集定义及运算	42
5.2.2 区间直觉模糊集定义及运算	44
5.3 直觉模糊 c 均值聚类算法	46
5.3.1 直觉模糊聚类算法	46
5.3.2 直觉模糊聚类有效性函数	47
5.3.3 实验仿真与分析	48
5.4 区间直觉模糊 c 均值聚类算法	49
5.4.1 区间直觉模糊聚类算法	49

5.4.2 区间直觉模糊聚类有效性函数.....	51
5.4.3 实验仿真与分析.....	52
5.5 本章小结.....	52
第六章 全文总结和研究展望.....	53
6.1 工作总结.....	53
6.2 研究展望.....	53
参考文献.....	55
攻读硕士学位期间发表的论文.....	62
致 谢.....	63

厦门大学博硕士学位论文摘要库

CONTENTS

Chapter1 Introduction	1
1.1 Significance of Reseach	1
1.2 Reseaching Status	2
1.3 Structure and Layout	4
Chapter2 Basic of Clustering Algorithm	7
2.1 Clustering Brief Introduction	7
2.2 Introduction of Every Clusterings	9
2.2.1 Hierarchical Clustering Algorithm	10
2.2.2 Partitioning Clustering Algorithm	11
2.2.3 Density-Based Method	11
2.2.4 Grid-Based Method	12
2.2.5 Model-Based Method	12
2.2.6 Restrain-Based Method	12
2.2.7 Other New Methods	13
2.3 Fuzzy C-Means Clustering	14
2.3.1 Step of FCM	15
2.3.2 Validity Indexs of Fuzzy Clustering	17
2.4 Summary	18
Chapter3 Clustering Algorithm based on Improved Projection Method	19

3.1 Theroy of Projection Method and Improved Projection Method.....	19
3.1.1 Theroy of Projection Method.....	19
3.1.2 Theroy of Improved Projection Method	20
3.2 Clustering Algorithm based on Improved Projection Method (PFCM).22	
3.2.1 Describe and Achieve of PFCM.....	22
3.2.2 Analysis of Algorithm	25
3.3 Simulation Experiment.....	25
3.4 Summary.....	30
Chapter4 Validity Indexs Based on Projection Method and Improved Projection Method.....	31
4.1 Validity Indexs Based on Projection Method	31
4.2 Validity Indexs Based on Improved Projection Method.....	32
4.3 Simulation Experiment.....	33
4.4 Summary.....	40
Chapter5 Clustering Alogirthm and Validity Index of Intuitionistic and Interval-Valued Intuitionistic Fuzzy Sets	41
5.1 Theroy and Applications of Intuitionistic and Interval-Valued Intuitionistic Fuzzy Sets	41
5.2 Definitions and Rules of Intuitionistic and Interval-Valued Intuitionistic Fuzzy Sets	42
5.2.1 Definitions and Rules of Intuitionistic Fuzzy Sets	42
5.2.2 Definitions and Rules of Interval-Valued Intuitionistic Fuzzy Sets	44

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库