FUZZY C-MEANS ALGORITHM BASED ON PRETREATMENT OF SIMILARITY RELATIONTP

Zhiping Jia[1] , Chenghui Zhang[2] and Rupeng Sun[1]

[1]School of Computer Science and Technology
Shandong University, Jinan, 250061,China
[2]School of Control Science and Engineering
Shandong University, Jinan, 250061,China

Abstract. In order to make up some deficiencies of the fuzzy c-means clustering algorithm, a new FCM algorithm based on pretreatment of similarity relation between samples is proposed in the paper, which is utilized to estimate the fuzzy clustering centers and the weight coefficient of samples effecting on the fuzzy clustering centers during iteration process. The new FCM algorithm makes the clustering quicker and more accurate. Finally, a simulation experiment is given in the paper to demonstrate the new FCM algorithm can avoid the limitations of the traditional FCM algorithm and the improvement is very effective.

Keywords. Fuzzy Set; Clustering Analysis; Fuzzy C-Means Algorithm; Similarity Relation.

AMS (MOS) subject classification: 62H30.

1 Introduction

Clustering is a process of dividing some datum into several categories, and the data object in the same category has more similarity. In the course of the clustering analysis, the fuzzy theory was used to instruct the clustering process, and so a new clustering algorithm-fuzzy c-means clustering algorithm(FCM) was proposed. The FCM algorithm classifies data objects into some different categories, and it can express the undefined degree of each object belonging to each category. Therefore, the FCM algorithm can reflect the actual instances more externally and gradually it is becoming the mainstream of clustering analysis and application([1,2,3,4]).

Many clustering methods based on the fuzzy theory have been proposed, such as the transitive closure clustering algorithm based on the fuzzy equal relationship, the largest tree algorithm based on fuzzy graph theory, and the clustering algorithm based on objective function and etc. The FCM algorithm is discussed in this paper, which belongs to an algorithm based on objective function. This clustering algorithm has a good real-time property and is suitable for the situation with a great deal of datum. Meanwhile, it is simply designed and easy to be realized on computer. So it has become the most popular clustering algorithm at present. Deep researches on the