Choosing DBSCAN Parameters Automatically using Differential Evolution

Abstract

Over the last several years, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) has been widely applied in many areas of science due to its simplicity, robustness against noise (outlier) and ability to discover clusters of arbitrary shapes. However, DBSCAN algorithm requires two initial input parameters, namely Eps (the radius of the cluster) and MinPts (the minimum data objects required inside the cluster) which both have a significant influence on the clustering results. Hence, DBSCAN is sensitive to its input parameters and it is hard to determine them a priori. This paper presents an efficient and effective hybrid clustering method, named BDE-DBSCAN, that combines Binary Differential Evolution and DBSCAN algorithm to simultaneously quickly and automatically specify appropriate parameter values for Eps and MinPts. Since the Eps parameter can largely degrades the efficiency of the DBSCAN algorithm, the combination of an analytical way for estimating Eps and Tournament Selection (TS) method is also employed. Experimental results indicate the proposed method is precise in determining appropriate input parameters of DBSCAN algorithm.
Choosing DBSCAN Parameters Automatically using Differential Evolution

- Tran Manh Thang and Juntatee Kim. The anomaly detection by using dbscan clustering with multiple parameters. In International Conference on Information Science and Applications (ICISA), pages 1–5, 2011.

Choosing DBSCAN Parameters Automatically using Differential Evolution

Choosing DBSCAN Parameters Automatically using Differential Evolution


Index Terms

Computer Science

Keywords

Clustering Analysis  DBSCAN  Differential Evolution  Tournament Selection