TELKOM UNIVERSITY

Abstract

School of Computing
The Graduate School

Master of Engineering

Spatial Clustering Based on Dissimilarity Region Using CLARANS with Polygon Dissimilarity Function

by Ichwanul Muslim Karo Karo

Most of unsupervised learning algorithms use a dissimilarity function to measures similarity between the objects within the dataset. However, traditionally dissimilarity functions did not design and fail to treat all spatial attributes of region or just solve partial kinds of region since incomplete representation of structural of region and other spatial information contained within the region datasets. In this research, we modified polygonal dissimilarity function (PDF) that comprehensively integrates both the spatial and the non-spatial attributes of a polygon to specifically consider the density and distribution that exist within the region datasets and work well to regular region, but not for irregular region. We represent a polygon as a set of intrinsic spatial attributes by slice vertices and structural region, extrinsic spatial attributes, and non-spatial attributes. Modified PDF was applied on cluster validity method by Davies Bouldin (DB) and spatial clustering by using CLARANS. Spatial clustering by CLARANS with modified PDF using two characteristically different sets of data, (a) regular geometric shapes (dummy region) and (b) irregular geometric shapes, Jakarta crime as case study on spatial clustering. Modified PDF is working and does not has member of disjoint cluster for more unstructured regions compared to origin PDF, has two percents smallest SSE than PDF, and accurately 36 percent than PDF on spatial clustering. Completely spatial information has above fifty percents significances and best cluster result for all dataset. In addition, modified PDF with DB (DB^P) can evaluated result of spatial region clustering than Silhouette index.