Mining Significant Co-Location Patterns at Multiple Distances

Abstract : Data that defines a location on earth surface is called spatial data. Typically, spatial data is represented using coordinates, shape, size and orientation. A co-location pattern is a subset of boolean spatial features whose instances are frequently located together. To identify a co-location pattern, we have some existing works like association rule mining, statistical representations using k-function, multi resolution pruning techniques to eliminate invalid patterns, instance lookup schemes etc. Identifying a valid co-location pattern includes two constraints namely, Distance threshold and Prevalence threshold. Euclidean distance was considered to find the distance between two co-location patterns, but this euclidean distance is a straight line (single scale) distance which might not be appropriate to mine all kind of patterns. Hence, we propose a multi-scale approach to find significant co-location patterns.

Key words : Boolean Spatial Features, Co-location Patterns, Distance Threshold, Prevalence Threshold, Euclidean Distance, Multi Scale.