Modeling Spatial Dependencies for Mining Geospatial   Data: PLUMS Approach

Sanjay Chawla , Shashi Shekhar, Weili Wu

Contact author: Weili Wu (wuw@cs.umn.edu)

University of Minnesota

Minneapolis, MN 55455, USA

Geo-spatial data mining is a process to discover   interesting and potentially useful spatial patterns embedded in spatial databases. Efficient tools for extracting information from geo-spatial data sets   can be of Importance to organizations which   own, generate and manage large geo-spatial data sets. The current approach towards solving spatial data mining problems is to use classical data mining tools after "materializing" spatial relationships and assuming independence between different data points. However, the key property of spatial data is that of spatial autocorrelation. Like temporal data, spatial data values are influenced by values in their immediate vicinity. Ignoring spatial autocorrelation in the modeling process leads to results which are a poor-fit and unreliable. In this paper we will propose PLUMS(Predicting Locations Using Map Similarity), a new approach for supervised spatial data mining problems. PLUMS searches the space of solutions using a map-similarity measure which is more appropriate in the context of spatial data. We will show that compared to state-of-the-art spatial statistics approaches, PLUMS achives comparable accuracy but at a fraction of the computational cost. Furthermore, PLUMS provides a general framework for specializing   other data mining techniques for mining spatial data.