LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-ABS-547781

A heuristic for the local region covering problem

Y. J. Fan, C. Kamath

April 12, 2012

21st International Symposium on Mathematical Programming
Berlin, Germany
August 19, 2012 through August 24, 2012

A Heuristic for the Local Region Covering Problem

Ya Ju Fan and Chandrika Kamath
Lawrence Livermore National Lab

Abstract

A variation of set covering problem is formulated to solve a local region problem that can be used to estimate the intrinsic dimensionality of multidimensional datasets. Each local region contains k nearest neighbors of a data point. The cost of selecting a region is the sum of the distances from the point to its nearest neighbors. A desired set of local regions should cover all points with the minimum cost. Hence the estimation is obtained based on a fair selection of local regions. We reformulate the set covering utilizing the distance matrix and the k nearest neighbors and obtain an equivalent set covering problem. This problem can be seen as a facility problem with the number of facilities being optimized in order to maintain the service to at most k cities per facility. To solve the equivalent problem, we present a simple and easy to implement heuristic method that is a variant of the greedy algorithm.