© 2013 FIZ Karlsruhe

io-port 01997943

Guo, Diansheng; Peuquet, Donna J.; Gahegan, Mark

ICEAGE: Interactive clustering and exploration of large and high-dimensional geodata.

GeoInformatica 7, No. 3, 229-253 (2003).

Summary: The unprecedented large size and high dimensionality of existing geographic datasets make the complex patterns that potentially lurk in the data hard to find. Clustering is one of the most important techniques for geographic knowledge discovery. However, existing clustering methods have two severe drawbacks for this purpose. First, spatial clustering methods focus on the specific characteristics of distributions in 2- or 3-D space, while general-purpose high-dimensional clustering methods have limited power in recognizing spatial patterns that involve neighbors. Second, clustering methods in general are not geared toward allowing the human-computer interaction needed to effectively tease-out complex patterns. In the current paper, an approach is proposed to open up the "black box" of the clustering process for easy understanding, steering, focusing and interpretation, and thus to support an effective exploration of large and high dimensional geographic data. The proposed approach involves building a hierarchical spatial cluster structure within the high-dimensional feature space, and using this combined space for discovering multi-dimensional (combined spatial and non-spatial) patterns with efficient computational clustering methods and highly interactive visualization techniques. More specifically, this includes the integration of: (1) a hierarchical spatial clustering method to generate a 1-D spatial cluster ordering that preserves the hierarchical cluster structure, and (2) a density- and grid-based technique to effectively support the interactive identification of interesting subspaces and subsequent searching for clusters in each subspace. The implementation of the proposed approach is in a fully open and interactive manner supported by various visualization techniques.

Classification: H.2.8 J.2
Keywords: geographic knowledge discovery; spatial clustering and ordering; hierarchical subspace clustering; visualization and interaction
doi:10.1023/A:1025101015202