

Chapter 10

SPAM: An Effective and Efficient Spatial Algorithm for Mining Grid Data

Ritu Chauhan

Amity University, India

Harleen Kaur

Hamdard University, India

ABSTRACT

The tremendous increase in spatial database technology has created furious interest among the researchers to adopt new methodologies for discovery of interesting patterns among large databases. But the raw data gathered from various resources such as Geographic Information Systems (GIS), business organizations, medical databases, climatic, market survey, remote sensing and several other resources might consist of data, which can be relevant, irrelevant or noisy in nature. However, retrieval of patterns from such databases can lead to serious concerns, which can frame inconsistent or irrelevant futuristic results. To deal with such issues, feature selection techniques are adopted to remove irrelevant, redundant and noisy features. Our approach focuses on retrieval of effective and efficient spatial clusters from large number of medical databases. In this chapter, we have defined our novel framework SpaGrid and SPAM algorithm to retrieve clusters of variant shape and size from large databases. The application of our framework is used with spatial medical databases where the implementation details are discussed with Matlab 7.1.

DOI: 10.4018/978-1-4666-8465-2.ch010

INTRODUCTION

In past decades, there has been explosive growth in spatial database technology generated from various sources. Such databases are usually evolved from numerous domains such as medical technology, business oriented organizations, market based analysis, social analysis, science exploration, geographical information studies, and several other automation technologies for retrieval and storage of data. There are constant efforts laid down by researchers and scientists throughout the globe to retrieve hidden information from large spatial databases. However, spatial data analysis tends to be major necessity for discovery of undefined rules and patterns from large spatial databases.

In pervasive term Spatial databases comprise of spatial as well as non-spatial attributes. Spatial attributes consist of features related to specific areas or regions, whereas non-spatial attributes are height, weight, temperature, etc. of such areas. There exist general relationship characteristics between spatial and non-spatial attributes for extraction of interesting spatial patterns from raw data (Lui & Han, 1993). Spatial data consists of data related to medical images which include thematic maps of specific area, remote sensing and several other sources where spatial data is bonded with non-spatial attributes of specific location (Lui & Han, 1993; Rigaux & Scholl, 2002; Zaiane, 2002). However, the discrete nature of spatial data can be represented in different form such as lattice data which consists of regular or irregular data units, the second representation of spatial data includes point pattern data which occurs in space at different location and finally geostatistical data which is associated with continuous variation of data in space (Cressie, 1991).

There exists different representation of spatial data in space such as raster and vector representation. The Raster data representation divides the space into rectangular cells and each cell corresponds to some value such as maximum, minimum or entire value of cell is made into consideration. They are represented as points with respect to collection of area in specific space, whereas vector data is represented by line, point and polygon of specific area. Point belonging to these spatial objects represents spatial data in space. The comparison between raster and vector data is made, raster data is applied with simpler data structure as spatial objects are divided into smaller subsets. Vector data can be easily stored, transformed and rotated hence can efficiently store without the high memory requirement. Therefore, it was found that vector data tends to work more efficiently than raster data.

Such databases are multidimensional in nature; hence proper spatial data structures are required for storage and retrieval of raw data. Therefore, discovery of knowledge should be formed as an automated process for retrieval of information from spatial databases (Seeger & Kriegel, 1988).

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/spam/136107

Related Content

Geospatial Resource Integration in Support of Homeland Defense and Security

David Foster and Christopher Mayfield (2016). *International Journal of Applied Geospatial Research* (pp. 53-63).

www.irma-international.org/article/geospatial-resource-integration-in-support-of-homeland-defense-and-security/160759

Characterization of Fire Regime Descriptors in Botswana Using Remotely Sensed Data

John Isaac Molefe (2018). *Handbook of Research on Geospatial Science and Technologies* (pp. 86-100).

www.irma-international.org/chapter/characterization-of-fire-regime-descriptors-in-botswana-using-remotely-sensed-data/187718

Alternative Methods for Developing and Assessing the Accuracy of UAV-Derived DEMs

Dion J. Wiseman and Jurjen van der Sluijs (2019). *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 1175-1196).

www.irma-international.org/chapter/alternative-methods-for-developing-and-assessing-the-accuracy-of-uav-derived-dems/222942

Campaign Optimization through Mobility Network Analysis

Yaniv Altshuler, Erez Shmueli, Guy Zyskind, Oren Lederman, Nuria Oliver and Alex "Sandy" Pentland (2015). *Geo-Intelligence and Visualization through Big Data Trends* (pp. 33-75).

www.irma-international.org/chapter/campaign-optimization-through-mobility-network-analysis/136099

Geospatial Evaluation for Urban Agriculture Land Inventory: Roanoke, Virginia USA

Tammy E. Parece and James B. Campbell (2017). *International Journal of Applied Geospatial Research* (pp. 43-63).

www.irma-international.org/article/geospatial-evaluation-for-urban-agriculture-land-inventory/169736