

The Methods of Spatial Data Mining and Its Application

Jiandong Chen

Ningbo University of Technology, Ningbo 315000, China.

Abstract: With the arrive of big-data age, it is increasingly important to know about Spatial Data Mining. This paper is designed to offer effective reference about future development trend of the spatial data mining as well as its several common methods and their application.

Keywords: The Technology and Methods of Mining; Future Development Trend; Application

1. Analysis On Development Directions

There are three main future development directions of spatial data mining methods. First of all, optimize traditional methods. Recent years, many scholars come up with new effective methods for the open problems of traditional spatial data mining. For example. Second, strengthen research on spatial data. With the development of spatial data, more and more methods are launched in new areas such as uncertainty. For example, He Binbin put forward a new mining algorithm model based on EM and Apriori algorithm and it can help to improve the authenticity and objectivity of mining. Third, use cloud computing platform. Data mining system based on cloud computing not only has unparalleled processing speed on spatial data, but services for users transparently. Users can gain valuable knowledge, not having to understand the principles and procedure of the system if he know how to select proper algorithm.

2. Common Methods

2.1 Method of clustering and classification

Both clustering and classification divide data into a series of differentiated groups according to certain distance or similarity measures and then to find out the whole distribution pattern of the data collection. Clustering method can be divided into the clustering method based on block, the clustering method based on hierarchical, the clustering method based on density, the clustering method based on grid and the clustering method based on pattern by different calculation algorithm of cluster analysis. Its specific algorithm can be divided into K-means algorithm, K-medoids algorithm, Clara algorithm, Clarans algorithm and so on. The essential difference between clustering and classification is that the typical feature of the data is clear for classification while not for clustering.

2.2 Statistic analysis

Statistical method is a common method in spatial mining that makes use of the finite information or uncertain information of the spatial target for statistics and then analyses, forecasts and assesses the attribute, feature and spatial pattern of the spatial target. In typical statistic analysis, the data is hard to meet the assumption of statistical incoherence as spatial data generally gathers on highly correlated and similar object. In addition, researchers are highly required because professional statistical knowledge or knowledge of corresponding field is necessary and this method generally is implemented by experts with statistical experience.

2.3 Method based on spatial association rule

Method based on spatial association rule is a method that mines, extracts or gets the relationships between the data in the source of spatial data. Actually it is a kind of algorithm and first put forward by Agrawal as Apriori algorithm. The core of the

Apriori algorithm is a recursive algorithm based on two-phase frequent item set thought, belonging to one-dimensional, one-layer and boolean association rule. In order to improve the performance of the Apriori algorithm, Jiawei Han and his team put forward the FP-growth algorithm based on frequent item set generated by FP-tree. Later, Jihua Cheng, Longfei Xu and others put forward multilayer association rule algorithm and general association rule model.

2.4 Induction learning

Induction learning is a method that generalizes and extracts general rules and patterns from the sea of empirical data. Most of algorithms of the Induction learning come from the field of machine learning. There are a lot of algorithms of induction learning, for example, AQ11 and AQ15 of Michaski, AE1 and AE9 of Jiarong Hong, CLS of Hunt, ID3 of Quinlan and so on.

2.5 Method of spatial analysis

Method of spatial analysis is a method that use various models of spatial analysis and spatial operation to handle the data of spatial database for deep processing. This can help to find out the association rule of the target on spatial connection, adjacency and symbiosis, or the knowledge to assist decision-making between targets like shortest path and optimal path.

3. Application

3.1 Retail industry of commodity

Early in 1994, the association rule among various market produces was discovered in sales transaction database, the placement regulation of goods and storage rack, the determination of the volume of purchases of goods and the application of digital network platform are all involved in and take advantage of the association rule in data mining.

3.2 Geological prospecting

Data mining method is mainly used to forecast geological minerals in the field of geology and mineral. Because of the diversity, difference and unique spatial properties of mineral data, when exploring the texture and structure of underground rock formations, not only mathematical model is needed, but also the spatial mining method is to collect related geological model from the sea of spatial database and then analyse, mine and transform. This allows spatial mining method achieves good effect in geological prediction.

4. Concluding remarks

Spatial data is various in forms and carries huge value. It has many features like multidimensional property, complexity, spatiality, scalability, and uncertainty. Its common methods are clustering, classification, statistic analysis, method based on spatial association rule, induction learning, method of spatial analysis and so on. It has been applied to some fields like retail industry of commodity and geological prospecting.

References

- [1] Xie YF, Liu Y, Li HJ, A review of spatial data mining methods [J]. Global Positioning System, 2010, 35(05): 65-68+75.
- [2] Du ZY, Tang WX, Deng K, Research on Spatial Data Mining Technology and Application [J]. Engineering Technology Research, 2018(06):67-68.
- [3] Du JY, Bian FL, Overview of Spatial Data Mining for Big Data [J]. Geospatial Information, 2017, 15(01): 8-12.
- [4] Cui GF, Research on statistical methods and their application in data mining [D]. Southwest Petroleum University, 2014.
- [5] Wang L, Application Research of Spatial Data Mining Technology Based on GIS [J]. Surveying and Mapping and Spatial Geographic Information, 2013, 36(06): 121-123.
- [6] Wang YM, Application Research of Spatial Data Mining Technology [J]. Computer CD-ROM Software and Application,

2014, 17(21): 310+312.

[7] Liu DJ, Overview of the development of spatial big data mining technology [J]. Information and Computer (Theoretical Edition), 2018(09):138-139.

[8] Gu RL, Analysis on the Application of Spatial Data Mining Technology [J]. Journal of Jilin Provincial Institute of Education (Middle), 2013, 29(05): 139-140.

[9] Lu XH, Research and application of spatial data mining technology based on GIS [D]. Jiangsu University of Science and Technology, 2013.

[10] Chen GF, Research and application of spatial data mining technology for precision agriculture [D]. Jilin University, 2009.

[11] Fu XJ, Application Technology of Spatial Data Mining [J]. Journal of China Environmental Management Cadre College, 2011, 21(02): 89-91.