Chapter 7
A Review on Spatial Big Data Analytics and Visualization

Bangaru Kamatchi Seethapathy
VIT University, India

Parvathi R
VIT University, India

ABSTRACT

Spatial dataset, which is becoming nontraditional due to the increase in usage of social media sensor networks, gaming and many other new emerging technologies and applications. The wide variety of sensors are used in solving real time problems like natural calamities, traffic analysis, analyzing climatic conditions and the usage of GPS, GPRS in mobile phones all together creates huge amount of spatial data which really exceeds the traditional spatial data analytics platform and become spatial big data .Spatial big data provide new demanding situations for their size, analysis, and exploration. This chapter discusses about the analysis of spatial data and how it gets descriptive manipulation, so that one can understand how multi variant variables get interact with each other along with the different visualization tools which make the understanding of spatial data easier.

INTRODUCTION

The data generation becomes easier in these modern days, for example taxi booking which stores the traveler location and the frequent routes travelled by the user and much more information. Handling of those huge data and making it more productive become the most difficult task. Thus "big data" comes into the picture, the data

which is bigger in both size and complexity can be analyzed and processed using different big data methodologies. The sources of big data are originating from different fields like social networks, sensor information, messaging system and much more. One of the most interesting things about big, data is that it can handle unstructured data which means that data which is not in pre-defined manner example photos, videos, voice recordings, etc. Interpreting such unstructured data shows big data as a new door of innovation and one among such unstructured data is the spatial data. Spatial data is the information or data that distinguishes the geographic area of elements and limits on Earth, such as natural or built components, seas, and many more. Spatial information is normally put away as directions and topology, and is information that can be mapped. Spatial information is regularly gotten to, control or broke down through Geographic Information Systems (GIS) (Aji, Wang, Vo et al., 2013) . Large percentage of data produced all over the world have the spatial component According to the recent survey, 80% of data is geographic which indicates the importance of handling geospatial big data especially societal applications like disaster management, disease surveillance, transportation monitoring and critical infrastructure but there are a huge number of constraints.Big data is characterized by the following(Chen & Zhang, 2014)

- **Volume:** Petabyte archives for remotely sensed imagery data, ever increasing volume of real-time sensor observations and location-based social media data, vast amount of VGI data, etc.., as well as the continuous increase of these data, raise not only data storage issues but also a massive analysis issue.
- **Variety:** Map data, imagery, data, geotagged text data, structured and unstructured data, raster and vector data, all these different types of data – many with complex structures – calls for more efficient models, structures, indexes and data management strategies and technologies.
- **Velocity:** Imagery data with frequent revisits at high resolution, continuous streaming of sensor observations, Internet of Things (IoT), real-time GNSS trajectory and social media data all require matching the speed of data generation and the speed of data processing to meet demand.
- **Veracity:** Much of geospatial big data are from unverified sources with low or unknown accuracy, level of accuracy varies depending on data sources, raising issues on quality assessment of source data and how to ''statistically'' improve the quality of analysis results.
- **Visualization:** Provides valuable procedures, to impose human thinking into the big data analysis. Visualizations help analysts identifying patterns (such as outliers and clusters), leading to new hypotheses as well as efficient ways to partition the data for further computational analysis. Visualizations

Related Content

Effect Analysis of Nursing Intervention on Lower Extremity Deep Venous Thrombosis in Patients

Xuanyue Zhang (2023). *International Journal of Data Warehousing and Mining (pp. 1-15).*

www.irma-international.org/article/effect-analysis-of-nursing-intervention-on-lower-extremity-deep-venous-thrombosis-in-patients/319948

A Look back at the PAKDD Data Mining Competition 2006

Nathaniel B. Norieland Chew Lim Tan (2007). *International Journal of Data Warehousing and Mining (pp. 1-11).*

www.irma-international.org/article/look-back-pakdd-data-mining/1780

Sarcasm Detection Using RNN with Relation Vector

Satoshi Hiaiand Kazutaka Shimada (2019). *International Journal of Data Warehousing and Mining (pp. 66-78).*

www.irma-international.org/article/sarcasm-detection-using-rnn-with-relation-vector/237138

Association Rule

 (2023). *Principles and Theories of Data Mining With RapidMiner (pp. 160-176).*

www.irma-international.org/chapter/association-rule/323373

Measuring Semantic-Based Structural Similarity in Multi-Relational Networks

Yunchuan Sun, Rongfang Bieand Junsheng Zhang (2016). *International Journal of Data Warehousing and Mining (pp. 20-33).*

www.irma-international.org/article/measuring-semantic-based-structural-similarity-in-multi-relational-networks/143713