Abstract

As stated in literature by several authors, there has been literally big-bang explosion in data acquired in recent times. This is especially so about the geographical or geospatial data. The huge volume of data acquired in different formats, structured, unstructured ways, having large complexity and non-stop generation of these data have posed an insurmountable challenge in scientific and business world alike. The conventional tools, techniques and hardware existing about a decade ago have met with the limitations in handling such data. Hence, such data are termed as big data. This has necessitated inventing new software tools and techniques as well as parallel computing hardware architectures to meet the requirement of timely and efficient handling of the big data. The field of data mining has been benefitted from these evolutions as well. This article reviews the evolution of data mining techniques over last two decades and efforts made in developing big data analytics, especially as applied to geospatial big data. This is still a very actively evolving field. There will be no surprise if some new techniques are published before this article appears in print.
References


Index Terms

Computer Science

Information Sciences

Keywords

Data mining, Distributed Computing, Hadoop, Big Data, Geospatial, Radoop, SpatialHadoop, Hadoop-GIS, Pigeon