

Preface

Data fundamentally serve in the development of a variety of knowledge bases, from understanding the real world to creating and enjoying the information world. Initially, humans used stones and shells to count according to the principle of one-to-one and kept records by tying knots. Later, picture notes were conceived to historically preserve more accurate records by using simple graphics along with a perceptual cue. Once pictures became relatively fixed by shape symbols and associated with words in the human language, text was created. The advent of text abstracted and generalized the world, promoted cultural understanding, and prepared the necessary foundation for the development of science. To break through the restrictions that the written symbols depended on, artificial transcription, carving or engraving, and human-operated machines were used after the industrial revolution to batch mechanization production and subsequently improve the efficiency of cultural transmission.

Thereafter, the computer was given the ability to listen to high-speed computing and spin off software within the hardware, which enabled the dissemination of information “electronically” and “automatically.” The internet then focused on networks by interrelating computers, essentially breaking through the local information limitation. Mobile communication released humans from their computers by enabling the computer to follow the user’s movements. The Internet of Things cares for the application of human-machine interoperation by automatically identifying objects, and allowing information sharing between humans and things. Cloud computing centralized these information services by consolidating the expertise and optimizing the allocation of resources. All the while, data have been continuously created and accumulated in big data centers that are closely associated with various applications for which spatial data mining is a necessity.

Spatial data account for the vast majority of data mining because most objects are now associated with their geo-spatial positions. Data mining attempts to discover the previously unknown, potentially useful, and ultimately understandable patterns of big data. The complex types, intrinsic relationships, and implicit auto-correlations in spatial big data make it more difficult to extract the useful patterns from spatial datasets than from conventional numeric and categorical datasets. To penetrate the volume, variety, velocity, and veracity of the values in big data, at

times it may be necessary for spatial data mining to take advantage of population data instead of sample data.

In this monograph, we present our novel theories and methods of spatial data mining as well as our successful applications of them in the realm of big data. A data field depicts object interactions by diffusing the data contribution from the universe of samples to the universe of population. The cloud model bridges the mutual transformation between qualitative concepts and quantitative data; and the mining view of spatial data mining hierarchically distinguishes the mining requirements with different scales or granularities. The weighted iteration method is used to clean spatial data of errors using the principles of post-variance estimation. A pyramid of spatial data mining visually illustrates the mining mechanism. All of these applications concentrate on the bottlenecks that occur in spatial data mining in areas such as GIS and remote sensing.

We were urged by scholars throughout the world to share our innovative interdisciplinary approach to spatial data mining; this monograph in English is the result of our 10-year foray into publishing our collective wisdom. The first Chinese edition was published in 2006 by Science Press and was funded by the National Foundation for Academy Publication in Science and Technology (NFAPST) in China. Simultaneously, we began writing an English edition. The first Chinese edition, meanwhile, was well received by readers and sold out in a short time, and an unplanned second printing was necessary. In 2013, writing a second Chinese edition was encouraged for publication in Science Press on the basis of our new contributions to spatial data mining. As a result, the English edition, although unfinished, was reorganized and updated. In 2014, Mr. Alfred Hofmann, the Vice President of Publishing for Springer, came to Beijing Institute of Technology personally to sign the contract for the English edition for worldwide publication. In 2015, the second Chinese edition won the Fifth China Outstanding Publication Award, which is a unique award by Science Press; Ms. Zhu Haiyai, the President of Publishing on Geomatics for Science Press, also personally wrote a long article on the book's publication process, which appeared in the Chinese Publication Newspaper. Following the criteria of contribution to the field, originality of the research, practicality of research/results, quality of writing, rigor of the research, substantive research and methodology, the data field method was awarded the Fifth Annual InfoSci®-Journals Excellence in Research Awards of IGI. The contributions were further reported by VerticalNews journalists, collected in *Issues in Artificial Intelligence, Robotics and Machine Learning* by ScholarlyEditions, *Developments in Data Extraction, Management, and Analysis* by IGI Global.

To finish writing the monograph in English as soon as possible, we made it a priority. We tried many translation methods, which taught us that only we, the authors, could most effectively represent the monograph. Although it took nearly 10 years to finish it, the final product was worth the wait!

Deren Li
Shuliang Wang
Deyi Li

Spatial Data Mining

Theory and Application

Li, D.; Wang, S.; Li, D.

2015, XXVIII, 308 p. 103 illus., 81 illus. in color.,

Hardcover

ISBN: 978-3-662-48536-1