

self-correlated. The economies of a region tend to be similar. A change in natural resources, wildlife, and temperature varies gradually from space to space. The spatial statistics study states, an area within statistics dedicated to the analysis of spatial data, then the property is called as spatial autocorrelation. For example, Figure 2 shows the value distributions of an attribute in a spatial framework for an independent identical distribution and a distribution with spatial autocorrelation.

In this review I have presented the features of spatial data mining that distinguish it from classical data mining in the following categories: description, input and statistical foundation.

I have also discussed major research accomplishments and techniques in spatial data mining, especially those related to four important output patterns: predictive models, spatial outliers, spatial correlation rules, and spatial clusters. I have also identified research needs and requirements for spatial data mining.

References

- [1] J. Han and M. Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2000.
- [2] Zhang, J. 2004. Polygon-based spatial clustering and its application in watershed study. MS Thesis, Department of Computer Science and Engineering, University of Nebraska-Lincoln, December 2004.
- [3] M. Ankerst, M. Breunig, H. -P. Kriegel, and J. Sander. OPTICS: ordering points to identify the clustering structure. In *Proc. 1999 ACM-SIGMOD Int. Conf. Management of Data (SIGMOD '99)*, Philadelphia, PA, June 1999.
- [4] Published in Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)-"A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise" by Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu
- [5] R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghavan. Automatic subspace clustering of high dimensional data for data mining applications. In *Proc. 1998 ACM-SIGMOD Int. Conf. Management of Data (SIGMOD '98)*, Seattle, WA, June 1998.
- [6] P. Bradley, U. Fayyad, and C. Reina. Scaling clustering algorithms to large databases. In *Proc. 1998 Int. Conf. Knowledge Discovery and Data Mining (KDD '98)*, pages, New York, NY, Aug. 1998.
- [7] Shashi Shekhar, Pusheng Zhang, Yan Huang, and Ranga Raju Vatsavai "Trends in spatial data mining" (Roddick and Spiliopoulou 1999, Shekhar and Chawla 2002)
- [8] S. Guha, R. Rastogi, and K. Shim. CURE: An efficient clustering algorithm for large data bases. . In *Proc. 1998 ACM-SIGMOD Int. Conf. Management of Data (SIGMOD '98)*, Seattle, WA, June 1998.
- [9] Hinneburg and D. A. Kein. An efficient approach to clustering in large multimedia databases with noise. In *Proc. 1998 Int. Conf. Knowledge Discovery and Data Mining (KDD '98)*, New York, NY, Aug. 1998.
- [10] J. Hou, a M.Sc. thesis submitted by Jean Fen-ju Hou, in the School of Computer Science, Simon Fraser University.
- [11] Y. Huang, S. Shekhar & H. Xiong, *Discovering Spatial Co-location Patterns from Spatial Datasets: A General Approach*, IEEE Transactions on Knowledge and Data Eng., 2004.
- [12] L. Kaufman and P.J. Rousseeuw. Finding groups in Data: an Introduction to Cluster Analysis. John Wiley & Sons, 1990.
- [13] R. Ng, and J. Han. Efficient and effective clustering method for spatial data mining. In *Proc. 1994 Int. Conf. Very Large Data Bases (VLDB '94)*, Santiago, Chile, Sept. 1994.
- [14] J. O'Rourke. *Computational Geometry in C (2nd Ed.)*. Cambridge University Press, 1998.
- [15] G. Sheikholeslami, S. Chatterjee, and A. Zhang. WaveCluster: A multi-resolution clustering approach for very large spatial databases. . *Very Large Data Bases, 1998*.
- [16] Silicon Graphics. Inc. BSP Tree: Frequently asked questions. <http://reality.sgi.com/bspfag/index.shtml>, 1997.
- [17] K. H. Tung, J. Hou, & J. Han. Spatial Clustering in the Presence of Obstacles. In *Proc. 2001 Int. Conf. Data Engineering (ICDE '01)*, Apr. 2001b.
- [18] T. Zhang, R. Ramakrishnan, and M. Livny. BIRCH: an efficient data clustering method for very large databases. In *Proc. 1996 ACM-SIGMOD Int. Conf. Management of Data (SIGMOD '96)*, Montreal, Canada, June 1996
- [19] Margrat M. Dunham, "Data mining introductory and advanced Topics", Pearson-2004.
- [20] Ji Zhang, Wynne Hsu, Mong Lee. "Clustering in Dynamic Spatial Databases", January-2005.
- [21] Emin Erkan, Ken Barker, "Intelligent Data Analysis", March-2006.
- [22] Wei-Bang Chen, "Spatial Clustering with Obs An efficient hybrid clustering Algorithm", March-2006.
- [23] Chuang-Cheng Chiu, Cheh-Yuan T sai, "A K-Anonymity clustering Method", Aug-2007.
- [24] Lian Duan, Li da Xu, Feng Guo, "A Local-density based Spatial Clustering Algorithm" Nov-2007.
- [25] Jens keller, 2008, "Clustering biological data using a hybrid approach".
- [26] Bin Jiang, Jing Chang Pan, "A Data Mining Application in Stellar Spectra", Dec2008.

Author Profile



V. Padmavati received the B.Sc(IT) and M.Sc(C.S.) degrees in IT and Computer Science from MCNUJC, Bhopal in 2003 and 2006, respectively. After wards she completed M.Phil(C.S.) in 2010 from C.V. Raman University, Bilaspur (C.G.). She is now in Pt. Ravishnakar Shukla University, Dep. Of comp. sci. & IT, Raipur(C.G.) as contract Assistant Professor.