计算机应用 2012, 32(03) 847-851 DOI: 10.3724/SP.J.1087.2012.00847 ISSN: 1001-9081 CN: 51-1307/TP



典型应用

# 基于DBSCAN聚类算法的闪电临近预报模型

侯荣涛[1,2],朱斌[1,2],冯民学[3],史鑫明[1,2],路郁[1,2]

1.南京信息工程大学 江苏省网络监控中心,南京 210044;
2.南京信息工程大学 计算机与软件学院,南京 210044;
3.江苏省防雷中心,南京 210009

**摘要:** 针对闪电定位仪中庞大而杂乱的定位数据,提出一种基于改进DBSCAN聚类算法(IDBSCAN)进行闪电聚类分析的方法。该方法依据闪电定位系统中的实时监控数据,搜索闪电密度大于阈值范围的地闪点,建立密度可达最大值的地闪聚类簇,并找到该簇类中的核心地闪点。同时,应用邻接表结构对DBSCAN算法进行改进,使得初始地闪数据的搜索集的建立时间和空间得到大大减少。在聚类分析结果基础上,对核心地闪点的移动路径进行拟合,从而预报下一时刻的核心地闪点位置。实验证明,将IDBSCAN算法应用在闪电临近预报中是有效的。

**关键词:** 闪电临近预报 定位资料 DBSCAN算法 邻接表 空间聚类

## Prediction model for lightning nowcasting based on DBSCAN

HOU Rong-tao[1,2],ZHU Bin[1,2],FENG Min-xue[3],SHI Xin-ming[1,2],LU Yu[1,2]

1.Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing Jiangsu 210044, China;
2.College of Computer and Software, Nanjing University of Information Science and Technology, Nanjing Jiangsu 210044, China;
3.Lightning Protection Center of Jiangsu Province, Nanjing Jiangsu 210009, China

Abstract: Against the massive monitoring data of lightning locating system, a lightning nowcasting model based on Improved Density-Based Spatial Clustering of Application with Noise (IDBSCAN) clustering algorithm was put forward. Based on the lightning location data in real-time monitoring system, this method searched for lightning-density flash point greater than the threshold value of the land, built the cluster with up to the maximum ground flash density, and located the core of the cluster. Besides, with the application of adjacency list search algorithm, time and space consumed for the initial search set of lightning data had been greatly reduced. Furthermore, using regression fitting algorithm, the proposed algorithm can predict the path of movement of lightning cluster. The experimental results show that IDBSCAN algorithm used in the lightning nowcasting is effective.

Keywords: lightning nowcasting location information DBSCAN algorithm adjacency list spatial clustering

**收稿日期** 2011-08-22 **修回日期** 2011-12-12 **网络版发布日期** 2012-03-01

DOI: 10.3724/SP.J.1087.2012.00847

**基金项目**:

国家自然科学基金资助项目(50875175);天津市自然科学基金资助项目(10JCZDJC23400);江苏高校优势学科建设工程项目。

**通讯作者**: 朱斌

**作者简介**: 侯荣涛(1957-),男,河北唐山人,教授,博士生导师,主要研究方向:混沌信息识别、特征提取、机器视觉、信号处理;朱斌(1987-),男,江苏常州人,硕士研究生,主要研究方向:信息安全、气象信息处理、数据挖掘;冯民学(1962-),男,江苏无锡人,正研级高级工程师,博士,主要研究方向:雷电监测与防护;史鑫明(1988-),男,江苏扬州人,硕士研究生,主要研究方向:数据挖掘、人工智能;路郁(1988-),男,江苏苏州人,硕士研究生,主要研究方向:数据挖掘。

**作者**Email: binzhu1987@163.com

**参考文献：**

[1]HAND D, MANNILA H, SMYTH P. Principles of data mining [M]影印版.北京: 机械工业出版社, 2003.

[2]周水庚,周傲英,曹晶,等.一种基于密度的快速聚类算法[J].计算机研究与发展, 2000, 37(11): 1187-1192.

[3]陈燕,耿国华,郑建国. 一种改进的基于密度的聚类算法[J]. 微机发展,2005,15(3):84-89.

[4]李霞,汪庆森,巩晴霞,等. 苏州地区雷电分布规律分析[J]. 气象科技,2006,26(4):443-449.

[5]李照荣,张强, 陈添宇,等. 一次强冰雹暴雨天气过程闪电特征分析[J]. 干旱区研究,2007,24(3):321-327.

[6]刘卫宁,曾婵娟, 孙棣华. 基于DBSCAN算法的营运车辆超速点聚类分析[J]. 计算机工程,2009,35(5):268-270.

[7]CHEN YAN-HUA, MANJEET REGE, DONG MING, et al. Non-negative matrix factorization for semi-supervised data clustering[J]. Knowledge and Information Systems,2008,17(3):355-379.

[8]TSAI CHENG-FA, YEH HENG-FU, CHANG JUI-FANG, et al. An efficient data clustering scheme using partition space technique for knowledge discovery in large databases[J]. Applied Intelligence,2010,33 (1): 39-53.

[9]WU YI-PU, GUO JIN-JIANG, ZHANG XUE-JIE.A linear DBSCAN algorithm based on LSH[C]// Proceedings of the Sixth International Conference on Machine Learning and Cybernetics. Piscataway, NJ: IEEE Press,2007:2608-2614.

[10]MARTIN E. A density-based algorithm for discovering clusters in large spatial databases with noise [C]// Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining. Portland, USA:AAAI Press,1996:226-231.

[11]MARZENA K R Y, PIOTR L A. TI-DBSCAN: Clustering with DBSCAN by means of the triangle inequality[C]// Proceedings of the 7th International Conference on Rough Sets and Current Trends in Computing. Berlin: Springer-Verlag,2010:60-69.

[12]DERYA BIRANT, ALP KU. ST-DBSCAN: An algorithm for clustering spatial-temporal data[J]. Data & Knowledge Engineering,2007,60(1):208-221.

[13]BORAH B, BHATTACHARYYA D K. An improved sampling-based DBSCAN for large spatial databases[C]// Proceedings of Intelligent Sensing and Information Processing. Piscataway, NJ: IEEE Press,2004:92-96.

[14]PEI JIAN, HAN JIA-WEI, MORTAZAVI-ASL B. PrefixSpan: Mining sequential patterns by prefix-projected growth[C]// Proceedings of the 17th IEEE International Conference on Data Engineering. Piscataway, NJ: IEEE Press,2001:215-224.

**本刊中的类似文章**

1．余莉 李佳田 李佳 段平 王华.二维空间聚类的树ART2模型[J]. 计算机应用, 2011,31(05): 1328-1330

2．李建国 胡学钢.高效的混合聚类算法及其在异常检测中的应用[J]. 计算机应用, 2010,30(07): 1916-1918

3．杨杨 孙志伟 赵政.一种处理障碍约束的基于密度的空间聚类算法[J]. 计算机应用, 2007,27(7): 1688-1691

4．赵红蕊 唐中实 .基于图像空间聚类的滤波技术[J]. 计算机应用, 2006,26(11): 2691-2693

5．王生生;刘大有;曹斌;刘杰.一种高维空间数据的子空间聚类算法[J]. 计算机应用, 2005,25(11): 2615-2617

6．吴安阳，赵卫东.基于多最小支持度的空间关联规则发现[J]. 计算机应用, 2005,25(09): 2171-2174

7．孙志伟，赵政.DBSCAN在非空间属性处理上的扩展[J]. 计算机应用, 2005,25(06): 1379-1381