引文格式:周相兵.GNSS 数据的智能聚类学习算法研究[J].测绘学报,2019,48(8):1072. DOI:10.11947/j.AGCS.2019.20180448.
ZHOU Xiangbing.Research on intelligent clustering learning algorithm for GNSS data[J]. Acta Geodaetica et Cartographica Sinica,2019,48(8):1072. DOI:10.11947/j.AGCS.2019.20180448.

GNSS 数据的智能聚类学习算法研究

周相兵[1,2]

1. 中国科学院水利部成都山地灾害与环境研究所数字山地与遥感应用研究中心,四川 成都 610041;
2. 四川旅游学院信息与工程学院,四川 成都 610059

Research on intelligent clustering learning algorithm for GNSS data

ZHOU Xiangbing[1,2]

1. Research Center for Digital Mountain and Remote Sensing Application Institute of Mountain and Environment, CAS, Chengdu 610041, China; 2. School of information and Engineering, Sichuan Tourism University, Chengdu 610059, China

随着 GNSS 终端设备的普及与广泛应用,海量的、带丰富位置信息的数据所隐藏的地质与空间变迁信息正在支撑智慧城市的发展。论文以 10 组不同城市出租车 GPS 数据为研究对象,以遗传算法、粒子群算法和蚁群算法 3 种智能算法为研究基础,以聚焦划分聚类算法为自动聚类的基本算法;提出了基于智能优化的 GPS 数据自动聚类学习算法,这些算法通过所构建的模糊系统和初始化种群技术,有效地克服了基于划分聚类算法的聚类数目不易确定、预设参数过多、敏感于初始种子点、难以将上一代优秀聚类结果保存到下一代、易陷入局部最优等长期以来存在的缺陷。试验验证表明,本文所提出的算法效率能提升 6%及以上,最高可达 20%,而且算法计算复杂度能控制在 O(n)左右(n 是 GPS 数据点的数目);能更好地发掘城市热点、人群聚集区、城市运行状态等有价值信息。因此,本文主要内容包括:

(1) 结合噪音、改进 Canopy、密度估计与 k-means++分别提出了 3 种遗传算法的初始化种群技术;提出了一种基于密度估计的共享小生境技术和改进型基因重排技术,用于避免遗传算法早熟现象且处理不等长的染色体;并结合自适应交叉、变异、精英操作实现遗传优化,找到最优染色体(个体)实现城市出租车 GPS 数据的 k-means 自动聚类。整个算法有效确保种群的多样性,达到全局优化。

(2) 将基于噪音的初始化种群方法、遗传算法与自适应模糊的粒子群算法进行融合,提出了一种新的模糊粒子群算法与遗传算法融合的自动划分聚类算法。一方面捕获到最优染色体实现 k-means 自动聚类;另一方面自动生成噪音半径和增加了模糊系统输出量,并将 k-means 融合到粒子群与遗传算法混合中实现 GPS 数据自动聚类。

(3) 将基于噪音的初始化种群方法、遗传算法与提出的自适应模糊蚁群系统进行全面混合,研究了一种遗传算法与蚁群系统相结合的自适应模糊蚁群 k-means 自动聚类算法。先用遗传操作获得最优染色体,后再进入蚁群系统算法中实现蚁群 k-means 自适应聚类,有效解决了 k-means 聚类敏感于初始化聚类中心和聚类陷入局部最优化缺陷。

(4) 在云计算环境下,提出了一种基于 MapReduce 的遗传 k-means 聚类算法,用于处理大体量 GPS 数据。算法以 Mahout 中的 Canopy、k-means 为基础,用采样频率改进 Canopy 和 k-means++(是基于 MapReduce 的)并使其产生适当聚类数目,生成遗传操作的初始化种群,实现大体量 GPS 数据自适应遗传自动聚类。

(5) 以出租车 GPS 数据特征为基础,提出了一种基于轨迹角度划分和余弦约束的 GPS 轨迹片段生成方法(含有 3 个 GPS 点的轨迹片段),再结合基于拉格朗日定理改进了 Fuzzy c-means 聚类算法,并对所生成的轨迹片段进行模糊聚类。同时,结合基于噪音的初始化种群方法和模糊系统,提出了模糊自适应遗传算法,并用于实现改进型 Fuzzy c-means 轨迹片段的自动聚类,最后采用最小二乘算法回归生成 GPS 平滑轨迹,为城市道路规划及路网更新提供了新方法。

中图分类号:P228 **文献标识码**:D

文章编号:1001-1595(2019)08-1072-01

基金项目:四川省科技计划(2016GZ0140;2018GZ0177);国家自然科学基金(61071121);四川省高校科研创新团队(15DT0039)

收稿日期:2018-10-08

作者简介:周相兵(1980—),男,2018 年 6 月毕业于成都理工大学,获工学博士学位(指导教师:苗放教授,马洪江教授),研究方向为空间数据机器学习算法研究。

Author: ZHOU Xiangbing (1980—), male, received his doctoral degree from Chengdu University of Technology on June, 2018, majors in spatial data machine learning algorithm.

E-mail: 3dsmaxmaya@163.com