

改进小波聚类算法在 QAR 数据中的应用

杨 慧, 李 振, 霍纬纲

(中国民航大学 计算机科学与技术学院, 天津 300300)

摘 要: 传统小波聚类算法标记满足密度阈值的连通单元为同一个簇, 而不满足密度阈值的网格有可能存在属于簇的数据对象, 数据的每维属性有时差距较大, 不合适再划分均匀网格。为此, 提出一种改进的小波聚类算法 CWaveCluster, 划分非均匀网格, 进一步细化边界网格, 对不满足密度阈值的网格进行处理, 最终形成聚类。在指定的快速存取记录器(QAR)数据集上的实验结果表明, 改进的小波聚类算法能根据数据特点划分网格, 区分簇与簇的边界, 有效解决 QAR 数据异常点检测问题。

关键词: 连通单元; 小波聚类; 边界网格; 快速存取记录器; 密度阈值

中文引用格式: 杨 慧, 李 振, 霍纬纲. 改进小波聚类算法在 QAR 数据中的应用[J]. 计算机工程, 2017, 43(9): 29-33, 38.

英文引用格式: YANG Hui, LI Zhen, HUO Weigang. Application of Improved Wavelet Clustering Algorithm in QAR Data[J]. Computer Engineering, 2017, 43(9): 29-33, 38.

Application of Improved Wavelet Clustering Algorithm in QAR Data

YANG Hui, LI Zhen, HUO Weigang

(School of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

[Abstract] Traditional wavelet clustering algorithm labels the communication unit satisfying density threshold as the same cluster, the mesh which does not meet the density threshold may have the data objects belonging to the cluster, and each dimension attribute of the data sometimes has a big gap, so that subdividing the mesh into uniform grid is not appropriate. Thus, an improved wavelet clustering algorithm is proposed. The method is used to divided the non-uniform grid, and refines further the boundary of the grid which does not satisfy the density threshold, and formats the clusters finally. By applying on the specified Quick Access Recorder (QAR) data sets, experimental results show that the improved wave cluster algorithm can effectiveing distinguish between cluster and boundary of the cluster, according to the characteristics of the data mesh, this method solves the question of the QAR data anomaly detection effectively.

[Key words] communication unit; wavelet clustering; border grid; Quick Access Recorder (QAR); density threshold

DOI: 10.3969/j.issn.1000-3428.2017.09.006

0 概述

快速存取记录器^[1] (Quick Access Recorder, QAR) 是目前国内航空公司普遍采用的记录飞机飞行过程中的一系列参数。QAR 数据涵盖了飞机飞行过程中的绝大多数参数, 主要用于日常机务维护、飞行检查、故障检测等需要。因此, 通过挖掘 QAR 数据隐含的飞机安全隐患并为飞机安全管理提供参考具有重大意义。

由于 QAR 数据集比较复杂, 数据量庞大, 故障的产生通常是多种属性共同作用的结果, 并且检测异常数据需要较高的精度, 再使用传统的异常点检

测算法很难获得理想的结果。

小波聚类算法^[2] 是一种基于密度和网格相结合的混合聚类算法, 该算法将数据看做多维信号来处理, 能够有效地处理大数据集, 发现任意形状的簇, 具有自动去噪、对输入的数据顺序不敏感等优点, 被广泛地应用于大数据集的处理。

目前, 国内外专家已经在小波聚类算法 (WaveCluster) 的研究和应用上取得了较大进展, 如文献[3]将小波聚类和 MPI 相结合将其应用到曙光 TC 1700 上; 文献[4]将小波聚类方法用于药片种类判别当中; 文献[5]将小波聚类算法应用到股票收益序列时间模式挖掘中; 文献[6]将小波聚类算法应用

基金项目: 国家自然科学基金(61301245); 国家自然科学基金与中国民航联合基金(61179063)。

作者简介: 杨 慧 (1957—), 女, 教授、博士, 主研方向为智能数据处理; 李 振, 硕士研究生; 霍纬纲, 副教授、博士。

收稿日期: 2016-08-15 **修回日期:** 2016-09-28 **E-mail:** yanghui_z@ sina. com

到遥感图像检索上。然而,到目前为止,小波聚类算法还没有被应用到 QAR 数据集异常点检测的研究中。

小波聚类算法的主要思想是通过小波变换转换特征空间,并在其中分别以不同尺度和不同分辨率寻找连通单元形成聚类。小波聚类算法会面临 MAUP(Modifiable Areal Unit Problem)问题,即可塑性面积单元问题,这也是其他网格类聚类算法通常会遇到的问题。所谓可塑性面积单元问题是指随着基本面积单元(网格)粒度大小选取的不同,聚类分析结果也会发生变化,同一粒度情况下选取不同的聚合方式分析结果也不尽相同。小波聚类另一个可能出现的问题是网格中的点有可能处于不同的簇中,同时由于网格类算法是对网格的计算代替对数据点的计算,即时间复杂度降低是以丢失大部分点的信息最终降低聚类精度的代价为前提,目前,对小波聚类算法的改进^[7-10]基本都是基于对网格的重新划分^[11-13],但是这些改进算法很少考虑到聚类簇的边界问题。

由此,本文提出一种改进小波聚类簇边缘精度的聚类算法,首先按照数据分布信息划分非均匀网格,并根据数据对象在网格中的分布划分稠密网格和稀疏网格,对稠密单元中的边界网格细化并通过寻找连通单元形成主聚类,在主聚类的基础上对稀疏网格根据距离和寻找连通单元进一步划分形成聚类。本文给定相关定义,对算法框架进行描述,分析其时间复杂度,并将其应用到 QAR 数据异常点检测中。

1 基本概念

定义 1 $A = A_1, A_2, \dots, A_d, P = A_1 \times A_2 \times \dots \times A_d$ 为一个 d 维特征空间,其中, A_1, A_2, \dots, A_d 为 P 的维数,且对于 $\forall a_{ij} \in A_i$, 有 $[a_{i\min}, a_{i\max}] \in A_i$, 其中, $a_{i\min}, a_{i\max}$ 分别为 A_i 属性的最小值与最大值。

定义 2 将 d 维特征空间分割成不重叠的长方形网格单元,即将每维 A_i 分割成相同数目且不重叠的长方形网格,数目为 m ,根据经验公式得:

$$m = (\text{int}) \sqrt[d]{N} + 1 \quad (1)$$

其中, N 为特征空间中数据的个数; d 为特征空间的维数,期望属性差为:

$$dec = \frac{\sum_{i=1}^n (a_{i\max} - a_{i\min})}{n} \quad (2)$$

其中, n 为特征空间的维数,则每维属性划分的网格数目为:

$$m_i = \frac{(a_{i\max} - a_{i\min})}{dec} \times m \quad (3)$$

定义 3 设密度阈值为 q , 则相对密度阈值为:

$$q_i = \frac{m_i}{m} \times q \quad (4)$$

定义 4 在特征空间中大于相对密度阈值 q_i 为稠密网格,小于相对密度阈值 q_i 并且大于 0 的为稀疏网格^[14]。

定义 5 引入 partition and judge the grid C_{ij} 算法,该算法细分网格并判断一个网格是否为边缘网格,主要思想是判断当前网格是否存在反方向单元,所谓反方向单元即满足密度阈值的网格单元邻域中存在着与该网格单元是直线上的、且对于该网格单元是递增的 2 个网格单元,如果该直线是对角线,则称为对角线反方向,如果非对角线,则称为非对角线反方向^[15]。用 partition and judge the grid C_{ij} 算法计算一个单元网格是否为簇与簇的边界网格,并统计趋向的反方向簇数记为 $count$ 。

定义 6 在图 1 中,网格单元 ε 为稠密网格, a, b, e 均为稀疏网格,网格单元 ε 中的点的坐标分别为:

$$(o_{11}, o_{12}, \dots, o_{1d}), (o_{21}, o_{32}, \dots, o_{ad}), \dots, (o_{n1}, o_{n2}, \dots, o_{nd})$$

其中, d 为特征空间的维数,则网格单元 ε 中数据对象的质心坐标为:

$$o_1 = \frac{\sum_{i=1}^n o_{i1}}{n}, o_2 = \frac{\sum_{i=1}^n o_{i2}}{n}, \dots, o_d = \frac{\sum_{i=1}^n o_{id}}{n}$$

质心到网格单元 ε 中点的最远距离为:

$$r = \max(\sqrt{[(o_1 - o_{11})^2 + (o_2 - o_{12})^2 + \dots + (o_d - o_{1d})^2]}) \quad (5)$$

以质心为圆心、 r 为半径的圆中的数据对象属于网格单元 ε 中的簇号。

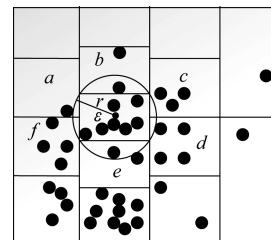


图 1 小波聚类算法聚类网格单元的质心

如图 2 所示,图 2(a) 中 A2 即坐标系中第 2 行第 A 列的网格和 C4 即坐标系中第 4 行第 C 列的网格有可能被当成噪声来处理,图 2(b) 中有 B2 即坐标系中第 2 行第 B 列的网格的出现,使原本的 2 个

簇变成了 1 个簇,并且其中明显存在不属于任何簇的点。网格中数据对象分布不一,稠密网格有可能是几个簇的边界,导致不同的簇最后被标记为同一个簇等情况,引入定义 5 中的 partition and judge the grid C_{ij} 算法,找出满足上述条件的边缘网格(如果存在这种边界网格),稀疏网格中的数据对象有可能属于与其连通的簇。

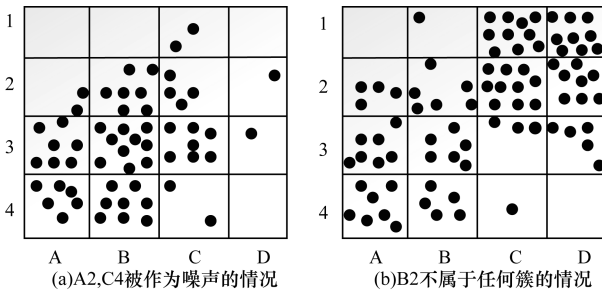


图 2 小波聚类算法聚类结果分析

1) 改进小波聚类算法

算法步骤如下:

步骤 1 计算每维相对网格数,划分网格。对于 d 维特征空间的 A_i 维, $A_i = \{a_{i1}, a_{i2}, \dots, a_{in}\}$, 按照式(2)求得每维属性相对网格数为:

$$m_i = (\text{int}) \frac{(a_{imax} - a_{imin})}{dec} \times m \quad (6)$$

并划分每维网格数为 m_i 。

步骤 2 量化特征空间。输入数据集为 $E = \{e_1, e_2, \dots, e_d\}$, 将数据集 E 映射到划分好的相对网格中去,即对于网格单元 $c_i, c_i = \{c_{i1}, c_{i2}, c_{id}\}$, 如果 $l_{ij} \leq e_{ki} \leq h_{ij}$, 则网格单元 c_i 包含数据对象 $e_k = \{e_{k1}, e_{k2}, \dots, e_{kd}\}$ 。

步骤 3 在特征空间运用小波变换。通过对几种小波函数的比较后,选择 Daubechies 小波 (db. 2 小波), db. 2 小波具有精度高、速度快等优点。然后在特征空间上运用 db. 2 小波变换。

步骤 4 查找稠密网格和稀疏网格。密度阈值 q , 由式(2)与式(3)求出网格相对密度阈值 q_i , 按照定义 4 划分稠密网格和稀疏网格。

步骤 5 寻找稠密连通单元。在稠密网格中寻找连在一起的部分形成稠密连通单元并分配相应的簇标识号。

步骤 6 查找边界网格,分配相应的簇标识号。按照定义 5 中的 partition and judge the grid C_{ij} 算法求出稠密网格中的簇与簇的边界网格,并求得数据对象分别趋向的簇,即趋向的簇数为 $count$, 运用 K-Means 算法聚类 ($k = count + 1$) 并

分配趋向簇的簇号,统计剩余的数据对象数,如果大于等于相对密度阈值 q_i , 则分配新的簇号,否则标记为异常点。

步骤 7 给稀疏单元分配簇标识号。查找与稠密连通单元连通的稀疏网格,如果存在则按照定义 6 中的方法求出数据对象的质心及圆的半径 r , 对圆中的数据对象分配网格单元的簇号,标记剩下的稀疏网格中的数据对象为异常点。

步骤 8 构造查找表并输出。构造一个查找表 T 来给出簇 w 中量化空间中的数据和原始数据集的对应关系并输出。

2) 复杂度分析

算法初始是读取数据时间复杂度为 $O(n)$, 其中 n 是数据的个数,然后量化特征空间并进行小波变换最坏情况下需要对每个网格单元变换一次,则时间复杂度为 $O(w \times d \times m_1 \times m_2 \times \dots \times m_n)$, 其中, w 为小波卷积系数; d 为数据维度; m_i 表示每维划分网格数均为常数,寻找连通单元的时间复杂度为 $O(m_1 \times m_2 \times \dots \times m_n)$, 在网格中运用 partition and judge the grid C_{ij} 算法的时间复杂度为 $O(n_1)$, K-Means 算法的时间复杂度为 $O(n_2 \times k \times t)$, $n_2 \times k \times t \ll n$, 即由网格类聚类算法待定网格代替数据点的特性得到时间复杂度为 $O(n)$ 。改进算法更改了划分网格的方式,空间复杂度发生了变换,由划分方法容易得出空间复杂度为 $O(m_1 \times m_2 \times \dots \times m_n)$ 。如表 1 为小波聚类算法 WaveCluster 与改进的小波聚类算法 CWaveCluster 的时间复杂度与空间复杂度对比。

表 1 小波聚类算法与改进后的算法时间复杂度与参数比较

算法	时间复杂度	空间复杂度
WaveCluster 算法	$O(n)$	$O(m^d)$
CWaveCluster 算法	$O(n)$	$O(m_1 \times m_2 \times \dots \times m_n)$

2 实验结果与分析

将本文提出的改进算法与小波聚类算法分别应用到 QAR 数据异常点检测中,比较两者之间的效率。实验采用波音 737-800 的 QAR 飞行数据,选取 2 种不同的故障数据对 CWaveCluster 算法进行验证。算法采用 Matlab R2014a 实现,实验电脑配置为 Intel Core i7, 2.25 GHz, 8 GB 内存实现,操作系统为 Windows 7, 实验流程和实验数据分别见图 3 和表 2。

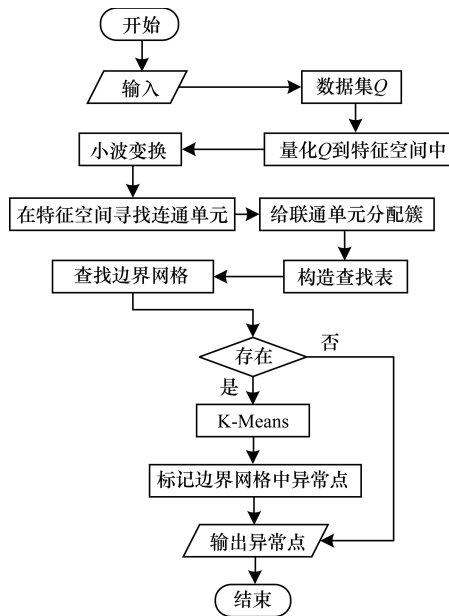


图3 改进小波聚类算法 CWaveCluster 实验流程

表2 实验数据信息

故障类型	数据对象数	选取属性
空中停车	13 386	Air, N2(高压转子转速)
空中颠簸	32 712	Air/Ground, AccelVert

用启发式方法选取不同的量化值 m (从小到大) 得到相对量化值 m_i 来找到合适的簇数 k , 使用小波聚类算法 WaveCluster 和改进后的小波聚类算法 CWaveCluster 对选取的 QAR 数据集聚类分析。

实验1 选取空中停车数据, 根据专家经验获知飞机在空中是否有空中停车的故障发生主要与 N2 (高压转子转速) 的转速有关, 当 N2 转速小于 25 时发生空中停车故障, 由于空中停车发生在飞机飞行起飞之后, 因此截取 QAR 数据的 Air/Ground (空中或陆地) 和 N2 为所用属性, 并且对 Air/Ground 预处理, 这里只选择 Air/Ground 属性列中为 Air 的 N2 数据, 该数据为波音 737-800 空中巡航阶段的 13 386 条数据, 其中含空中停车数据 13 条, 取密度阈值为 20, 实验结果见表 3、表 4。

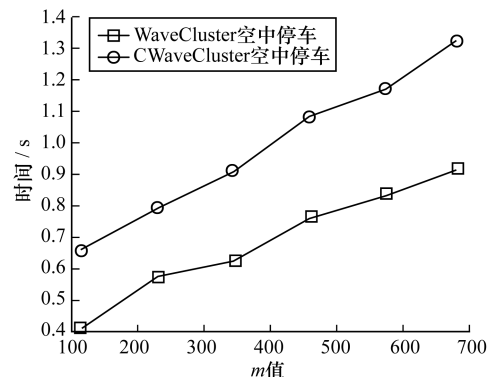
表3 小波聚类算法在空中停车故障中应用效果

m	k	所得异常点数	实际异常点数	时间 /s	准确率 /%
115	3	33	13	0.41	39.39
230	4	21	13	0.57	61.90
345	5	16	13	0.62	81.25
460	6	17	13	0.76	76.47
575	7	19	13	0.83	68.42
680	8	35	13	0.91	37.14

表4 改进小波聚类算法在空中停车故障中应用效果

m	k	所得异常点数	实际异常点数	时间 /s	准确率 /%
115	3	31	13	0.66	41.90
230	4	17	13	0.79	76.47
345	6	13	13	0.91	100.00
460	6	13	13	1.08	100.00
575	7	16	13	1.17	81.25
680	8	28	13	1.32	46.42

从表 3、表 4 中可以看出, 在均为均匀划分的前提下, 当 m 值较小时, 2 种算法均将正常数据划分为异常数据, 但是改进后的算法将正常数据看做异常数据的个数明显较少; 随着 m 值的增加, 2 种算法的聚类精度都有不同程度的提高, 聚类效果增强, 异常数据中出现正常数据的个数越来越少, 对比表 3 和表 4 容易看出, 当量化值 m 为 345 时 WaveCluster 算法将边界网格两端的簇当成了 1 个簇, 而 CWaveCluster 算法正确区分 2 个簇; 当 m 值增加到一定程度后, 空中停车数据的聚类效果又开始下降, 但是 CWaveCluster 算法仍然比 WaveCluster 算法效率要高。综上所述, 改进后的算法效率明显要比改进之前更高。由表 3、表 4 的实验结果得到 2 个算法在不同 m 值时所用的时间对比, 如图 4 所示。

图4 空中停车故障不同 m 值下的运行时间

由图 4 可以看出, 改进的小波聚类算法 CWaveCluster 运行时间大于小波聚类算法 WaveCluster 的运行时间, 但是处于同一数量级, 结合表 1 改进小波聚类算法 CWaveCluster 与小波聚类算法 WaveCluster 的时间复杂度均为 $O(n)$, 说明改进算法可行。

实验2 选取空中颠簸数据, 根据专家经验波音 737-800 飞行过程中发生空中颠簸故障, 主要和飞机飞行过程中的垂直加速度有关, 当飞机飞行过程中的垂直加速度位于 $[0.5, 1.5]$ 之间时飞机无明显颠簸感觉, 当飞机飞行过程中的垂直加速度小于 0.5 或者大于 1.5 时颠簸感觉较强烈, 由此选取 QAR 数

据中飞机的 Air/Ground(预处理)和 AccelVert(垂直加速度)属性,截取波音 737-800 空中巡航阶段的 32 712 条数据,其中含空中颠簸数据 21 条,取密度阈值为 30,实验结果见表 5、表 6。

表 5 小波聚类算法实验结果

m	k	所得异常 点数	实际异常 点数	时间 /s	准确率 /%
181	3	45	21	0.87	46.67
362	4	33	21	0.99	63.64
543	5	25	21	1.09	84.00
724	6	23	21	1.22	91.30
905	7	24	21	1.34	87.50
1 086	8	31	21	1.47	67.74
1 276	9	47	21	1.54	44.68

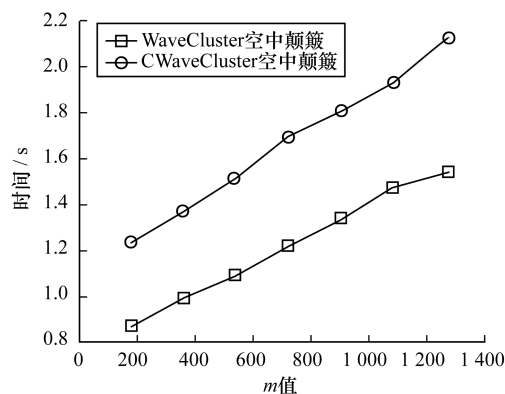
表 6 改进小波聚类算法实验结果

m	k	所得异常 点数	实际异常 点数	时间 /s	准确率 /%
181	3	40	21	1.23	52.50
362	4	28	21	1.37	75.00
543	5	23	21	1.52	91.30
724	6	21	21	1.69	100.00
905	7	22	21	1.80	95.45
1 086	8	27	21	1.93	77.78
1 276	9	38	21	2.12	55.26

从表 5、表 6 中可以看出,在非均为均匀划分的前提下,当 m 值较小时,2 种算法均将正常数据划分为异常数据;随着 m 值的增加,WaveCluster 算法和 CWaveCluster 算法的聚类精度均不同程度提高,异常数据中出现正常数据的个数越来越小并且改进后的算法效率更高;当 m 值增加到一定程度后,WaveCluster 算法和 CWaveCluster 算法的聚类精度均开始下降,但是 CWaveCluster 算法仍然比 WaveCluster 算法效率要高。由表 5、表 6 的实验结果得到 2 个算法在不同 m 值时所用的时间对比,如图 5 所示。

由图 5 可以看出,改进的小波聚类算法 CWaveCluster 运行时间要大于小波聚类算法 WaveCluster,但是处于同一数量级,结合表 1 改进小波聚类算法 CWaveCluster 与小波聚类算法 WaveCluster 的时间复杂度均为 $O(n)$,说明改进算法可行。

对比实验 1 与实验 2 的实验结果可以得出,CWaveCluster 在不增加原有算法时间复杂度数量级的情况下增加了聚类精度,提高了聚类效率。

图 5 空中颠簸故障不同 m 值下的运行时间

3 结束语

本文分析了小波聚类算法的特点,在此基础上对小波聚类算法进行改进,并将其应用到 QAR 数据集异常检测中。该算法根据 QAR 数据特点对网格进行非均匀划分,将其量化到网格中并运用小波变换寻找网格中的稠密连通单元,运用 partition and judge the grid C_{ij} 算法寻找簇与簇的边界,利用 K-Means 算法划分边界并标记为相应的簇编号,最后将属于某个簇的稀疏网格分配给相应的簇并标记出异常点。通过在 QAR 数据集中的实验结果验证了该算法的有效性和可行性。CWaveCluster 算法在应用中还需要根据实际情况进行改进,如每维数据换算后差距太大可能导致划分的网格每维数量差距过大,从而影响聚类精度;算法应用到高维空间中也存在一些问题,如 K-Means 算法是基于距离判断边缘网格中的数据对象属于某个簇,但是在高维数据空间中距离相等的点较多,影响算法的聚类精度,这将是下一步需要研究的工作。

参考文献

- [1] 曹琳. QAR 原理分析及在飞机维护中的应用[J]. 江苏航空, 2011(1): 36-37.
- [2] SHEIKHOLESLAMI G, CHATTJRC S, ZHANG Aidong. Wave Cluster: A Multi-resolution Clustering Approach for Very Large Spatial Database [C]//Proceedings of International Conference on Very Large Data Base. New York, USA: ACM Press, 1998: 428-439.
- [3] 赵光嗣, 李伟生. 基于 MPI 的并行小波聚类算法在曙光 TC1700 上的实现[J]. 计算机应用, 2006, 26(3): 645-646.
- [4] 方利民, 林敏. 小波聚类方法和近红外光谱技术用于药片种类判别[J]. 光谱学与光谱分析, 2010, 30(11): 2958-2961.

(下转第 38 页)

初阶段就被划分到同一个处理节点上,结合本文的动态调整算法,使分布式信息网数据库管理系统达到更好的处理性能。

参考文献

- [1] 杨小虎,王新宇,毛明.基于数据划分的分布式模型及其负载均衡算法[J].浙江大学学报(工学版),2008,42(4):602-607.
- [2] HENDRICKSON B, LELAND R. A Multi-level Algorithm for Partitioning Graphs [C]//Proceedings of 1995 ACM/IEEE Conference on Supercomputing. Washington D. C., USA; IEEE Press, 1995:28.
- [3] KARYPIS G, KUMAR V. Parallel Multilevel Graph Partitioning Schemes [D]. Minneapolis, USA: University of Minneapolis, 1995.
- [4] ARORA A, KAUR K. Enhanced Multilevel Hybrid Algorithm for Graph Partitioning [J]. International Journal of Computer Applications, 2015, 120(6):16-19.
- [5] LIU Mengchi, HU Jie. Information Networking Model [M]// LAENDER A H F, CASTANO S, DAYAL U, et al. Conceptual Modeling-ER 2009. Berlin, Germany: Springer, 2009:131-144.
- [6] 胡捷,刘梦赤.信息网模型 INM 研究 [M]. 武汉:科学出版社,2011.
- [7] 徐倩,胡婕,刘梦赤.复杂语义关系的描述与操作 [J]. 计算机科学与探索, 2014, 8(12):1432-1441.
- [8] GAREY M R, JOHNSON D S, STOCKMEYER L. Some Simplified NP-Complete Problems [C]//Proceedings of ACM Symposium on Theory of Computing. New York, USA; ACM Press, 1974:47-63.
- [9] ANDREEV K, RACKE H. Balanced Graph Partitioning [J]. Theory of Computing Systems, 2006, 39(6):929-939.
- [10] GUY E, JOSEPH N, SATISH R, et al. Fast Approximate Graph Partitioning Algorithms [C]//Proceedings of the 8th ACM-SIAM Symposium on Discrete Algorithms. [S. l.]: Society for Industrial and Applied Mathematics, 1997:2187-2214.
- [11] TATAROWICZ A L, CURINO C, JONES E P C, et al. Lookup Tables; Fine-grained Partitioning for Distributed Data-bases [C]// Proceedings of the 28th IEEE International Conference on Data Engineering. Washington D. C., USA; IEEE Press, 2012:102-113.
- [12] GUTTMANN-BECK N, HASSIN R. Approximation Algorithms for Minimum K-cut [J]. Algorithmica, 2000, 27(2):198-207.
- [13] GHAFARI M, KUHN F. Distributed Minimum Cut Approximation [M]// Afek Y. Distributed Computing. Berlin, Germany: Springer, 2013:1-15.
- [14] NAGAMUCHI H, IBARAKI T. A Fast Algorithm for Computing Minimum 3-way and 4-way Cuts [J]. Mathematical Programming, 2000, 88(3):507-520.
- [15] DAN G, ÉVA T. A Faster Parametric Minimum-cut Algorithm [J]. Algorithmica, 1994, 11(3):278-290.
- [16] XU Ning, CHEN Lei, CUI Bin. LogGP: A Log-based Dynamic Graph Partitioning Method [J]. Proceedings of the VLDB Endowment, 2014, 7(14):1917-1928.
- [5] 陈佐,谢赤,陈晖.基于小波聚类方法的股票收益序列时间模式与挖掘 [J]. 系统工程, 2005, 23(11):102-107.
- [6] MAHESHWARY P, SRIVASTAVA N. Wave Cluster for Remote Sensing Image Retrieval [J]. International Journal on Computer Science and Engineering, 2011, 3(2):976-979.
- [7] MICHAEL B. Mammographic Segmentation Using Wave Cluster [J]. Algorithms, 2012, 5(3):318-329.
- [8] LI Xingyi, LU Junyun, SHI Huaji, et al. An Approach for Treatment of the Incomplete Data Based on Wave Cluter and Weighted 1-nearest Neighbor [C]// Proceedings of 2009 International Association of Computer Science and Information Technology Conference. Berlin, Germany: Springer, 2009:3-8.
- [9] ANGGRAINI E L, SUCIATI N, SUADI W. Parallel Computing of Wave Cluster Algorithm for Face Recognition Application [C]//Proceedings of International Conference on QIR. Washington D. C., IEEE Press, 2013:1195-1203.
- [10] YILDIRIM A A, ÖZDOĞAN C. Parallel Wavelet-based Clustering Algorithm on GPUs Using CUDA [J]. Procedia Computer Science, 2011, 3(1):396-400.
- [11] 王治和,杨晏.基于双层网格和密度的数据流聚类算法 [J]. 计算机工程, 2014, 40(4):146-150.
- [12] 林秀丹,毛国君.基于密度网格的分布式数据流聚类算法 [J]. 计算机工程, 2012, 38(16):70-73.
- [13] 张书春,孙秀英.基于网格结构的 CLARANS 改进算法 [J]. 计算机工程, 2012, 38(6):56-59.
- [14] 单世民,张宁,江贺,等.基于网格和密度的簇边缘精度增强聚类算法 [J]. 计算机工程与应用, 2008, 44(8):143-146.
- [15] 余灿玲,王丽珍,张元武.基于网格密度方向的聚类簇边缘精度加强算法 [J]. 计算机研究与发展, 2010, 15(5):816-822.

编辑 金胡考

编辑 索书志

(上接第33页)