

论文

一种基于自我聚类的异常检测学习方法

李娜娜 赵政 刘伯颖 顾军华

天津大学 计算机科学与技术学院; 河北工业大学 计算机科学与软件学院 天津大学 河北工业大学 教务处 河北工业大学 计算机科学与软件学院

摘要:

提出一种新的基于正选择的异常检测方法, 该方法通过聚类学习正常空间特征, 在每个类中选择有代表性的自我样本构造检测器集, 之后利用正选择算法进行异常检测。这种方法既能适用于自我样本集较多的情形, 克服了T. Stibor提出的正选择的局限, 又具备了一定的学习能力。同时, 该方法还避免了负选择中随机选择样本带来的弊端。通过实验分析, 该方法比VDetector具备更好的检测性能, 是一种有效的异常检测方法。

关键词: 聚类 异常检测 负选择 正选择

Anomaly detection method by clustering normal data

Abstract:

A new anomaly detection method was proposed based on positive selection. The method learned the characteristic of "self" space by clustering, and then selected typical samples from every cluster to construct detectors. And positive selection was used to detect anomalies. The new algorithm is not only effective in certain application with large number of "self" samples, but also avoids the shortcoming by randomly selecting sample in VDetector. Experimental results on Ring data and biomedical data show that the new method is more effective in anomaly detection.

Keywords: cluster anomaly detection negative selection positive selection

收稿日期 2007-12-10 修回日期 2008-02-25 网络版发布日期

DOI:

基金项目:

通讯作者: 李娜娜

作者简介:

参考文献:

本刊中的类似文章

1. 刘晓锋 吴亚娟 李明东. 一种基于模糊聚类的资源发现策略[J]. 计算机应用, 2007,27(9): 2227-2229
2. 许宁 张毅坤. 基于正交分层聚类算法软件可靠性模型的预测分析[J]. 计算机应用, 2007,27(3): 635-637
3. 张鲲 王士同. 分水岭算法和基于MRF的层次聚类相结合的混合无监督图像分割算法[J]. 计算机应用, 2007,27(3): 673-676
4. 潘登 郑应平 徐立鸿 陈俊. 基于RBF神经网络的网格数据聚类方法[J]. 计算机应用, 2007,27(2): 333-336
5. 覃俊华 张洪伟 赵世政. 基于遗传算法的模糊聚类研究及其应用[J]. 计算机应用, 2007,27(1): 52-55

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(634KB)
- ▶ [HTML全文]
- ▶ 参考文献

服务与反馈

- ▶ 把本文推荐给朋友
- ▶ 加入我的书架
- ▶ 加入引用管理器
- ▶ 引用本文
- ▶ Email Alert
- ▶ 文章反馈
- ▶ 浏览反馈信息

本文关键词相关文章

- ▶ 聚类
- ▶ 异常检测
- ▶ 负选择
- ▶ 正选择

本文作者相关文章

- ▶ 李娜娜
- ▶ 赵政
- ▶ 刘伯颖
- ▶ 顾军华

PubMed

- ▶ Article by
- ▶ Article by
- ▶ Article by
- ▶ Article by

6. 陈静 张晓敏 .无线传感器网络簇头优化分簇算法及其性能仿真[J]. 计算机应用, 2006,26(12): 2787-2788
7. 张文良 黄亚楼 倪维健 .一种基于聚类的文本特征选择方法[J]. 计算机应用, 2007,27(1): 205-206
8. 唐玉华 杨晓元 张敏情 韩鹏 .多超球面OC-SVM算法在隐秘图像检测中的应用[J]. 计算机应用, 2006,26(12): 2887-2889
9. 郭晨娟 李战怀 .基于概念的网页相似度处理算法研究[J]. 计算机应用, 2006,26(12): 3030-3032
10. 祁云平 张其善 佟雨兵.基于PSNR与SSIM联合的图像质量评价模型[J]. 计算机应用, 2007,27(2): 503-506
11. 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
12. 孙志伟 赵政 .SOFM神经网络在处理非空间属性中的应用[J]. 计算机应用, 2006,26(11): 2667-2669
13. 彭敦陆 周傲英 .基于方法聚类的Web服务检索技术[J]. 计算机应用, 2007,27(10): 2365-2368
14. 曹洪其 孙志挥 .基于网格技术的高维大数据集离群点挖掘算法[J]. 计算机应用, 2007,27(10): 2369-2371
15. 魏文红 高大利 孙镇涛 .一种基于Cayley图的确定性小世界网络的研究[J]. 计算机应用, 2007,27(10): 2406-2408
16. 文娟娟 柴玉梅 .基于免参数数据挖掘的相异度度量研究[J]. 计算机应用, 2006,26(12): 2982-2984
17. 艾轶博 穆志纯 陈静 .利用双向SOFM网络模拟汉字认知过程的研究[J]. 计算机应用, 2006,26(12): 2971-2973
18. 陈增照 杨扬 何秀玲 喻莹 董才林 .基于核聚类的SVM多类分类方法[J]. 计算机应用, 2007,27(1): 47-49
19. 蒋先刚 .基于各向异性扩散的图像平滑及在三维重构预处理中的应用[J]. 计算机应用, 2007,27(1): 249-251
20. 张喆 白琳 .一种基于克隆网络聚类的入侵检测方法[J]. 计算机应用, 2007,27(1): 128-131
21. 高原; 耿国华; 王怡.基于动态矩形的聚类方法的设计与实现[J]. 计算机应用, 2006,26(4): 870-871
22. 郑小慎.基于频繁特征项集的文档聚类研究[J]. 计算机应用, 2006,26(4): 875-877
23. 王威 卢虎生 刘存福 .求解非规则非致密分布数据的一种算法[J]. 计算机应用, 2006,26(7): 1669-1672
24. 匡泰; 朱清新; 孙跃.FCM算法用于灰度图像分割的初始化方法的研究[J]. 计算机应用, 2006,26(4): 784-786
25. 蔺旭东 曾晓宁 薄静仪.一种基于支持向量的镜头聚类算法[J]. 计算机应用, 2007,27(9): 2143-2146
26. 咎鑫 郑庆华 范宇倩 韩九强.攻击案例综合学习系统研究[J]. 计算机应用, 2007,27(9): 2177-2179
27. 曹晓莉 江朝元 甘思源.基于聚类支持向量机的船用污水处理装置故障诊断[J]. 计算机应用, 2008,28(10): 2648-2651
28. 陆虎 李永忠.不确定聚类算法及其在入侵检测系统中应用[J]. 计算机应用, 2008,28(10): 2715-2717
29. 刘纯青 杨莘元 张颖 .基于文化算法的聚类分析[J]. 计算机应用, 2006,26(12): 2953-2955
30. 徐雪松 .一种基于距离的再聚类的离群数据发现算法[J]. 计算机应用, 2006,26(10): 2398-2400
31. 陈治平 王雷 .基于密度梯度的聚类算法研究[J]. 计算机应用, 2006,26(10): 2389-2392
32. 张军 苏璞睿 冯登国 .基于系统调用的入侵检测系统设计与实现[J]. 计算机应用, 2006,26(9): 2137-2139
33. 柯永振 张家万 孙济洲 张怡 周小舟 .结合支持向量机与C均值聚类的图像分割[J]. 计算机应用, 2006,26(9): 2081-2083
34. 谷瑞军 须文波 .基于核方法的彩色图像量化研究[J]. 计算机应用, 2006,26(9): 2063-2064
35. 郭炜 何丕廉 王中 .数据挖掘在客户关系管理中的应用[J]. 计算机应用, 2006,26(8): 1996-1997
36. 康恺 林坤辉 周昌乐 .基于主题词频数特征的文本主题划分[J]. 计算机应用, 2006,26(8): 1993-1995
37. 郭依林 李中华 毛宗源 .自适应人工免疫算法在数据挖掘中的应用[J]. 计算机应用, 2006,26(8): 1943-1946
38. 蒋志为 陶宏才 白学祥 .基于模糊集的蚁群聚类算法的改进[J]. 计算机应用, 2006,26(8): 1950-1952
39. 范新南 沈红斌 陈学忠 .特征空间属性加权模糊核聚类算法[J]. 计算机应用, 2006,26(8): 1888-1889
40. 王燕 .基于信息熵的标称变量聚类算法研究[J]. 计算机应用, 2006,26(8): 1904-1905
41. 彭敦陆 周傲英 .基于向量空间的Web服务发现模糊方法[J]. 计算机应用, 2006,26(9): 2009-2012
42. 刘纯平 .基于Kohonen神经网络聚类方法在遥感分类中的比较[J]. 计算机应用, 2006,26(7): 1744-1746
43. 周涓; 熊忠阳; 张玉芳; 任芳.基于最大最小距离法的多中心聚类算法[J]. 计算机应用, 2006,26(6): 1425-1427
44. 苏晶; 潘蕾; 徐汀荣.结合内容挖掘的事务识别改进[J]. 计算机应用, 2006,26(6): 1411-1414
45. 俞研 黄皓 .一种半聚类的异常入侵检测算法[J]. 计算机应用, 2006,26(7): 1640-1642
46. 戴维迪; 何丕廉; 刘振华; 张倩.一种基于窗口的快速聚类算法——F-CABDET[J]. 计算机应用, 2006,26(5): 1127-1129

47. 邓健爽; 郑启伦; 彭宏. 基于提取网站层次结构的网页分类方法[J]. 计算机应用, 2006,26(5): 1134-1136
48. 李章华 刘高嵩 刘柯萍. 一种P2P系统节点聚类及信息检索算法[J]. 计算机应用, 2007,27(4): 821-824
49. 陈春燕 张伟. 一种基于概率密度的数据流聚类算法[J]. 计算机应用, 2007,27(4): 881-883
50. 王辉 高利军 王听忠. 个性化服务中基于用户聚类的协同过滤推荐[J]. 计算机应用, 2007,27(5): 1225-1227
51. 张钊 王锁柱. 一种基于SOM和PAM的聚类算法[J]. 计算机应用, 2007,27(6): 1400-1402
52. 沈鸿 颜晓龙. 一种适用于高维数据流的子空间聚类方法[J]. 计算机应用, 2007,27(7): 1680-1684
53. 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
54. 曾志雄. 一种有效的基于划分和层次的混合聚类算法[J]. 计算机应用, 2007,27(7): 1692-1694
55. 张建锦 吴渝 刘小霞. 一种改进的密度偏差抽样算法[J]. 计算机应用, 2007,27(7): 1695-1698
56. 周瑞英 顾军华 李娜娜 谭庆. 一种基于免疫和Hopfield神经网络的多峰值优化算法[J]. 计算机应用, 2007,27(7): 1751-1753
57. 楼若岩 许晓东 朱士瑞. 面向校园网的IP地址逐步优化层次聚类算法[J]. 计算机应用, 2007,27(8): 1862-1864
58. 王国良 梁德群 王演 王彦春. 基于区域与光照不变性的运动阴影检测算法[J]. 计算机应用, 2007,27(9): 2152-2153
59. 陈峻 邹凌君 屠莉. 多数据流的实时聚类算法[J]. 计算机应用, 2007,27(8): 1976-1979
60. 丁卫平 邓 伟. 一种基于约束关系的电子病历图像分割核聚类算法[J]. 计算机应用, 2007,27(8): 2066-2068
61. 倪霖 郑洪英. 基于聚类和支撑向量机的入侵检测研究[J]. 计算机应用, 2007,27(10): 2440-2442
62. 龚雪晶 慈林林 姚康泽. 基于邻域信息的遥感图像模糊聚类及并行算法设计[J]. 计算机应用, 2007,27(10): 2512-2514
63. 陈治平 胡宇舟 顾学道. 聚类算法在电信客户细分中的应用研究[J]. 计算机应用, 2007,27(10): 2566-2569
64. 林甲祥 樊明辉 陈崇成 江先伟. 二阶段近似KNN离群挖掘算法与应用[J]. 计算机应用, 2007,27(10): 2598-2601
65. 康大伟 陈天滋. 一种基于密度的面向线段的聚类算法[J]. 计算机应用, 2007,27(11): 2760-2762
66. 王飒 郑链. 基于Fisher准则和特征聚类的特征选择[J]. 计算机应用, 2007,27(11): 2812-2813
67. 衣治安 刘杨. 基于二叉树的多分类SVM算法在电子邮件过滤中的应用[J]. 计算机应用, 2007,27(11): 2860-2862
68. 冷明伟 陈晓云 颜清. 一种基于影响因子的快速K-均值算法[J]. 计算机应用, 2007,(12): 3042-3044
69. 唐勇智. 基于聚类的RBF-LBF串联神经网络学习算法[J]. 计算机应用, 2007,(12): 2916-2918
70. 高俊波 安博文 王晓峰. 在线论坛中潜在影响力主题的发现研究[J]. 计算机应用, 2008,28(1): 140-142
71. 於时才 陈涓. 基于小生境技术的神经网络进化集成[J]. 计算机应用, 2008,28(12): 3052-3054
72. 白如珍 田青 徐海江. 一种新的基于分水岭变换的聚类分析算法[J]. 计算机应用, 2008,28(12): 3240-3243
73. 孙秀娟 刘希玉. 基于新聚类有效性函数的改进K-means算法[J]. 计算机应用, 2008,28(12): 3244-3247
74. 李恒杰. 多子系统似然度评分融合说话人识别[J]. 计算机应用, 2008,28(1): 116-119,119
75. 刘雪燕 李明 张亚芬. 基于PCA和多约简SVM的多级说话人辨识[J]. 计算机应用, 2008,28(1): 127-130
76. 刘天鹏 周娅. P2P网络中的数据挖掘[J]. 计算机应用, 2008,28(1): 162-164,
77. 郭森 柳伟 王建华. 基于Mean-shift的粘连人体目标分割算法[J]. 计算机应用, 2009,29(1): 51-53
78. 戴维 张申生. 基于二值化聚类的图像文字提取算法[J]. 计算机应用, 2009,29(1): 57-59,7
79. 蔡茹 雷英杰 申晓勇 雷阳. 基于直觉模糊等价相异矩阵的聚类方法[J]. 计算机应用, 2009,29(1): 123-126
80. 蒋盛益 李霞. 一种改进的BIRCH聚类算法[J]. 计算机应用, 2009,29(1): 293-296
81. 李海芳 温侠 门立环. 基于蚁群的图像情感特征聚类算法的研究[J]. 计算机应用, 2009,29(2): 360-363
82. 李园敏 江桦 李霞. 数字信号调制识别的RBFNN分类器算法研究与实现[J]. 计算机应用, 2009,29(3): 798-800
83. 李刚成 刘赞波 曾庆光. 一种基于模糊聚类的构造进化树方法[J]. 计算机应用, 2009,29(3): 836-838
84. 汤周文 叶东毅. 基于层次聚类的差异化属性约简算法[J]. 计算机应用, 2009,29(2): 419-420
85. 况夯 罗军. 基于遗传FCM算法的文本聚类[J]. 计算机应用, 2009,29(2): 558-560
86. 唐少先 蔡文君. 基于无监督聚类混合遗传算法的入侵检测方法[J]. 计算机应用, 2008,28(2): 409-411
87. 黄崇本 陶剑文 程光华. 一种面向图包容搜索的图索引模型[J]. 计算机应用, 2008,28(2): 479-483
88. 孙越恒 侯越先 何丕廉. 非线性维数约减算法在文档聚类中的应用[J]. 计算机应用, 2008,28(2): 488-490

89. 杨振亚 王勇 杨振东 王成道.一种全新的RGB色差度量公式[J]. 计算机应用, 2009,29(2): 465-467
90. 王小乐 刘青宝 陆昌辉 陈文凯.一种处理障碍约束的聚类算法[J]. 计算机应用, 2009,29(2): 406-408
91. 梁敏君 倪志伟 倪丽萍 杨葛钟啸.基于网格与分形维数的聚类算法[J]. 计算机应用, 2009,29(3): 830-832
92. 邱保志 刘洋 陈本华.基于网格熵的边界点检测算法[J]. 计算机应用, 2008,28(3): 732-734
93. 张磊 陈曙.一个新的基于能量和距离的传感器网络协议[J]. 计算机应用, 2008,28(5): 1117-1119
94. 孙亚 钱洪波 叶亮.数据挖掘算法在交通状态量化及识别的应用[J]. 计算机应用, 2008,28(3): 738-741
95. 谭颖 胡瑞飞 殷国富.多密度阈值的DBSCAN改进算法[J]. 计算机应用, 2008,28(3): 745-748
96. 覃晓.基于遗传算法和自组织特征映射网络的文本聚类方法[J]. 计算机应用, 2008,28(3): 757-760
97. 蒋鸿 胡永祥 朱文球.应用快速高斯变换估计图像互信息[J]. 计算机应用, 2008,28(3): 680-682
98. 陆宏菊 刘培玉 崔嘉.结合模糊聚类的遗传算法在网络信息过滤中的应用[J]. 计算机应用, 2008,28(3): 703-705
99. 张松顺 李朝锋 吴小俊 高翠芳.改进微分进化算法的半监督模糊聚类[J]. 计算机应用, 2009,29(4): 1046-1047
100. 王明春 唐万生 江琪 刘鑫.基于相对距离的改进粗K-means方法[J]. 计算机应用, 2009,29(4): 1102-1105
101. 熊丽琼 郭帆 余敏.一种基于遗传聚类的报警聚合方法[J]. 计算机应用, 2008,28(4): 896-898
102. 郎焰 郭秀清.基于改进最短邻聚类的最优模糊在线辨识[J]. 计算机应用, 2008,28(7): 1659-1661
103. 谢信喜 王士同.适用于区间数据的基于相互距离的相似性传播聚类[J]. 计算机应用, 2008,28(6): 1441-1443
104. 陈秀敏 邹开其 闫忠文 祝美宁 闫娟娟 杨艳萍 阎丹丹.一种基于改进神经网络的高效模糊聚类算法[J]. 计算机应用, 2008,28(5): 1190-1193
105. 宋娇 葛临东.一种遗传模糊聚类算法及其应用[J]. 计算机应用, 2008,28(5): 1197-1199
106. 范莹 计华 张化祥.一种新的基于模糊聚类的组合分类器算法[J]. 计算机应用, 2008,28(5): 1204-1207
107. 金慧珍 赵辽英.多层自动确定类别的谱聚类算法[J]. 计算机应用, 2008,28(5): 1229-1231
108. 方新 赵卫东 杨晓春.基于Ant-Tree聚类算法的图像分割[J]. 计算机应用, 2008,28(5): 1240-1243
109. 胡彧 闫巧梅.滑动窗口模型下的优化数据流聚类算法[J]. 计算机应用, 2008,28(6): 1414-1416
110. 朱嘉贤 白伟华.基于服务虚拟化的业务编程模型[J]. 计算机应用, 2008,28(6): 1635-1638
111. 陈治平.基于多模式分类算法及其在客户保持中的应用[J]. 计算机应用, 2008,28(6): 1417-1419
112. 唐敏 阳爱民.一种高效的图像数据库检索方法[J]. 计算机应用, 2008,28(6): 1454-1456
113. 王顺风 张建伟.基于Gibbs场与模糊C均值聚类的脑MR图像分割[J]. 计算机应用, 2008,28(7): 1750-1752
114. 宋凌 李枚毅 李孝源.一种新的半监督入侵检测算法[J]. 计算机应用, 2008,28(7): 1781-1783
115. 陈应良 王士同.权重马氏距离高斯核在谱分割中的应用[J]. 计算机应用, 2008,28(7): 1738-1741
116. 罗 冯国灿 成秋生.结合聚类和改进的C-V演化方程在医学图像分割中的应用[J]. 计算机应用, 2008,28(9): 2288-2291
117. 李江华 杨书新 刘利峰.基于概念格的文本聚类[J]. 计算机应用, 2008,28(9): 2328-2330
118. 孔攀 邓辉文 江欢 黄艳艳.改进的基于核函数的模糊聚类算法[J]. 计算机应用, 2008,28(9): 2338-2340
119. 周大镗 吴晓丽 闫红灿.一种高效的多变量时间序列相似查询算法[J]. 计算机应用, 2008,28(10): 2541-2543

文章评论 (请注意:本站实行文责自负, 请不要发表与学术无关的内容!评论内容不代表本站观点.)

反馈人	<input type="text"/>	邮箱地址	<input type="text"/>
反馈标题	<input type="text"/>	验证码	<input type="text" value="4429"/>