

杨 静,李文平,张健沛.大数据典型相关分析的云模型方法[J].通信学报,2013,(10):121~134

## 大数据典型相关分析的云模型方法

### Canonical correlation analysis of big data based on cloud model

投稿时间: 2013-04-21

DOI: 10.3969/j.issn.1000-436x.2013.10.015

中文关键词: [大数据](#) [典型相关分析](#) [云模型](#) [云运算](#) [云计算](#)

英文关键词: [big data](#) [canonical correlation analysis \(CCA\)](#) [cloud model](#) [cloud operation](#) [cloud computing](#)

基金项目: 国家自然科学基金资助项目(61370083, 61073043, 61073041); 高等学校博士学科点专项科研基金资助项目(20112304110011, 20122304110012); 黑龙江省自然科学基金资助项目(F200901); 哈尔滨市科技创新人才研究专项基金(优秀学科带头人)资助项目(2011RFXXG015)

作者

单位

[杨 静, 李文平, 张健沛](#)

[哈尔滨工程大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001](#)

摘要点击次数: 247

全文下载次数: 82

中文摘要:

针对传统大数据典型相关分析(CCA, canonical correlation analysis)方法的高复杂度在面临大数据PB级数据规模时不再适应的现状,提出了一种基于云模型的大数据CCA方法。该方法在云计算架构的基础上,通过云运算将各端点云合并为中心云,并据此产生中心云滴,以中心云滴作为大数据的不确定性复原小样本,在其上施以CCA运算,中心云滴的较小数据量提高了运算效率。在真实数据集上的实验结果验证了该方法的有效性。

英文摘要:

The complexity of traditional CCA methods is too high to meet the requirements to analyze big data due to their huge scale which is reaching the level of peta-byte. A novel approach to CCA was proposed to mine the big data by introducing the cloud model which is a brand-new theory about the uncertainty artificial intelligence. A distributed architecture based on cloud computing was established. All of the clouds distributing on the nodes of the distributed architecture were combined to a center cloud via cloud operation (where cloud is a synopsis of data and which is a concept coming from the cloud theory). A type of virtual sample of data called cloud drops created based on the center cloud. Finally the computing of CCA was imposed on the cloud drops. The CCA was imposed on the cloud drops with less volume, which improves the efficiency. Experimental results on real data sets indicate the effectiveness of this method.

[查看全文](#) [查看/发表评论](#) [下载PDF阅读器](#)

关闭

版权所有:《通信学报》

地址:北京市丰台区成寿寺路11号邮电出版大厦8层814室 电话:010-81055478, 81055479

81055480, 81055482 电子邮件: [xuebao@ptpress.com.cn](mailto:xuebao@ptpress.com.cn)

技术支持:北京勤云科技发展有限公司