

网络中心在云计算、大数据分析的关键技术研发及服务创新上取得系列成果

文章来源：计算机网络信息中心

发布时间：2014-09-22

【字号：小 中 大】

中国科学院计算机网络信息中心科学数据中心通过需求牵引在云计算及大数据分析的关键技术研发及服务创新上取得系列成果。

在云服务平台方面，网络中心科学数据中心云计算团队以满足用户的需求和云平台的稳定性、可靠性为目标进行关键技术研发，共申请国家发明专利3项；其中“一种集群系统应用软件快速部署方法”的专利在2014年9月17日获得授权（已应用于云计算平台ECCP，实现操作系统和相关应用20分钟内在120个结点的快速部署），其他2项发明专利正在实质审查阶段。

在大数据分析方面，科学数据中心大数据分析团队连续在国际数据挖掘著名会议发表多篇论文，相关成果也支持学科领域的科研人员在学术期刊*Nature Communications* 和*Lancet* 发表论文，服务于科研创新。

兰金松、黎建辉等基于出租车产生的海量轨迹时空数据，对城市交通异常事件的检测算法进行了深入的研究。把热量传播模型（Heat Diffusion Model）引入到交通异常扩散的研究领域，以模型来描述异常扩散的现象，提出异常源发现算法；并在真实的数据上进行实验，验证了算法的准确性和高效性；相关成果发表在国际数据挖掘一流会议SDM 2014上。

随着开源项目的兴起，如何定位软件工件（software artifact）是软件工程领域的重要研究方向之一。传统软件工件的检索是基于文本查询的方式，郭旦怀、熊辉、周园春等在对大量开源代码分析的基础上，抽取代码在程序之间的结构特点等相关属性，并提出属性之间的距离的度量算法，最后在大量的开源代码进行测试，验证了算法的有效性，相关成果已被国际数据挖掘著名会议ICDM 2014接收为regular paper（录用率为9.7%）。

H7N9最近两年在中国的爆发，是近年来中国较大影响的公共卫生事件。患者的时空分布，活禽市场与病例爆发的之间的关联。是关系到流行病急需回答的问题。郭旦怀等支持中国CDC和国际流行病学专家为此做出定量的回答，为H7N9的迅速控制与政府决策提供了依据，相关成果发表于医学期刊*Lancet*

(<http://www.sciencedirect.com/science/article/pii/S0140673613619042>) 和*Nature communication*

(<http://www.nature.com/ncomms/2014/140617/ncomms5116/full/ncomms5116.html>) 等刊物。

狂犬病是一种人犬共患的高度致死性流行病，中国是世界第二大狂犬病受害国，郭旦怀、周园春、黎建辉等人运用数据挖掘的方法重新定义了狂犬病的时空聚集，提出了为更高效地进行狂犬病防控修改原来的防控标准的建议，相关成果发表在*Plos One*等期刊

(<http://www.plosone.org/article/info:doi/10.1371/journal.pone.0072352>) 上。

在大数据时代，空间数据的传输效率是许多应用的制约瓶颈，空间数据的压缩引起了极大的关注，郭旦怀等积极组织参加ACM GIS CUP 2014中的空间数据的压缩竞赛，与来自十多个国家的同行同台竞技，最终脱颖而出，初次参赛取得了第二名的好成绩。

