

## 农业云服务可适性技术研究进展\*

陈天恩<sup>1</sup>, 刘军萍<sup>2\*</sup>, 王登位<sup>1</sup>, 史晓慧<sup>1</sup>

(1. 北京农业信息技术研究中心 / 农业部农业信息技术重点实验室, 北京 100097;

2. 北京市农村工作委员会信息中心, 北京, 100081)

**摘要:**【目的】作为农业信息化应用新的交付模式, 农业云服务在降低农业信息化技术及成本门槛、响应个性化服务需求等方面具有明显优势。但农业行业应用的特殊性、用户需求地区差异性和动态变化性等特征, 也对通用云计算应用模型的适用性构成挑战, 研究突破可适性技术瓶颈, 是云计算在农业领域深入应用所面临的首要问题。【方法】文章围绕农业云服务涉及的数据管理、知识发现和服务提供等关键环节, 系统阐述了农业云服务可适性技术在农业数据存储方法与模型、农业数据挖掘与知识发现、农业知识服务组合、农业应用按需服务这4个方面的研究进展情况, 并提出了一种可适性农业云服务参考模型。【结果/结论】揭示了云服务可适性技术的研究潜力, 为开展云计算行业应用和大规模农业云服务应用研究提供参考。

**关键词:** 农业云服务; 可适性; 农业时空数据; 一体化存储; 农业知识发现; 知识服务选择与优化组合; 按需服务

DOI: 10.12105/j.issn.1672-0423.20180107

## 0 引言

近30年以来, 互联网的迅速发展在世界范围催生了诸多以网络为基础的经济发展模式 and 新型产业业态。特别是近10年来, 互联网已经从信息互联互通的渠道逐渐演进为世界各国经济发展不可或缺的基础设施, 一批互联网经济体已经成为中美等国经济发展中最具活力的重要组成部分。在此背景下, 基于互联网、特别是移动互联网的信息(技术)服务业正在成为重要的战略性新兴产业<sup>[1]</sup>。研究探索网络环境下应用服务本身的科学问题与系列关键技术, 实现网络应用的服务质量提升, 提供更好的应用服务用户体验, 正成为一个越来越受到重视的研究方向。

当前各类互联网应用服务场景中, 用户需求的多元化和动态化, 与网络服务资源供给技术、模型之间存在突出的矛盾, 映射在行业互联网应用中, 行业特殊性及其应用服务个性化需求与云计算为代表的互联网服务提供共性技术之间的矛盾, 制约了互联网行

收稿日期: 2018-02-03

第一作者简介: 陈天恩(1978—), 男, 汉族, 江苏常州人, 博士, 研究员。研究方向: 农业数据智能分析处理技术、农业云服务可适性技术、农业信息化标准的研发与应用推广。Email: chente@nrcita.org.cn

※ 通信作者简介: 刘军萍(1965—), 女, 汉族, 山西忻州人, 博士, 研究员。研究方向: 农业农村信息化。Email: liujunping@bjaginfo.gov.cn

\* 基金项目: 北京市自然科学基金重点项目“农业云服务可适性关键技术及应用模型研究”(4151001); 北京市科技计划课题“服务乡村振兴的都市农业双创平台建设与成果示范应用”(Z181100002418002)

业应用的深入发展，也对行业相关产业借助互联网创新发展造成了障碍。在此背景下，互联网应用服务的可适性问题被提出，用以描述互联网服务需求多变本质与服务系统计算模型复杂难控现象之间的矛盾。互联网应用服务的可适性与服务背后的数据管理模式、平台构建模式和应用服务模型密切相关。

农业云服务作为云计算技术面向农业行业应用的载体，也是农业信息化应用交付模式不断演进的最新结果。近年来，随着农业互联网应用的爆发式增长，通用的标准云服务技术在农业领域的适用性问题得到越来越多的关注，农业行业网络应用的特殊性、需求的地区差异性和动态变化性等，对标准云计算模型的可适性构成了事实上的挑战。农业生产的特点决定了农业云服务必须依赖多源异构农田数据，以农业知识为核心服务内容，并最终通过 SaaS (Software as a Service, 软件即服务) 或按需信息服务的形式落地。由此，农业云服务模型的构建需要突破“数据—知识—服务”3 个层面的可适性关键技术瓶颈，以确保服务的高效能、高可靠性和高可适性。因此，研究突破农业云服务可适性技术瓶颈，是云计算在农业领域发挥作用面临的首要问题。

近年来，学术界针对农业云服务背后的农田数据存储方法与模型、农业知识发现算法、服务组合及平台构建模式，以及按需服务模型开展了深入研究，取得了积极的进展。

## 1 农业云服务可适性技术

### 1.1 农业数据存储方法与模型研究

农业数据内容和来源较为复杂，其中与农业生产和作物生长密切相关的多尺度农田环境数据（光、温、水、气、热）、土壤数据（温度、湿度、质地、养分、电导率、地下水位等）、作物数据（基因、植株生理、群体特性）、遥感数据等兼具时间和空间的耦合特性，且农业决策主要依赖数据时空特性分析应用。当前，以 GPFS、Google GFS 等为代表的海量结构化数据高效存储管理技术已经较为成熟，但是海量时空耦合数据的存储模型及技术研究相对落后，多数情况下人们采用单独考虑数据的空间特性或时序特征的办法进行描述，并以空间或时态数据库形式存储；Google、Microsoft 等大型云服务商解决了基于 NoSQL 技术的 PB 级别栅格数据存储管理技术，并提供相应的地图服务产品；ESRI 的 Geodatabase 及 MrSID 栅格数据模型可以实现海量矢栅空间数据高效率的存储和访问；国内 2012 年张桂刚等<sup>[2]</sup>提出了一种基于互联网计算模型的海量空间信息处理的云模型，2013 年陈崇成等<sup>[3]</sup>引入分布式图数据库 Neo4J 和并行图计算框架，实现矢量栅格数据一体化的海量空间数据分布式云存储管理与访问，2010 年赵春江等在 863 计划的支持下，研究了海量农业知识资源的空间服务技术，支撑了北京农业 221 平台的建设。在时态数据存储研究方面，美国 J. Ben Zvi 于 20 世纪 80 年代对时态数据库作了开创性研究，AspenTech 的 InfoPlus 时态数据库实现了时序属性数据的高水平压缩和快速存取；2013 年刘东宁等<sup>[4]</sup>提出一种时态信息模型，重点描述了时态数据模型和时态知识表达模式，并开发了时态数据原型系统，此外，地学领域的学者从 GIS 技术发展的角度提出了实时 GIS 时空数据模型的设想<sup>[5]</sup>。单独考虑数据的时间和空间特性有利于设计数据存储结构，但

2018年2月

不利于从应用角度出发实现数据的高效检索和同时基于时空序列特征的处理。农业数据在时间和空间上的变异不是互相孤立的，它们之间存在紧密的耦合关系，变化的相互影响较大，要发掘其后隐藏的规律指导农业生产，必须复合考虑时空的变异，现有空间及时态特性分离的数据存储管理技术无法支撑。特别随着最近几年卫星遥感和专用传感器/传感网技术在农业领域的成熟应用，数据获取更加便捷，更加廉价，农业数据迅猛增长，实时感知数据正在农业生产分析决策中占据越来越重要的地位，开展适用于农业云服务的海量时空耦合数据的一体化存储模型与分布式存储技术研究显得尤为紧迫和重要。

2015年以来，国内外学者主要针对传统数据存储冗余大、检索效率低、属性关联弱等问题进行了研究。王剑等<sup>[6]</sup>尝试在Hadoop框架下研究了面向海量农业数据的分布式计算与存储技术，提出了基于Hadoop的相邻区域边缘化存储和混合式索引分布与检索技术，采用了“中心控制节点—数据节点”的存储体系，通过报文通信技术和混合式索引分布策略，实现了对海量数据的高并发式存储和检索。2016年封孝生等<sup>[7]</sup>针对在HBase中如何进行有效的半结构化时空数据存储和查询问题展开研究，对该问题进行形式化描述，并利用半结构化处理方法TwigStack提出HBase的半结构化时空数据存储模型，在此基础上开展了半结构化的时空范围查询和kNN查询。2017年郑浩泉等<sup>[8]</sup>实验得出，按照数据访问的特点为轨迹数据挖掘算法选择合适的轨迹存储方法，可以有效地提高挖掘算法的执行效率。2018年Mohamed A.Gad等<sup>[9]</sup>提出通过沿轨迹调整和空间插值相结合的方法，可确定在任何给定时间内的空间可变散射数据的位置。

为了提高时空数据的查询和分析效率，2017年夏慧琼等<sup>[10]</sup>提出一种基于拉普拉斯特征映射(Laplacian eigenmaps, LE)的时空划分方法，该方法能够较好地保持地理对象时空关系的完整性，能够获取数据在低维流形空间中的结构特征，使得时空信息的组织和管理变得更加容易。2017年程星华等<sup>[11]</sup>探索了多粒度时空对象的建模方法，对政区进行了多粒度时空对象表达，并采用可视化表达的方法进行了实验验证，能够更好地支持时空对象的查询、分析和可视化表达。在建立时空数据存储管理模型方面，2016年Hung-Ming Chen<sup>[12]</sup>提出了一种基于云计算的系统架构基于BigTable和MapReduce的数据存储和处理模式，提供基于网络的服务存储和分析海量的建筑信息模型(BIMS)。2017年李寅超等<sup>[13]</sup>提出了基于对象和快照的混合时空数据模型，用面向对象描述地表覆盖的斑块对象时空过程，组织管理斑块对象时空事件和空间、属性信息，同时用快照描述地表覆盖整体时空分布，组织管理栅格快照，两者通过基于时间和空间位置的逻辑关联关系形成混合模型。综上，农业海量时空耦合数据存储模型和分布式存储技术的研究已经取得一定进展，具有实用价值的时空耦合农田数据模型研究有望在未来几年取得突破。

## 1.2 农业数据挖掘与知识发现技术研究

农业知识是农业云服务的核心内容，如何利用云计算能力实现更高水平的农业数据挖掘和知识发现，是农业云服务的重要使命之一。然而，农业是一个异常复杂的巨系统，农田土壤类型众多，作物品种复杂，环境辅复杂多变，病虫害发生频繁且症状不断变化，土、肥、水、作物以及气候相互之间的关系错综复杂，许多规律还未被人们所准确认知，农业数据具有大量、多维、动态、不完整、不确定等特性，其之多之复杂是任何其他一

个领域所没有的。这给农业数据挖掘研究带来很大的挑战。目前农业数据挖掘研究基本停留在对单一来源数据挖掘的研究层次，2009 年陈桂芬等<sup>[14]</sup>利用空间模糊聚类算法等空间数据挖掘技术，提出了解决玉米精准施肥、土壤肥力评价、地力等级分类和产量预测等问题，2008 年郑向群等<sup>[15]</sup>基于数据仓库和工作流挖掘技术开展了土壤肥力评价规则挖掘算法研究，指导农业生产科学施肥；2005 年查骏雄<sup>[16]</sup>提出了一种土壤侵蚀分析的数据挖掘方法，通过对大量的水土保持信息数据的分析，从中抽取潜在的土壤侵蚀的变化规律与变化模式；2011 年陈天恩等<sup>[17]</sup>利用地统计分析方法对不同尺度下土壤养分的空间自相关规律进行挖掘，得出不同尺度下土壤养分采样的最优布局策略，2004 年赵恽甦等<sup>[18]</sup>采用模糊评价算法在综合考虑农业环境各污染物因子的基础上按地区实际条件评测环境，并采用 Apriori 算法找出污染物因子之间的相关性，从而挖掘出导致环境差异的原因。此外，李干琼等<sup>[19]</sup>、许世卫等<sup>[20]</sup>针对农产品市场行情数据进行挖掘研究，提出了价格预警与行情趋势分析的算法。

2014 年黄魁建等<sup>[21]</sup>结合本体推理以及数据挖掘技术，实现标准知识查询和 GAP 控制点的择优推荐；2016 年方永美等<sup>[22]</sup>针对蔬菜质量安全，以蔬菜农药残留检测数据为基础，基于贝叶斯网络数据挖掘进行研究，得到不同品种蔬菜的农药残留含量分布，以及一品种在不同时间的分布趋势；2016 年朱玉梅<sup>[23]</sup>针对棉花上的棉铃虫、棉叶螨、棉蚜 3 种害虫的发生趋势进行综合分析，采用 Apriori 算法数据挖掘关联规则，得出气候因素与棉花三大害虫的发生有密切的关系；2017 年丁晓颖等<sup>[24]</sup>通过各种传感器采集日光温室种植过程中的关键要素数据，利用数据挖掘技术，挖掘出本地设施内的光、温、水、气等多种作物生长因子的变化规律，搭建了远程专家诊断指导系统；2015 年曹敏杰<sup>[25]</sup>针对浙江近岸海域海洋生态环境时空分析及预测关键技术进行研究，设计了海洋生态环境三元混合时空格局分析与关联规则的挖掘框架，形成了从多源监测数据管理、时空分析与挖掘、动态预测预警的服务链条；刘军萍等 2016 年起基于部署的探针和北京移动大数据平台等第三方数据进行了北京房山区幽岚山景区的游客行为分析和景点竞争力评价指数研究。2017 年樊磊等<sup>[26]</sup>基于多源数据的土壤水分估算及森林火灾风险评估应用进行研究，结合地面观测、光学遥感、微波遥感等多源数据，提高了近地表土壤水分的空间分辨率及估算精度；2017 年顾侃等<sup>[27]</sup>基于惯性检验的序列数据趋势识别方法、序列复合弯曲距离的趋势相似性度量方法及趋势事件的多维关联规则，研究了序列型数据中的趋势性知识。

从农业云服务应用的角度看，农业数据挖掘和知识发现的需求主要集中在两个方面，一是面向农业特定应用需求汇聚多源异构数据实现聚合分析挖掘，二是农田实时感知数据的在线挖掘。限于算法自身结构和效率，传统农业数据挖掘算法不仅不具备针对多源或实时数据进行在线分析挖掘的能力，也不具备利用云端并行计算特点的能力。随着卫星 / 无人机遥感、传感器和个体识别技术在农业领域的大量应用，实时在线感知农业数据前所未有地呈现出爆发增长态势，如何充分利用多源实时数据服务为农业生产提供决策成为迫切需要解决的问题，突破云服务条件下的多源农田感知数据在线挖掘和知识发现方法，将为上述问题的解决提供关键支撑。

2018年2月

### 1.3 农业知识服务组合技术研究

农业知识服务组合作为面向服务架构 (Service-Oriented Architecture, SOA) 的主要实现技术,旨在实现不同功能知识服务的有机聚合,以形成功能更加强大的粗粒度服务,是实现农业 SaaS 按需服务的关键支撑技术之一,也是构建良好的农业云服务知识资源汇聚体系的重要手段。服务组合技术主要分为基于流程建模、基于语义网、基于部署的3种基本策略。基于流程建模分为过程驱动和模型驱动两种类型,由工作流模型、状态演算模型、进程代数模型和知识复用模型等,代表方法主要有 BPEL4WS、Petri 网模型、 $\pi$ -演算模型、ViPen、MDA 等;基于语义网主要是语义驱动类型,建立语义 Web 模型,代表方法主要有 CoSMoS、YASA-M 等<sup>[28-31]</sup>;基于部署的主要是服务部署驱动类型,由 P2P 架构模型和网格计算模型等,代表方法主要有 SELF-SERV、WSRF 等。

目前,学术界的研究兴趣主要集中在流程驱动的服务组合机制、任务驱动的服务组合机制,以及服务组合结果的评价方法等方面<sup>[31-35]</sup>。在流程驱动的服务组合中,研究相对最为成熟是静态人工绑定服务组合方式<sup>[36]</sup>,该方式具有很高的执行效率,已有 SOAP、UDDI、WSDL、BPEL 等一套完善的协议支持,IBM 的 WebSphere Enterprise Service Bus (WESB) 企业服务总线,SAP 的 SAP NetWeaver 系统和 Oracle 的 Fusion Architecture 等都是基于该服务组合方式实现,由于在大规模服务环境下缺少灵活性和针对动态环境的 QoS 自动优化,该方式将逐步被执行前自动绑定服务组合的方式取代<sup>[37-43]</sup>。执行前自动绑定服务组合方式在服务建模时只考虑流程和接口问题。等到需要执行服务组合时,再从候选实体服务库中依据优化目标函数和约束,挑选出最适合的一组实体服务执行,在灵活性、适用性和扩展性方面,都有了很大提升<sup>[44-48]</sup>。国内外学者围绕该组合方式及其服务选择策略开展了很多研究<sup>[49]</sup>,Anis Charfi<sup>[50]</sup>提出了一种集成规则到 Web 服务组合中的方法,并特别针对于 BEPL 进行了说明;Colom<sup>[51]</sup>在 SCENE 系统中提出了使用 ECA 规则模型解决在服务组合过程中服务出错和环境变化的情况下绑定不同的服务的问题;国内学者汤景凡<sup>[52]</sup>针对于传统服务组合模型进行了扩展;王文彬等<sup>[53]</sup>提出了一种面向动态 Web 服务选择的离散微粒群算法,并结合服务选择研究背景,提出了3种速度计算算子和1种位置进化方程;张成文等<sup>[54]</sup>提出一种用于 QoS 感知的 Web 服务选择的遗传算法,该算法采用关系矩阵编码方式,克服了一维编码方式表示的局限性;张燕平等<sup>[55]</sup>提出一种基于离散粒子群智能优化算法的 DDPSO 算法,以解决动态 Web 服务组合问题,且首先引入 Skyline 技术来剔除冗余候选服务,以降低服务选择时空开销;承松等<sup>[56]</sup>研究了混沌蚁群算法的 Web 服务组合,避免优化过程中出现局部最优解。执行前自动绑定服务组合模式能够根据 QoS 有效地自动选取实体服务,很好地顺应了云计算环境下在线服务快速增长的潮流,但是流程驱动的服务组合只适用于想对固定的组合流程,不能避免人工参与设定,并且在运行期间组合流程不会发生变化。而任务驱动的服务组合方法强调任务即时求解和流程自动生成,通过服务注册时的自描述信息,利用语义、人工智能和图论中的相关方法,通过计算机的自动推理,进行服务组合,逐渐成为新的研究热点。国外学者先后提出了基于语义和基于图论的服务组合方法<sup>[57-59]</sup>,国内邓水光等<sup>[60]</sup>提出了基于回溯树的 Web 服务自动组合,大大降低了搜索的范围和搜索的复杂度,加快

了服务组合的效率；张俊娜等<sup>[61]</sup>提出融合网络环境下快速可靠的服务组合容错方法，在故障排除率、故障处理时间与组合最优度方面均有所提高。基于任务的自动服务组合方式也面临着在规模化应用背景下服务关系构建和求解时间开销成指数级上涨的困境，往往无法在可接受时间内完成求解。王磊等<sup>[62]</sup>基于区块链机制的云计算环境下服务组合策略的研究，提出了跨服务覆盖层的基于链路预测生存时间及服务强度最优策略的高效服务路径生成算法等。

农业云服务环境下的知识服务组合主要用于 SaaS 系统的按需构建，在农业云服务大规模知识服务环境下，充分考量农学专家等服务提供者的服务策略和服务组合的效益水平，兼顾组合的灵活性和执行效率，降低服务实体变动情况下的重新组合成本，对于提高农业云服务 SaaS 系统的质量和服务水平具有重要意义。

#### 1.4 农业应用按需服务技术研究

按需服务的前提是了解不同类别用户消费服务的习惯差异，基于用户的信息使用行为、习惯、偏好和特点，提供满足用户个性化需求的服务。农业本身具有强烈的地域性和季节性，其种植类型与种植管理技术存在巨大差异。加上农业政府部门管理者、生产者、经营者、消费者对农业知识的不同理解和诉求，农业云服务消费者的需求呈现出鲜明的差异化、个性化、动态化特征。满足这类差异化需求需要农业云服务平台具备提供高度个性化服务的能力。研究如何整合、分析大量农业云服务用户兴趣爱好的特征，建立更加贴合用户偏好需求的用户模型，提升信息推送技术和定制服务技术，是整个农业云个性化服务体系的关键<sup>[63-65]</sup>。

当前，针对农业云计算个性化服务的用户模型研究还相对较少；大多数个性化系统都是采用传统的关键词或关键词向量空间模型表示用户的兴趣偏好。这种方法虽然理论成熟、计算方法简单，但并不能十分准确地反映语义信息，制约了不同系统中个性化信息的共享以及用户模型的重用<sup>[66-70]</sup>。如重庆工学院提出了一种个性化服务用户模型构架，建立了数字图书馆个性化服务用户模型；宋雯斐等<sup>[71]</sup>人构建了基于层次语义情景的用户模型，提出了针对数字图书馆个性化服务地用户兴趣模型。为了实现用户模型的共享与重用性，近年来，国内外开始针对重用性用户模型的研究<sup>[72-93]</sup>；2004 年德国 Fraunhofer IPSI 研究院的 Bhaskar 等人提出了一种支持跨系统个性化服务的方法，以及用统一上下文模型和上下文护照实现的跨系统个性化用户建模技术；2005 年 Heckmann D. 等人开发了通用用户本体 GUMO，用于智能语义丰富环境中的分布式用户模型的统一描述<sup>[94]</sup>。但以上的研究都存在一定局限性，前者要求服务提供方必须遵循 CSCSP 协议而且必须采用 UUCM 模型的子集作为系统的用户模型；后者的使用则需要建立不同领域的应用机制。基于以上两方面的分析，2014 年刘飞等<sup>[95]</sup>提出一种面向个性化云服务基于用户类型和隐私保护的信任模型，引入信任评估代理作为信任评估的主体，并且设计了基于用户类型的信任值评估方法；2015 年 Amin Javari 等<sup>[96]</sup>提出了一个可调节的多样性和精确度的混合模型，该模型在解决数据精度与多样性并存的情况表现较好，可广泛应用于实际商用系统中。2015 年张启宇等<sup>[97]</sup>重点研究了用户类别兴趣向量、用户特征词喜好向量和文档特征向量，提出了个性化服务推荐算法，该模型可根据用户兴趣制定推

2018年2月

荐,为用户提供有价值的信息,满足用户个性化需求;2016年杨凤萍等<sup>[98]</sup>通过产品特征属性与用户特征属性的映射,建立用户偏好模型,引入神经网络集成的机器学习方法来模拟偏好模型。2016年 Moon-Hee Park 等<sup>[99]</sup>提出了一个基于地图的个性化推荐系统,系统从移动设备收集上下文信息、位置、时间、天气和用户请求,并通过显示在微型地图上,推荐出最合适的项目以提供适当的服务。

在信息按需推送技术方面,2015年高升等<sup>[100]</sup>综合分析了不同领域内用户对信息对象的评价数据,并通过传递、共享跨域共性特征信息的方式缓解了目标领域的信息稀疏性问题<sup>[95]</sup>;2017年崔艳萍等<sup>[101]</sup>提出依据用户浏览行为将其兴趣度的作用划分为正、逆两种方向,分析互联网用户兴趣度属性之间的依赖关系对信息推送问题建模,提出一种基于 MQTT Proxy 的互联网信息资源用户获取优化推送方法;2017年余本功等<sup>[102]</sup>设计了基于用户偏好的产品知识设计的混合推送算法,将知识文本向量空间化并由本体标注,采用设计任务与用户偏好共同构造知识需求向量,通过计算知识需求向量与知识文本向量的语义相似性得到知识推送结果。在信息定制服务技术方面,2016年张驰等<sup>[103]</sup>通过设计标准的数据通讯协议实现异构物联网监测设备的统一接入;采用组件库为系统提供独立的功能单元,复用软件资源;通过云服务的方式对组件进行组装集成,快速响应用户需求;2017年刘军萍等围绕北京休闲农业数据开展进一步的深入挖掘和分析,提出休闲农业服务指数计算方法,以及休闲游消费者偏好信息服务模型,提高了信息化服务的质量。孙琦宗等<sup>[104]</sup>对产品特征类别、客户需求、用户情感偏好等方面进行了分析,并建立了产品特征与用户需求关系的映射模型,提出了一种应用于个性化定制服务中的定制优先级的判定方法。

农业云知识需求用户种类繁多,不同农业从业人员的知识水平,所处环境的差异性也较大,目前针对农业云知识需求用户特点的个性化服务用户模型研究报道很少。此外,现有的个性化服务用户模型大多都是基于 Web 浏览的用户模型;而影响农业云知识推送用户需求的因素不仅包含用户自身行为兴趣,还受到环境地理位置等客观因素的影响;因此针对农业云知识需求用户的特点,研究通用的个性化服务策略,构建农业云用户需求模型,对于实现农业知识按需云服务具有重要价值。

## 2 可适性农业云服务参考模型研究

2014年国家农业信息化工程技术研究中心赵春江、陈天恩等首次提出了可适性农业云服务的概念,并对海量时空耦合农业数据一体化存储技术、农田感知数据知识在线发现理论、农业知识服务组合优化模型和地域性差异化需求快速响应机制等农业云服务可适性关键技术开展了较为全面、系统的研究,提出了一种可适性农业云服务参考模型。其中在农业数据描述与存储模型方面,分析了农业多源高维数据的时空属性特点和农田数据时态拓扑关系,构建了农田时空耦合数据统一描述模型;研究了多层次一体化农业时空数据资源虚拟化技术,实现了多类型资源统一管理与存储模型;为解决海量农业数据的检索效率问题,提出一种多级时空索引树结点分裂更新方法。在农田感知数据在线

挖掘方面, 研究分析了农业环境监测数据多源、异构、强实时性等特点, 提出了包括时空变量的时序分解、变异函数建模和时空 Kriging 插值等算法在内的时空变异规则挖掘与预处理方法; 分析了结构化作物生长环境参数与非结构化图像、视频等跨媒体数据特点, 围绕作物病害监测提出一种基于深度神经网络的识别模型, 围绕菇房异常环境预警提出一种基于模糊推理算法的方法, 开发了多参数跨媒体数据的农作物病害智能决策系统原型。在农业知识服务组合优化方面, 构建了情景感知用户兴趣模型与农业知识服务本体结构, 实现了基于农业知识服务按需选择与最优化动态组合的施肥决策模型; 通过农户需求与农业知识相似度分析实现专家知识提炼与知识服务组合, 开发了基于农业实时数据的畜禽行为诊断按需服务系统原型。在农业知识按需服务方面, 通过融合情景感知信息和用户兴趣模型, 建立了病害、疾病、品种、农资、质量安全等领域农业知识图谱, 实现了基于情景感知和协同过滤的信息推荐算法; 提出一种农产品产销媒合模型, 开发了面向多业态、多用户的农产品供应链个性化服务系统原型。在技术研究的基础上, 结合国家重大科研任务提出了一种农业云服务参考模型, 构建了国家农科城云服务平台, 面向全国农业科技园区发布了在线农技服务、设施农业生产管理、寒地水稻精准施肥、生鲜农产品冷链物流监管服务等 10 余项农业云服务应用。

### 3 总结

由于农业地域差异大、涉及面广, 应用需求及专业技术复杂多变, 农业云服务研究与应用还处于起步阶段。云计算在农业领域发挥更大作用必须解决可适性关键问题, 当前迫切需要结合农业生产、经营、管理重大应用需求, 围绕农业云服务体系架构的完善和服务模型的形成, 构建科学合理的农业云计算可适性服务模型, 在海量农业时空数据高效管理、农田多源感知数据在线挖掘、农业知识服务组合优化和地域性差异化需求快速响应等关键技术取得实质性突破, 才能为构建高效能、高可适性的农业云计算服务模式提供基础技术支撑, 进而促进云计算技术和农业生产实现深度融合和良好衔接, 使云服务技术真正融入我国农业现代化发展进程, 并为改造和提升传统农业发挥重要作用。

### 参考文献

- [1] 陈伟.《促进大数据发展行动纲要》解读. 中国信息化, 2015, 10(10): 11~14.
- [2] 张桂刚. 一种基于海量信息处理的云存储模型研究. 计算机研究与发展, 2012, 49: 32~36.
- [3] 陈崇成. 基于 NoSQL 的海量空间数据云存储与服务方法. 地球信息科学学报, 2013, 15(2): 166~174.
- [4] 刘东宁. 基于时态数据库的极小子结构逻辑系统. 计算机学报, 2013, 36(8): 1592~1601.
- [5] 龚健雅, 李小龙, 吴华意. 实时 GIS 时空数据模型. 测绘学报, 2014, 43(03): 226~232, 275.
- [6] 王剑, 黄朝光, 王健, 等. 面向农业科学数据的分布式存储方法研究. 计算机工程与应用, 2016, 52(11): 248~253.
- [7] 封孝生, 张翀, 陈晓莹, 等. HBase 中半结构化时空数据存储与查询处理. 国防科技大学学报, 2016, 38(3): 174~181.
- [8] 郑浩泉, 何浩奇, 刘伽椰, 等. 时空轨迹数据存储方法研究. 南京师大学报(自然科学版), 2017, 40(03): 38~44.
- [9] Mohamed A. Gad, Mai H. Elshehaly, DenisGračanin et al. A Tracking Analyst for large 3D spatiotemporal data from

2018年2月

- multiple sources. *Computers & Geosciences. Automation in Construction*, 2018 (111): 283~293.
- [10] 夏慧琼, 林丽群. 拉普拉斯特征映射的时空数据划分方法. *测绘科学*, 2018 (06): 1~9 [2018-03-08]. <http://kns.cnki.net/kcms/detail/11.4415.P.20180122.1802.054.html>.
- [11] 程星华, 胡迪, 俞肇元, 等. 政区多粒度时空对象建模及其 Geodatabase 实现. *地球信息科学学报*, 2017, 19 (09): 1228~1237.
- [12] Hung-Ming Chen, Kai-Chuan Chang, et al. A cloud-based system framework for performing online viewing, storage, and analysis on big data of massive BIMs. *Automation in Construction*, 2016: 34~48.
- [13] 李寅超, 李建松. 一种面向 LUCC 的时空数据存储管理模型. *吉林大学学报 (地球科学版)*, 2017, 47 (01): 294~304.
- [14] 陈桂芬. 面向精准农业的空间数据挖掘技术研究与应用. 吉林大学, 2009.
- [15] 郑向群, 赵政, 刘东生. 基于数据仓库的土壤环境监测综合挖掘模型构架. *农业工程学报*, 2008 (08): 162~168.
- [16] 查骏雄. 数据挖掘在土壤侵蚀分析系统中的应用. *南昌工程学院学报*, 2005 (02): 46~48, 76.
- [17] 陈天恩, 董静, 陈立平, 等. 县域农田土壤采样布局多目标优化分析模型. *农业工程学报*, 2012, 28 (23): 67~73.
- [18] 赵恽甦. 数据挖掘在农业环境中的应用. 天津大学, 2004.
- [19] 李干琼, 许世卫, 李哲敏, 等. 蔬菜市场价格短期波动影响因素分析—基于 VAR 模型的实证研究. *中国食物与营养*, 2013, 19 (3): 45~49.
- [20] 许世卫. 农业大数据与农产品监测预警. *中国农业科技导报*, 2014, 16 (5): 14~20.
- [21] 黄魁建, 郝文革, 邢文英, 等. 基于本体推理和数据挖掘的良好农业规范专家系统. *中国农学通报*, 2014, 30 (35): 261~268.
- [22] 方永美, 熊俊涛, 杨振刚, 等. 基于贝叶斯网络数据挖掘的蔬菜质量安全分析. *湖北农业科学*, 2016, 55 (23): 6253~6257.
- [23] 朱玉梅. Apriori 算法在棉花病虫害分析中的应用. *农学学报*, 2016, 6 (11): 23~26.
- [24] 丁晓颖, 史昊玲, 李博. 基于物联网的设施农业环境监控与数据挖掘技术的研究应用. *现代农业科技*, 2017 (12): 240~241.
- [25] 曹敏杰. 浙江近岸海域海洋生态环境时空分析及预测关键技术研究. 浙江大学, 2015.
- [26] 樊磊. 基于多源数据的土壤水分估算及森林火灾风险评估应用. 中国科学院大学 (中国科学院遥感与数字地球研究所), 2017.
- [27] 顾侃. 序列数据趋势性知识发现. 北京科技大学, 2018.
- [28] Khalaf R, Mukhi N, Weerawarana S. Service-Oriented Composition in BPEL4WS. //2008.
- [29] Tang X, Jiang C, Zhou M. Automatic Web service composition based on Horn clauses and Petri nets. *Expert Systems with Applications*, 2011, 38 (10): 13024~13031.
- [30] 廖军, 谭浩, 刘锦德. 基于 Pi- 演算的 Web 服务组合的描述和验证. *计算机学报*, 2005, 28 (4): 635~643.
- [31] 倪晚成, 刘连臣, 吴澄. Web 服务组合方法综述. *计算机工程*, 2008, 34 (4): 79~81.
- [32] Ding W, Wang J, Han Y. ViPen: A Model Supporting Knowledge Provenance for Exploratory Service Composition// IEEE International Conference on Services Computing. *IEEE Computer Society*, 2010: 265~272.
- [33] Thomas D, Thomas D, Thomas D, et al. Model driven architecture: how far have we come, how far can we go? //Companion of the, ACM Sigplan Conference on Object-Oriented Programming, Systems, Languages, and Applications. ACM, 2003: 273~274.
- [34] Fujii K, Suda T. Semantics-Based Context-Aware Dynamic Service Composition. *Acm Transactions on Autonomous & Adaptive Systems*, 2009, 4 (2): 12.
- [35] Chahab Y, Tata S, Ozanne A. YASA-M: A Semantic Web Service Matchmaker//IEEE International Conference on Advanced Information NETWORKING and Applications. *IEEE*, 2010: 966~973.
- [36] Sheng Q Z, Benatallah B, Dumas M, et al. Chapter 99-Self-serv: A platform for rapid composition of web services in a peer-to-peer environment//ELSEVIER, 2002: 1051~1054.
- [37] 王勇, 代桂平, 侯亚荣. 信任感知的聚合动态选择方法. *计算机学报*, 2009, 32 (8): 1668~1675.
- [38] 鲍军鹏, 宋楠, 陶斌. 依赖知识的服务组合算法. *西安交通大学学报*, 2013, 47 (8): 1~6.
- [39] 朱兵舰, 彭晓明, 周敏. 静态与动态相结合的 Web 服务组合方法. *空军预警学院学报*, 2009, 23 (1): 75~78.
- [40] 孙维梁, 刘清堂, 杨宗凯, 等. 基于动态 QoS 的 Web 服务组合. *计算机科学*, 2012, 39 (2): 268~272.
- [41] 肖芳雄, 黄志球, 曹子宁, 等. Web 服务组合功能与 QoS 的形式化统一建模和分析. *软件学报*, 2011, 22 (11): 2698~2715.

- [42] 欧伟杰, 曾承, 项小明, 等. 基于概念松弛的高效 Web 服务查询方法. 计算机学报, 2011, 34 (12): 2381~2390.
- [43] 夏虹, 李增智. 粒子群算法求解 Web 服务组合中基于 QoS 的服务选择. 北京邮电大学学报, 2009, 32 (4): 63~66.
- [44] 邓永光, 黄龙涛, 尹建伟, 等. Web 服务组合技术框架及其研究进展. 计算机集成制造系统, 2011, 17 (2): 404~412.
- [45] 周涛. 基于策略的服务组合关键技术研究与应用. 浙江: 浙江大学, 2012.
- [46] 傅兵. 基于 SOA 的数字农务系统关键技术研究. 南京: 南京农业大学, 2012.
- [47] 温嘉佳. Web 服务组合及其相关技术的研究. 北京: 北京邮电大学, 2006.
- [48] 李德胜. 基于 Pi 演算的 Web 服务组合研究. 北京: 北京邮电大学, 2008.
- [49] Y. Liu, M. Zhang, S. P. Ma, et al. User Behavior Oriented Web Spam Detection. In Proceedings of the 17th International Conference on World Wide Web, 2008, 1039~1040.
- [50] Charfi A, Mezini M. AO4BPEL: An Aspect-oriented Extension to BPEL. *World Wide Web-internet & Web Information Systems*, 2007, 10 (3): 309~344.
- [51] Hennig P, Balke W T. Highly scalable web service composition using binary tree-based parallelization//Web Services (ICWS), 2010 IEEE International Conference on. IEEE, 2010: 123~130.
- [52] 汤景凡. 动态 Web 服务组合的关键技术研究. 浙江大学, 2005.
- [53] 王文彬, 孙其博, 赵新超, 等. 基于非均衡变异离散粒子群算法的 QoS 全局最优 Web 服务选择方法. 电子学报, 2010, 38 (12): 2774~2779.
- [54] 张成文, 苏森, 陈俊亮. 基于遗传算法的 QoS 感知的 Web 服务选择. 计算机学报, 2006, 29 (7): 1029~1037.
- [55] 张燕平, 荆紫慧, 张以文, 等. 基于离散粒子群算法的动态 Web 服务组合. 计算机科学, 2015, 42 (6): 71~75.
- [56] 承松, 周井泉, 常瑞云. 混沌蚁群算法的 Web 服务组合优化研究. 计算机技术与发展, 2017, 27 (2): 178~181.
- [57] Aydogan R, Zirtiloglu H. A Graph-Based Web Service Composition Technique Using Ontological Information//IEEE International Conference on Web Services. IEEE, 2007: 1154~1155.
- [58] Dustdar S, Schreiner W. A survey on web services composition. *International journal of web and grid services*, 2005, 1 (1): 1~30.
- [59] Tan W, Fan Y, Zhou M C. A petri net-based method for compatibility analysis and composition of web services in business process execution language. *Automation Science and Engineering*, IEEE Transactions on, 2009, 6 (1): 94~106.
- [60] 邓水光, 吴健, 李莹, 等. 基于回溯树的 Web 服务自动组合. 软件学报, 2007, 18 (8): 1896~1910.
- [61] 张俊娜, 王尚广, 孙其博, 等. 融合网络环境下快速可靠的服务组合容错方法. 软件学报, 2017, 28 (4): 940~958.
- [62] 王磊, 赵晓永. 基于区块链机制的云计算环境下服务组合策略的研究. 2019, 36 (1). [2018-01-10]. <http://www.aocmag.com/article/02-2019-01-027.html>.
- [63] 熊锦华, 虎嵩林, 刘晖. 云计算中的按需服务. 中兴通讯技术, 2010, 16 (4): 13~17.
- [64] 赵春江. 人工智能引领农业迈入崭新时代. 中国农村科技, 2018, 272: 29~31.
- [65] 覃子珍, 毛彧. 农业网络信息资源的个性化服务探讨. 热带农业科学, 2016, 36 (8): 90~99.
- [66] Liu Y, Gao B, Liu T, et al. BrowseRank: letting web users vote for page importance. Proceedings of the 31st Annual international ACM SIGIR Conference on Research and Development in information Retrieval. 2008, 451~458.
- [67] P. Singla, M. Richardson. Yes, there is a correlation - from social networks to personal behavior on the web. In Proceedings of the 17th International Conference on World Wide Web, 2008, 655~664.
- [68] Sadagopan N, Li J. Characterizing typical and atypical user sessions in clickstreams. Proceeding of the 17th international Conference on World Wide Web. 2008, 885~894.
- [69] 岑荣伟. 基于用户行为分析的搜索引擎评价研究. 清华大学. 2010.
- [70] 章毅. 神经网络及在网络用户行为分析中的应用研究. 电子科技大学. 2011.
- [71] 宋雯雯. 基于层次语义情景的用户模型构建. 情报科学, 2013, 31 (6): 24~27.
- [72] Hayes B. Cloud computing. *Communications of the ACM*, 2008, 51 (7): 9~11.
- [73] ZHAO Chunjiang, WU Huarui, SUN Xiang, et al. An Algorithm for Mining Association Rules with Multiple Minimum Supports Based on FP-Tree. *New Zealand Journal of Agricultural Research*, 2007, 50: 1375~1381.
- [74] 王彦集, 陈天恩, 徐刚. 基于 SOA 的农田测绘成图系统的设计和实现. 农业工程学报, 2009, 25 (S2): 267~

2018年2月

- 271.
- [ 75 ] 吴华瑞, 张凤霞, 赵春江. 一种多重最小支持度关联规则挖掘算法. 哈尔滨工业大学学报, 2008, 49 ( 9 ): 1447~1451.
- [ 76 ] 陈康, 郑纬民. 云计算: 系统实现与研究现状. 软件学报, 2009, 20 ( 5 ): 1337~1348.
- [ 77 ] Dean J, Ghemawat S. MapReduce: Simplifier date processing on large clusters. *Communication of the ACM*, 2008, 51 ( 1 ): 107~113.
- [ 78 ] Neumeyer L, Robbins B, Nair A, et al. S4: Distributed Stream Computing Platform. Proc of ICDM Workshops 2010, Piscataway, NJ: IEEE, 2010: 170~177.
- [ 79 ] Goodhope K, Koshy J, Kreps J, et al. Building Linked Ins Real-time Activity Data Pipeline. *Data Engineering*, 2012, 35 ( 2 ): 33~45.
- [ 80 ] Isard M, Budiu M, Yu Yuan, et al. Dryad: Distributed data-parallel programs from sequential building blocks. Proc of EuroSys 2007, New York: ACM, 2007: 59~72.
- [ 81 ] Gu Y, Grossman R L. Sector and sphere: The design and implementation of a high-performance data cloud. *Philosophical Transaction of the Royal society A: Physical, Mathematical and Engineering Sciences*, 2009, 367( 1897 ): 2429~2445.
- [ 82 ] Zhou Ao-Ying, Jin Che-Qing, Wang Guo-Ren, Li Jian-Zhong. A survey on the management of uncertain data. *Chinese Journal of Computers*, 2009, 32 ( 1 ): 1~16.
- [ 83 ] Guo Zhi-Mao, Zhao Ao-Ying. Research on data quality and data cleaning: A survey. *Journal of Software*, 2002, 13 ( 11 ): 2076~2082.
- [ 84 ] 张炜, 李建, 中刘禹. 一种基于概率模型的预测性时空区域查询处理. 软件学报. 2007, 18 ( 2 ): 279~290.
- [ 85 ] 刘大有. 时空数据挖掘研究进展. 计算机研究与发展. 2013, 50 ( 2 ): 225~239.
- [ 86 ] Roddick JF, Spiliopoulou M, Lister D, et al. Higher order mining. *ACM SIGKDD Explorations Newsletter*, 2008, 10 ( 1 ): 5~17.
- [ 87 ] Fu TC. A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 2011, 24 ( 1 ): 164~181.
- [ 88 ] Mennis HJ, Han J. Geographic Data Mining and Knowledge Discovery—An introduction. *Computers, Environment and Urban Systems*, 2009, 33 ( 6 ): 403~408.
- [ 89 ] Shekhar S, Zhang P, Huang Y. *Data Mining and Knowledge Discovery Handbook*. Berlin: Springer, 2010: 837~854.
- [ 90 ] Hsu KC, Li ST. Clustering spatial-temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources*, 2009, 33 ( 2 ): 190~200.
- [ 91 ] Rosswong J, Ghose K. Efficiently detecong clusters of mobile objects in the presence of dense noise. Proc of the 2010 ACM symp on Applied Computing. New York: ACM, 2010: 1095~1102.
- [ 92 ] Schultz REO, Centeno TM, Selleron G, et al. A soft computing-based approach to spatio-temporal prediction. *Int Journal of Approximate Reasoning*, 2009, 50 ( 1 ): 3~20.
- [ 93 ] Dietterich TG, Domingos P, Getoor L, et al. Structured machine learning: The next ten years. *Machine Learning*, 2008, 73 ( 1 ): 3~23.
- [ 94 ] Heckmann D, Brandherm B, Schmitz M, et al. GUMO: The General User Model Ontology. Proc. of International Conference on User Modeling. Edinburgh, Scotland: 2005, 428~432.
- [ 95 ] 刘飞, 罗永龙, 郭良敏, 等. 面向个性化云服务基于用户类型和隐私保护的信任模型. 计算机应用, 2014, 34 ( 4 ): 994~998.
- [ 96 ] Amin Javari, Mahdi Jalili. A probabilistic model to resolve diversity-accuracy challenge of recommendation systems. *Knowledge and Information Systems*. 2015, 44 ( 3 ): 609~627.
- [ 97 ] 张启宇, 郭承坤, 宋瑶, 等. 基于内容过滤的个性化农业信息推荐模型研究. 湖北农业科学, 2015, 54 ( 16 ): 4052~4056.
- [ 98 ] 杨凤萍, 张大斌. 差分进化神经网络集成的用户偏好模型构建. 微型机与应用, 2016, 35 ( 8 ): 44~47.
- [ 99 ] Moon-Hee Park, Jin-Hyuk Hong, Sung-Bae Cho. Location-Based Recommendation System Using Bayesian User's Preference Model in Mobile Devices. International Conference on Ubiquitous Intelligence and Computing. UIC 2007: Ubiquitous Intelligence and Computing, 2016: 1130~1139.
- [ 100 ] 高升, 任思婷, 郭军. 基于潜在因子模型的跨领域信息推荐算法. 电信科学, 2015, 7 ( 188 ): 1~6.
- [ 101 ] 崔艳萍, 阎知知, 王小巍, 等. 互联网信息资源用户获取优化推送仿真研究. 计算机仿真, 2017, 34 ( 7 ): 273~280.

- [ 102 ] 余本功, 张卫春, 汪柳. 基于用户偏好的产品知识推送算法研究. 科学技术与工程, 2017, 17(1): 265~271.
- [ 103 ] 张驰, 张晓, 王登位, 等. 基于组件库的生鲜农产品冷链物流云服务系统设计与实现. 农业工程学报, 2016, 32(12): 273~279.
- [ 104 ] 孙琦宗, 钟嘉信, 陈川, 等. 一种基于用户偏好的定制优先级判定方法. 机电工程 2017, 34(11): 1354~1358.

## Advances in adaptable technologies of agricultural cloud services

Chen Tian'en<sup>1</sup>, Liu Junping<sup>2\*</sup>, Wang Dengwei<sup>1</sup>, Shi Xiaohui<sup>1</sup>

(1. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097;

2. Information Center of Beijing Municipal Commission of Rural Affairs, Beijing 510275)

**Abstract:** [ **Purpose** ] As a new delivery mode of agricultural information technology applications, agricultural cloud service has obvious advantages in reducing technical and cost threshold, and better responding to personalized service requirements. But the particularity of agricultural applications, regional differences and dynamics based service requirements challenges the applicability of general cloud computing technologies. It is the most important factor of agricultural cloud service application to breakthrough adaptable technologies. [ **Method** ] In current paper, around key links of agricultural cloud service, such as data management, knowledge discovery and service, the advance of adaptable technologies in agriculture cloud service, which includes agriculture data storage method and model, agriculture data mining and knowledge discovery, agriculture knowledge service composition, on-demand services of agriculture application, were systematically summarized and presented, and a reference model of adaptive agriculture cloud service was put forward. [ **Result/conclusion** ] However, there are still some problems deserve more attentions in future, the research of adaptive agriculture cloud service revealed its great potential, methods presented in this paper can provide reference for the development of large-scale agricultural and other industry applications based on cloud services.

**Keywords:** adaptive agriculture cloud service; integrated storage of agriculture spatial-temporal data; agriculture knowledge discovery; knowledge services selection and optimum combination; on-demand services