

# 云平台中国土地资源信息存储检测仿真研究

栾绍鹏<sup>1,2,3</sup>, 侯西勇<sup>1,3</sup>

(1. 中国科学院烟台海岸带研究所, 山东 烟台 264003; 2. 中国科学院大学, 北京 100049;  
3. 中国科学院海岸带环境过程与生态修复重点实验室, 山东 烟台 264003)

**摘要:** 云平台中对土地资源信息进行检测, 能够有效的对土地资源信息进行合理的存储。对土地资源信息的存储, 需要将多源信息包分散的存储到每个节点中, 并形成存储信息包, 完成对土地资源信息的存储。传统方法先将信息存储的空间划分为不同的行列矩阵, 给出不同资源的稀疏列, 但忽略了存储信息包的建立, 导致信息存储精度偏低。提出基于云平台的土地资源信息分布式存储方法。上述方法先融合云计算理论得到传感器网络模型和能量模型, 给出信息分布式存储的问题方程, 组建以最大化信息资源存储质量为目的的分布式存储机制, 依据一定的接收概率将系统中多源信息包分散的存储到云计算平台下的每个节点中, 在每个节点上形成一个存储信息包, 以此为依据完成对土地资源信息分布式存储。实验结果证明, 所提方法存储精度较高, 可以有效地完成对土地资源信息分布式存储。

**关键词:** 云平台; 土地资源; 信息存储

**中图分类号:** TP391      **文献标识码:** B

## Based on the Cloud Platform of Land and Resources Data Distributed Storage Simulation

LUAN Shao-peng<sup>1,2,3</sup>, HOU Xi-yong<sup>1,3</sup>

(1. Yantai Institute of Coastal Zone Research, CAS, Yantai Shandong 264003, China;  
2. University of Chinese Academy of Sciences, Beijing 100049, China;  
3. Key Laboratory of Coastal Environmental Processes and Ecological Remediation, CAS, Yantai Shandong 264003, China)

**ABSTRACT:** A method for distributed information storage of land resource is proposed based on cloud platform. Firstly, the model of sensor network and energy model integrated with theory of cloud computing is obtained, and the equation of distributed information storage is provided. Then the mechanism of distributed storage aiming at maximizing quality of information resource storage is built, and the information packet with multisource in system is dispersedly stored into each node under cloud computing platform. Moreover, a storage information packet at each node is generated. Finally, the distributed information storage is completed. Experimental results show that the method has high storage precision. It can complete the distributed information storage effectively.

**KEYWORDS:** Cloud platform; Land resources; Information storage

### 1 引言

在国内, 土地资源信息化逐渐成为国内经济信息化的核心组成部分, 也是全方位提升国土资源管理、保护、规划和高效利用的重要手段<sup>[1-3]</sup>。在“十二五”期间对国土资源信息化明确指出, 在大信息环境的基础下, 需要充分依靠国家信息基础设施建设成果, 围绕信息科技进步和新一代国土资源管理, 借助国土资源创新平台, 开展资源信息存储、云计算、智慧地球等关键技术 in 国土资源信息化中的应用研究<sup>[4-6]</sup>。

但是在现代信息技术的飞速发展过程中, 不同级别国土资源管理部门的资源存储服务设备越来越多, 在添加硬件存储设备的同时, 也无形中增加了国土资源信息管理系统基础架构的复杂繁琐程度和信息存储成本。在这种情况下如何有效地对国土资源信息进行存储, 成为了信息管理领域亟待解决的主要问题。而资源信息存储方法可以依据一定的接收概率将网络中多源信息包分散的存储到云计算平台下的每个节点中, 在每个节点上形成一个存储信息包, 完成对国土资源信息的分布式存储, 也是解决上述问题的根本途径, 引起了很多专家和学者的重视, 同时也出现了很多好的方法<sup>[7]</sup>。

文献[8]提出一种基于哈希的国土资源信息分布式存储

收稿日期: 2017-02-13    修回日期: 2017-03-03

方法。该方法先融合于哈希理论给出信息存储的分布式存储的存放空间 组建资源信息的散列矩阵,完成对信息的分布式存储。该方法信息存储稳定性较强,但是存在信息存储时间开销较大的问题。文献[9]重点讨论了一种基于行列混合存储的国土资源信息分布式存储方法。该方法先将信息存储的空间划分为不同的行列矩阵,给出不同信息资源的稀疏列,利用该列完成对国土资源信息分布式存储。该方法信息存储时间复杂度较低,但是采用当前方法进行资源信息存储时,难以保障信息的存储质量,存在信息存储误差大的问题。文献[10]重点提出一种基于一致性分布树的国土资源信息分布式存储方法。该方法先给出资源信息的一致性哈希分布,对每一层的信息进行定位,给出资源信息的备份策略,完成对国土资源信息分布式存储。该方法信息存储精度较高,但是存在资源信息存储过程较为繁琐、耗时长的问題。

针对上述问题,提出一种基于云平台的国土资源信息分布式存储方法。实验结果证明,所提方法存储精度较高,可以有效地完成对国土资源信息的分布式存储。

## 2 云平台中国土资源信息存储检测原理

在对国土资源信息分布式存储过程中,对系统不同层面的存储节点进行划分,给出存储系统的能量模型,对资源结构进行描述,表述资源信息特征的映射过程,针对不同的资源信息选择相应的存储节点,建立国土资源信息分布式存储机制,完成对国土资源信息分布式存储。具体步骤如下详述:

假设,由  $v_{pop}$  代表存储节点,  $E_{SF}$  代表感知信息,  $d_{dgg}$  代表其感知信息的大小,  $s_{tyuu}$  代表一个节点在单位时间内采集感知信息的次数, Sink 节点单位时间内收到的查询次数为  $Y_{ase}$ , 则利用式(1)对系统不同层面的存储节点进行划分

$$f_{sdff} = \frac{Y_{ase} \times v_{pop}}{E_{SF}} \times \frac{d_{dgg} \mp \{s_{tyuu} \times Y_{ase}\}}{g_{ij}} \quad (1)$$

式中  $g_{ij}$  代表反馈信息和存储信息之间的比率。

假设,由  $w_{sd}$  代表发送信息时的能量消耗模型,  $se_{fj}$  代表接收信息时的能量消耗模型,  $A_{djj}$  代表一次存储单位大小感知信息时的能量消耗,则利用式(2)对资源结构进行描述

$$c_{dpp} = \frac{A_{djj} \mp se_{fj}}{w_{sd}} \times e_{dpp} \quad (2)$$

式中  $e_{dpp}$  代表查询反馈信息的数量。

假设,由  $v_{wedd}$  代表以节点  $v$  为根的子树的能量消耗,  $pefrjj$  代表节点  $v$  的双亲节点,则利用式(3)给出资源信息特征的映射过程

$$Q_{asu} = \frac{v_{wedd} \mp pefrjj}{r_{sdff}} \times b_{dgg} \quad (3)$$

式中  $r_{sdff}$  代表存储节点和各个节点树的匹配概率,  $b_{dgg}$  代表存储节点的相似性。

假设,由  $s_{sdjj}$  代表存储节点的类型集合,  $w_{askk}$  代表存储节点的类型数量,  $e_{dggk}$  代表资源信息的大小,  $fg_{vij}$  代表源信息包

的接收频率,则利用式(4)针对不同的资源信息选择相应的存储节点

$$p_{jh} = \frac{fg_{vij} \times s_{sdjj}}{w_{askk}} \mp \{e_{dggk} * r_{djj}\} \times Q_{asu} \quad (4)$$

式中  $r_{djj}$  代表存储资源访问吞吐率。

假设,由  $e_{igtv}$  代表节点发送感知信息的大小,  $s_{akk}$  代表存储节点的能耗设定阈值,  $e_{pp}$  代表随机选取存储节点的集合,则利用式(5)建立国土资源信息分布式存储机制

$$W_{sdpp} = \frac{e_{pp} \times e_{igtv}}{s_{akk}} \mp f_{sdff} \quad (5)$$

假设,由  $\chi_{fg}$  代表信息资源类型的稀疏性,则利用式(6)完成对国土资源信息分布式存储

$$w_{er} = \frac{\chi_{fg} \times c_{dpp}}{Q_{asu}} \mp p_{jh} \times W_{sdpp} \quad (6)$$

综上所述可以说明,云平台中国土资源信息存储检测原理,利用该原理可以完成对国土资源信息存储检测。

## 3 云平台中国土资源信息分布式存储检测方法

### 3.1 资源分布式存储的可行性问题的转换

在对国土资源信息分布式存储过程中,将国土资源信息分布式存储系统定义为一个无相连通图,得到当前系统的可用存储空间,设计系统传感器和能量模型,组建以最大化信息资源存储质量为目标分布式存储机制。具体的步骤如下详述:

假设,由  $G(V, E, N)$  代表分布式存储系统的无相连通图,  $V$  代表  $N$  个均匀部署的传感器节点,  $E$  代表  $m$  条边的集合,将  $DG_i$  定义为源节点,将  $V_s$  定义为  $k$  个发生器,而且每个信息发生器感知到的资源信息都是具有一定单位尺寸的资源信息项序列,  $S_{sdpp}$  代表  $DG_i$  需要重新分配的资源信息项,则利用式(7)计算当前系统的可用存储空间

$$SEY_{sdkk} = \frac{S_{sdpp} \times G(V, E, N) \otimes m}{DG_i \times V_s \pm e_{sf} \times k} \times p_{sffg} \quad (7)$$

式中  $e_{sf}$  代表传感器节点均匀分布的资源信息项,  $p_{sffg}$  代表信息项的总跳数。

假设,由  $E_{sepp}$  代表接收能量,  $p_{dggk}$  代表系统需要重分配的信息项类型总数,  $r_{dgp}$  代表任意传感器节点  $r$  的初始能量,则利用式(8)组建系统的能量模型

$$T_{sdj} = \frac{p_{dggk}}{E_{sepp}} \times \frac{r \times \{r_{dgp}\}}{P_{sdpp}} \quad (8)$$

式中  $P_{sdpp}$  代表系统的总能耗。

假设,由  $F_{skkj}$  代表发送资源和接收资源每比特的能耗,  $\varepsilon_{vbjj}$  代表发送放大器每比特的能耗,则利用式(9)组建系统的传感器模型

$$l_{dgl} = \frac{\varepsilon_{vbjj} \mp F_{skkj}}{r_{dggj}} \times E_{sepp} \times T_{sdj} \quad (9)$$

假设,由  $e_{ds}$  代表分配信息项  $j$  的过程中节点  $i$  消耗的能量,  $p_{djp}$  代表分配函数,  $p_{jg}$  代表分配路径集,则利用式(10)组

建以最大化信息资源质量为目标分布式存储机制

$$E_{swe} = \frac{p_{fg} \mp e_{ds}}{p_{djp}} \mp u_{ple} \times l_{dgl} \quad (10)$$

式中  $\mu_{ple}$  代表节点存储空间变化量。

假设,由  $R_{min}$  代表包括节点  $i$  在内的  $i$  的所有邻居节点的最小剩余存储空间,  $R_{TH}$  代表设置的剩余存储空间阈值,则利用式(11)计算出  $R_{TH}$

$$S_{xij} = \frac{R_{TH} \times R_{min}}{E_{swe}} \times F_{sdij} \mp SEY_{sdkk} \quad (11)$$

式中  $F_{sdij}$  代表剩余存储空间。

假设,由  $p_{xij}$  代表存储空间的利用率,  $\mu_{dfg}$  代表溢出信息项的总个数,将资源存储时间定义为全部存储溢出资源项的目的节点的剩余能量之和,利用式(12)表述

$$e_{dfg} = a_{dfg} \times p_{xij} \times S_{xij} \times w_{sf} \quad (12)$$

式中  $w_{sf}$  代表传感器节点存储空间的信息项的限制。

综上所述可以说明,在对国土资源信息分布式存储过程中,将国土资源信息分布式存储系统定义为一个无相连接图,得到当前系统的可用存储空间,设计系统传感器和能量模型,组建以最大化信息资源存储质量为目标分布式存储机制,为实现对国土资源信息分布式优化存储奠定了坚实的基础。

### 3.2 基于接收概率的国土资源信息分布式存储

在对国土资源信息分布式存储过程中,以3.1节定义的资源存储时间  $e_{dfg}$  为依据,得到各个节点存储的初值,利用定向随机游走理论将接收到的源信息包异或到最新的存储信息包中,得到存储信息包的生成矩阵,完成对国土资源信息分布式存储。具体的步骤如下:

假设,由  $E_{pooi}$  代表节点通信半径,  $C_i(o)$  代表节点  $o$  的邻居集合,  $c_j(i)$  代表源信息包  $X_j$  已访问节点  $i$  的当前次数,以3.1节定义的资源存储时间  $e_{dfg}$  为依据,得到各节点存储的初值,利用式(13)表述

$$S_{sepp} = \frac{E_{pooi} \otimes C_i(o)}{c_j(i)} \times \frac{p_{ghj} \mp e_{dfg}}{X_j} \quad (13)$$

假设,由  $s_{ghj}$  代表传递每个源信息包的定向随机游走的步数,  $p_{spp}$  代表每个节点接收一个新到达的源信息包的概率,  $v_{dfgg}$  代表当前访问节点  $v$  的邻居节点集合,利用定向随机游走理论将接收到的源信息包异或到最新的存储信息包中,用式(14)表述

$$R_{spp} = \frac{v \times s_{ghj}}{e_{spp}} \mp v_{dfgg} \times \{t_{chll} \mp u_{cipp}\} \quad (14)$$

式中  $\mu_{cipp}$  代表每个到达节点  $u$  的源信息包,  $t_{chll}$  代表节点  $u$  的传递概率,  $R_{spp}$  代表源信息包的线性组合。

假设,由  $\mu_{ser}$  代表每个节点接收一个新到达的源信息包的概率,  $X_{k_p}^m$  代表  $k_p$  个未知的源信息包,  $g_i$  代表一个独立的且取值在二元域  $F_2$  上的随机向量,  $S_{(k+\varepsilon)}^m$  代表  $k+\varepsilon$  个存储信息包生成的行向量,则利用式(15)得到存储信息包的生成矩阵

$$p_{asec} = \frac{(k+\varepsilon) \times S_{(k+\varepsilon)}^m}{g_i \times F_2 \mp X_{k_p}^m} \mp k_p^m \pm \{\mu_{ser}\} \quad (15)$$

假设,将  $k'$  个源信息包  $\Phi_{fgh}$  分布式存储到了网络的  $n$  个节点中,而且不同的节点存储信息包都是包含很多个源信息包的异或和,则利用式(16)完成对国土资源信息分布式存储

$$Y_{sdpp} = \frac{\Phi_{fgh} \times p_{asec}}{R_{spp} \times k'} \times \frac{n \mp S_{sepp}}{R_{spp}} \quad (16)$$

## 4 实验及仿真证明

为了证明提出的基于云平台的国土资源信息分布式存储方法有效性需要进行一次实验。在 MATLAB 环境下搭建国土资源信息分布式存储实验平台。在  $120 \times 120$  的正方形区域随机均匀地分布了多个点模拟资源分布式存储系统中各个存储节点,其中,资源节点的比例设定为 40%,系统中每个节点的通信半径  $R$  满足条件  $R^2 = R \times 100^2$ 。

实验分为两组,第一组实验测试所提基于云平台的国土资源信息分布式存储系统的负载均衡性  $l_{is}$ ,存储延迟  $S$  将该部分实验定义为对该系统的源信息包存储有效性的验证。第二组实验测试所提基于云平台的国土资源信息分布式存储系统的存储空间利用率  $p_{xij}$  和资源存储代价  $P_{sfdpp}$ 。将该组实验定义为对该系统的源信息包存储效率的验证。

### 4.1 源信息包存储有效性验证

测试所提基于云平台的国土资源信息分布式存储系统的负载均衡性  $l_{is}$ ,存储延迟  $S$ ,测试结果如图1和图2所述。

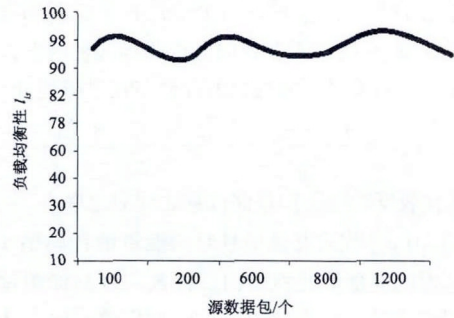


图1 分布式存储系统的负载均衡性

从图1和图2的仿真结果可以看出,随着源信息包的数量不断增加,利用所提基于云平台的分布式存储方法进行国土资源信息存储始终具有较好的均衡负载性和较低的存储延迟,可以充分的保证所提方法进行国土资源信息分布式存储的有效性。

### 4.2 源信息包存储效率的验证

测试所提基于云平台的国土资源信息分布式存储系统的存储空间利用率  $p_{xij}$  和存储代价  $P_{sfdpp}$ ,测试结果如图3和图4。

从图3和图4的仿真结果可以看出,随着源信息包的数

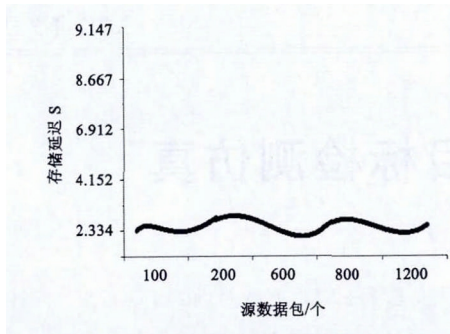


图2 分布式存储系统的存储延迟

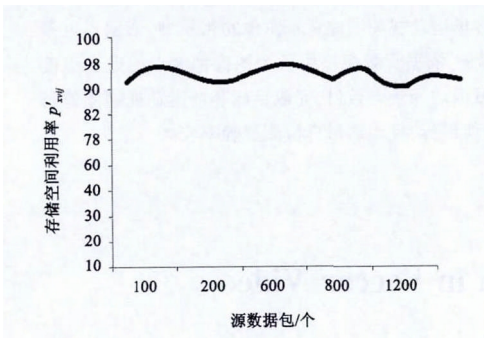


图3 分布式存储系统的存储空间利用率

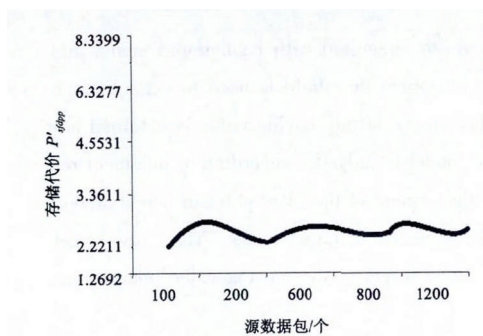


图4 分布式存储系统的存储代价

量不断增加,利用所提基于云平台的分布式存储方法进行国土资源信息存储的存储空间利用率较大,且进行资源信息存储的代价较低,极大的提升了所提方法进行国土资源信息分布式存储的效率。

实验结果证明,所提方法存储精度较高,可以有效地完成对国土资源信息分布式存储。

## 5 结束语

针对采用当前方法进行资源信息存储时,难以保障信息的存储质量,存在信息存储误差大的问题。提出一种基于云平台的国土资源信息分布式存储方法。实验结果证明,所提方法存储精度较高,有效地完成对国土资源信息分布式存储。

## 参考文献:

- [1] 刘春阳,等. 基于 XML 树-表结构的多元巡检数据存储方法[J]. 计算机仿真, 2014, 31(6): 133-136.
- [2] 蔡婷,等. 一种面向 SaaS 应用的差异数据存储方法[J]. 计算机应用与软件, 2016, 33(4): 39-43.
- [3] 刘高军,段然. 基于 MongoDB 的 CNONIX 数据存储方法研究[J]. 北方工业大学学报, 2016, 28(3): 40-48.
- [4] 付河. R5F104BC 单片机 Flash 中数据存储方法研究[J]. 单片机与嵌入式系统应用, 2016, 16(5): 31-34.
- [5] 杜燕忍,葛华勇,王龙. 基于 Lizard 框架 Ctrip Wireless H5 数据存储方式的优化[J]. 微型机与应用, 2016, 35(11): 14-17.
- [6] 郑义成,莫钦华,王海鸿. 基于 Elasticsearch 的海量 AIS 数据存储方法[J]. 指挥信息系统与技术, 2016, 7(3): 76-81.
- [7] 石常兴,周浩. 一种基于 EVC 的数据存储管理算法实现[J]. 电子设计工程, 2014, 22(20): 187-189.
- [8] 徐光伟,等. 远程不可信存储的数据完整性巡检算法研究[J]. 小型微型计算机系统, 2016, 37(9): 1930-1936.
- [9] 杨俊杰,廖卓凡,冯超超. 大数据存储架构和算法研究综述[J]. 计算机应用, 2016, 36(9): 2465-2471.
- [10] 向春枝,范颖. 云计算环境中分布式数据存储关键技术研究[J]. 现代电子技术, 2016, 39(3): 63-67.

## 【作者简介】



栾绍鹏(1978-)男(汉族),山东省招远市人,硕士研究生,高级工程师,研究方向:国土资源信息化、云GIS等;

侯西勇(1975-)男(汉族),山东省泰安市人,博士研究生,研究员,研究方向:海岸带土地利用/覆被变化(LUCC)、海岸线变化遥感监测、海岸带生态脆弱性评价、海岸带综合管理。