关于空间数据存储与区块链技术结合的探讨

姓　　名：张清昱

指导教师：闫志刚

班　　级：地理信息系统 17-1 班

学　　院：环境与测绘学院

2020 年 10 月

摘 要

当前，区块链技术已经深入到多个领域，通过介绍近些年来刚刚兴起的区块链技术及其分类，并延展到 GIS 方面的应用，探讨二者相结合的契合性，旨在展望空间区块链技术的可行性与未来前景。

关键词：区块链；GIS；数据存储

Abstract

At present, blockchain technology has penetrated into many fields, and by introducing the newly emerged blockchain technology and its classification in recent years, and extending to the application of GIS, and exploring the compatibility of the combination of the two, aiming to look forward to the feasibility and future prospects of spatial blockchain technology.

Key words: Blockchain; GIS; Data storage

目 录

第一章 绪论

继 2015 年国外开始大幅度在区块链技术上进行投资后，世界很多重大组织，包括高盛、花旗银行、英国央行、美国央行等机构纷纷在区块链技术上进行投资，而 2016 年可以说是中国区块链的元年，此时区块链技术开始受到极大的重视，首先是 1 月的时候，人民银行宣布使用数字货币，然后在 30 日以后，许多中国的组织单位开始投资区块链。区块链并不是一项新技术，而是一个新的技术组合，包括其中的 P2P 技术、密码学、共识机制等都是已经十年以上的老技术了，但是这些巧妙的结合在一起后，形成了如今的区块链技术。

大数据、云计算快速发展浪潮下，互联网信息爆炸式增长，人们对云存储的需求不断增加，百度网盘、腾讯云等云存储产品也随之应运而生。传统的中心化的云存储方式相对于本地存储虽有异地存取方便、空间灵活等的优点，但也存在用户隐私容易泄露、用户数据容易被篡改、用户对数据控制权减弱等缺点。[1]而参考到区块链的特性，即围绕着高安全、防篡改、可追溯三大特性，区块链脱离虚拟货币交易而主要发挥其数据存储的功能便被提了出来。而在各种数据上添加地理与时空信息，便可大致绘制出一张空间区块链的蓝图。

第二章　区块链简介

2.1 起源与发展

2.1.1 起源

2016 年 1 月 20 日，中国人民银行官方网站上发表了一篇题为《中国人民银行数字货币研讨会在京召开》的新闻，成为了推动区块链技术在国内迅速升温的导火线，相关首次给予了比特币底层技术——区块链技术高度评价。

所谓区块链，就是基于区块链技术形成的公共数据库，也有很多人称之为公共账本。区块链技术最早出现在用于比特币交换的概念中，而比特币的概念最早出现在中本聪（Satoshi Nakamoto）2008 年一篇题目为《比特币：一种点对点的电子现金系统》（Nakamoto， 2008）中[2]，是对他全新创造的货币体系的介绍，最初，比特币只是作为密码学的一种全新尝试而在一小群人之间传播，并没有人愿意用实在的货币进行交换，经过几年的发展，2011 年起，随着一系列交易市场的建立，比特币的价格也随之攀升。

2.1.2 发展

在区块链的发展过程中，最早出现的区块链 1.0，即区块链技术的基本版本，能够实现可编程货币，是与转账、汇款和数字化支付相关的密码学货币应用。通过这一层次的应用，区块链技术首先起到搅动金融市场的作用。

区块链 2.0 是指智能合约，智能合约与货币相结合，对金融领域提供了更加广泛的应用场景。区块链相对于金融场景有强大的天生优势.简单来说，如果银行进行跨国的转账，可能需要打通各种环境，货币兑换，转账操作，跨行问题等等。而区块链实现的点对点的操作，避免了第三方的介入，直接实现点对点的转账.提高了工作效率。

区块链 2.0 的代表是“以太坊”。以太坊是一个平台，它提供了各种模块让用户用以搭建应用。平台之上的应用，其实也就是合约，是以太坊技术的核心。以太坊提供了一个强大的合约编程环境，通过合约的开发，以太坊实现了各种商业与非商业环境下的复杂逻辑。以太坊的核心与比特币系统本身是没有本质的区别的。而以太坊的本质是智能合约的全面实现，支持了合约编成。让区块链技术不

仅仅是发币，而提供了更多的商业、非商业的应用场景。

区块链 3.0 是指区块链在金融行业之外的各行业的应用场景。能够满足更加复杂的商业逻辑。区块链 3.0 被称为互联网技术之后的新一代技术创新，足以推动更大的产业改革。区块链 3.0 涉及生活的方方面面。所以区块链 3.0 将更加的具有实用性，赋能各行业，包括地理信息行业。此时不再依赖于第三方或某机构获取信任与建立信用，能够通过实现信任的方式提高整体系统的工作效率。

也可以说，区块链 1.0 是区块链技术的萌芽，区块链 2.0 是区块链在金融，智能合约方向的技术落地。而区块链 3.0 是为了解决各行各业的互信问题与数据传递安全性的技术落地与实现。

2.2 分类

2.2.1 公共链

公共链对外开放，用户不用注册就能匿名参与，无需授权即可防伪网络和区块链。节点可选择自由出入网络。公共链上的区块可以被任何人查看，任何人也可以在公共链上发送交易，还可以随时参与网络上形成共识的过程，即决定哪个区块可以加入区块链并记录当前的网络状态。公共链完全去中心化， 如比特币等， 只要接入此链，都可读写、交易、参与共识等， 该链公开透明，但不受国家监管， 缺乏维权举证的法律保障。联盟链仅限联盟成员参与，共识过程由预选节点共同决定， 可有效提升联盟内跨组织工作效率，如 R3 银行联盟，其共享同一账本，可做到接近实时校验与自动结算。[3]

2.2.2 私有链

私有链则仅在私有组织使用，区块链上的读写权限、参与记账权限按私有组织规则来制定。私有链的应用场景一般是企业内部的应用，如数据库管理、审计等。也有一些比较特殊的组织情况，比如在政府行业的一些应用：政府的预算和执行，或者政府的行业统计数据。

这个一般来说由政府登记，但公众有权力监督。私有链的价值主要是提供安全、可追溯、不可篡改、自动执行的运算平台，可以同时防范来自内部和外部对数据的安全攻击，这个在传统的系统是很难做到的。央行发行数字货币可能就是一种私有链。私有链不会同步以太坊数据，不会消耗以太币，减少了成本和存储

的要求[4]。

2.2.3 联盟链

联盟链只选择部分节点来确定协商一致意见，其具有以下特性：首先，对于联盟链的一致性确定，由一组选定的节点负责对区块进行验证；其次，记录的读取权限可以是公开的，也可以是受限的；最后，半分布式并不是网络中所有节点都能加入联盟链的共识过程，共识效率高。[5]

2.3 共识机制

2.3.1 简介

在区块链上，每个人都会有一份记录链上所有交易的账本，链上产生一笔新的交易时，每个人接收到这个信息的时间是不一样的，想要攻击的人就有可能在这时发布一些错误的信息，这时就需要一个人把所有人接收到的信息进行验证，最后公布最正确的信息。它就像一个国家的法律，维系着区块链世界的正常运转。加密货币都是去中心化的，去中心化的基础就是 P2P 节点众多，那么如何吸引用户加入网络成为节点，有那些激励机制，同时，开发的重点是让多个节点维护一个数据库，那么如何决定哪个节点写入、何时写入，一旦写入，又怎么保证不被其他的节点更改（不可逆）。共识机制为解决上述问题而存在。

2.3.2 共识机制比较

（1）PoW(Proof-of-Work)

PoW 最著名的应用就是比特币以及以太坊和一些其他基于 PoW 协议的货币。其最大的缺点是会面临 51%以上的攻击，但由于区块链是不可变的，但是可以通过拒绝服务，来使得无法正常运行，并且需要消耗大量的算力大量的电力，并不适合树莓派来搭建，PoW 协议更多的适用于公链[6]。

（2）PoS(Proof-of-Stake)

PoS 要求证明人提供一定数量加密货币的所有权，此机制会根据每个节点拥有代币的比例和时间，依据算法等比例降低节点的挖矿难度，从而加快了寻找随机数的速度[6]。这种机制相较于 PoW 来说或许是一种好的选择，但是它本质上仍然需要网络中的系统但进行挖矿运算，不能为我们提供我们足够所需的控制和安全级别。

（3）PoA(Proof-of-Authority)

PoA 是相对 PoW 和 PoS 来说算是一个新概念，在这个概念中，您拥有许多预先批准的授权节点（即矿工）。想要添加任何新节点，必须由当前已有的矿工投票，这使您可以完全控制哪一些节点可以做为矿工。以太坊的 PoA 协议称为 Clique，它适用于私链，但不适用于公链。[7]

（4）Kafka

Kafka 是一个分布式的、高吞吐的、基于发布/订阅的消息系统。利用 kafka 技术可以在廉价 PC/Server 上搭建起大规模的消息系统。Kafka 具有消息持久化、高吞吐、分布式、实时、低耦合、多客户端支持、数据可靠等诸多特点，适合在线和离线的消息处理。[8]

2.4 关键技术

2.4.1 哈希链条

区块链采用密码学技术提高安全性，哈希函数是最基础也是最核心的技术。哈希函数通过输入任意大小的字符串得到固定大小的支出，哈希函数具有碰撞阻力、隐秘性和谜题友好三个特征。碰撞阻力指不同输入必须得到不同输出结果，可以对信息加密。创作者用哈希函数给作品加密后，作品处于锁定状态，任何人在修改账本后，都会改变账本的哈希值。

隐秘性可以为内容保密，当某人提出一个创意或草拟一份合同时，既希望对信息保密，也希望能够在日后证明创意或合同的真实性和有效性。哈希函数可以进行加密，隐秘性保证任何人都无法通过逆向运算的方式揭露创意或合同。当事人需要展现创意或合同时，可以计算哈希值来证明创意或合同的真实性和有效性，哈希函数上的时间戳还可以确定创意或合同的提交时间，任何人都无法更改时间戳。谜题友好指与输出值相对应的一部分输入值是随机的，没有计算输入值的捷径。在区块链技术中，通过计算谜题获得比特币作为奖励，计算的过程也称为挖矿，矿工通过挖矿计算哈希值，挖矿基本靠暴力计算。

哈希函数是一个数据块，哈希指针指向哈希函数的位置，由哈希指针组成的链表就是区块链。区块链中的哈希指针指向区块位置，并可以检验数据是否有过修改的记录。[9]

2.4.2 分布式账本

分布式对等网络是仅包含具有等效控制和操作能力节点的计算机网络。区块链信息系统底层拓扑结构是分布式对等网络，各个节点通过对等网络进行数据通信以支撑上层功能。网络的五层模型可以分为物理层、数据链路层、网络层、传输层、应用层。区块链系统小世界模型[10]的 P2P 网络是以 IP 协议、TCP 协议为基础存在于应用层面上的逻辑覆盖网络，特点主要有非中心化、扩展性强及负载均衡[10]，这些特点为区块链系统高效稳定运行提供了强有力的保证。

区块链系统维护一个在启动时可以连接的对等节点列表[12]，在系统新节点接入已有网络时，首先节点会通过“种子”得到对等节点 IP 列表。节点间通常采用 TCP 协议与相邻节点建立连接，建立连接时也会有认证“握手”的通信过程用来确定 P2P 协议版本、软件版本、节点 IP、区块高度等。为了能够被更多节点发现，新节点会将带有自身 IP 地址的信息发送给相邻节点，并要求其返回其已知的对等节点的 IP 地址列表。[13]

第三章　空间区块链

3.1 概念的提出

3.1.1 从理论上看

区块链技术发展到如今的 3.0 时代，带来了更多适用于商用以及非商业的使用平台，不再仅局限于挖矿与虚拟货币交易，对于地理信息而言，其公开性、安全性、可回溯性更是极其贴合新时代地理信息系统的进一步发展态势。将概念提取出来，仅将区块链作为一个存放数据的数据库，便能很容易让人联想到 GIS 上空间数据存储技术的革新。区块链具有严格的结构定义，每个块[14]由区块头和区块体构成，不会产生数据会丢失的忧虑。传统空间数据库的三大特点，数据量大，数据内包括空间数据、属性数据以及二者之间的关系，数据应用广泛，皆能被区块链继承。

对于数据量而言，根据曾志明教授在 2020GIS 软件技术大会上提供的数据，在 CPU Intel i5 2.70GHz，内存 16G 的硬件系统上，搭建 3 个 Peer 节点，1 个 Ordered 的节点，使用 Kafka 机制，TPS（Transaction Per Second）每秒交易量能够达到 920，而 Raft 则为 1032TPS，按照每秒能够支持 1000 笔交易进行计算，每天即为 2800 万笔，而拿全国的不动产登记来说，每天大约只有 30 万笔，是完全满足且有百倍预留的。事实上普通空间信息并不推荐使用区块链进行存储，因为与传统数据库读写相比，上时间上的数据读取并不理想，但其实对于每时间来说差的并不多。

3.1.2 从应用上看

数据应用方面，区块链技术早已可以深入到数据交易、身份认证、新能源、车联网、物流供应链、医疗、房地产、文化娱乐等领域。由于其有并不是新技术，而是新的技术的集合的这一特性，相信如有新的行业的出现，也可与区块链技术迅速进行结合，而事实上无论什么行业都离不开两种数据，一是时空数据，因为每笔交易都会在一定的时间下进行，而是地理数据，因为每笔交易都一定会伴随着地点的发生，这就给传统地理信息行业提供了拥有非常高普适性的数据接口。

对于区块链本身来说，如类似区块链本身已经进入的行业，食品、药品的溯

源监管，重要物品物流位置管理等等，将空间位置信息加入后，会有更好的位置服务。同时从空间资产管理的角度出发，也可用在国土空间规划与用途管制，不动产登记与交易，自然资源资产与权益管理等等，也可发挥区块链本身防篡改、透明性高的优势。建立分布式区块链系统日志网络，日志作为交易上传到区块链上[15]。

3.2 技术展望

3.2.1 与其他技术的结合

（1）5G

5G 是指第五代移动通信，是下一阶段的移动通信技术标准，5G 的出现能够加快数据的传输速度，使得基于 M/S 架构的 GIS 系统也能高速与联盟链契合，达到空间数据的高效传输，实现快速互联。5G 与测绘地理信息之间是相辅相成、相互促进的关系，5G 技术的研发使得测绘地理信息工作逐渐向智能化、现代化发展，而测绘地理信息技术在各行各业中的渗透又给 5G 技术带来更广阔的发展空间。将 5G 技术与测绘地理信息相结合，促使地理信息得到更深入和高效的利用，最终技术融合的成效会作用于社会公众的服务上，带来更多便捷、高质的生活方式。

（2）“云边端”

云计算可以带来就算是单一简单设备也能有的高算力优势，一方面来说，能够提供存储量小也能运操大量空间数据的美好设想，另一方面来说，云计算可以让计算更加分散，只要空间位置确定，该位置上的资源、类型、所有者信息均可以认为是依附于空间位置的抽象属性，该特征为数据的及时上链提供了理论依据。

与传统空间信息的“横向成网，纵向多级”特征相对应，云 GIS 呈现出“合而分之，分而合之”的特征。在对海量数据进行存储或者对大规模任务进行划分时，需要对资源数据进行划分，而这种划分过程必然保存了原有的空间信息和非空间的属性信息。

用户和数据均是云 GIS 的组成部分，两者均具有社会属性。前者表现为用户权限，后者表现为空间数据的社会属性（如河流名称、行政区划等）。社会科学视角下的云 GIS 平台呈现集中性特征。云 GIS 平台通常由“实力雄厚”的机构来创建基础设施和平台，其子机构和用户使用云平台提供的服务，可以达到资源

的有效控制和合理使用的双重目的。同时这样的特征与区块链的透明性结合能达到更多群众可以进行监管的效果。

3.2.2 展望

我们无需关心数据是怎么上的链，链上如何索引，数据怎么提取，在如今万物高速互联的时代背景下，我们只用关心数据的采集与提取出来后怎么用，中间这个黑匣子完全可以被封装成一个可用的工具，现在空间区块链技术的概念刚刚提出，能看到的只有少部分如超图公司对其进行的初步探索，如在一个城市宗地管理系统中，空间区块链保证宗地变更可信、可追溯，如图 1 所示。

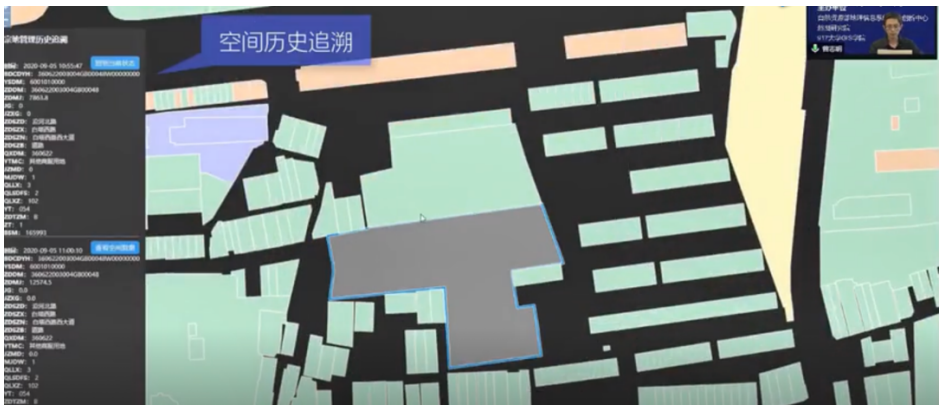

图 1 空间区块链管理下的宗地信息系统

GIS 的发展伴随着计算机技术的发展而发展，在新技术日新月异的今天，太多新的互联网技术事实上都是可以被 GIS 借鉴使用的，区块链只是其中的一个，而在这样的技术背景下，更是需要技术与技术之间互相扶持，产出更加人性化、面向未来的技术产业，空间区块链必能给传统的空间数据存储插上翅膀，走向更好的技术未来。

参考文献

[1] 赖力潜,刘学东,钟伟豪.基于区块链技术的轻量化分布式云存储系统设计[J].大众标准化,2020(18):42-44.

[2] 杨晓晨,张明.比特币:运行原理、典型特征与前景展望[J].金融评论,2014,6(01):38-53+124.

[3] 曾子明,万品玉.基于主权区块链网络的公共安全大数据资源管理体系研究[J].情报理论与实践,2019,42(08):110-115+77.

[4] 穆程刚,丁涛,董江彬,宁可儿,董晓博,贺元康,王永庆,陈天恩,刘健,Mohammad Shahidehpour.基于私有区块链的去中心化点对点多能源交易系统研制[J/OL].中国电机工程学报:1-12[2020-10-17].https://doi.org/10.13334/j.0258-8013.pcsee.200392.

[5] 王昊,吴天昊,朱孔林,张琳.交叉口场景下基于区块链技术的匿名车辆身份认证方案[J].网络与信息安全学报,2020,6(05):27-35.

[6] 刘童桐.区块链共识机制研究与分析[J].信息通信技术与政策,2018(07):26-33.

[7] 牛岩松,宗峰.一种利用 Raspberry 搭建区块链私有链全节点的方法研究[J].现代计算机,2020(19):105-108.

[8]胡聪,刘翠玲,张翠翠,徐敏. 基于 Kafka 分布式发布订阅消息系统的电网全业务统一数据中心-数据实时接入方法设计研究[C]. 中国电力科学研究院有限公司、国网电投（北京）科技中心、《计算机工程与应用》杂志社.第三届智能电网会议论文集——智能用电.中国电力科学研究院有限公司、国网电投（北京）科技中心、《计算机工程与应用》杂志社:国网电投(北京)科技中心,2019:93-96.

[9] 王亮.区块链在版权保护中的应用与实践[J].中国报业,2020(18):10-12.

[10] Zhu Y, Gan G H, Deng D, et al. Security architecture and key technologies of blockchain. J Inform Security Res, 2016, 2(12): 1090

[11] Antonopoulos A M. Mastering Bitcoin: Unlocking Digital Cryptocurrencies. California, O'Reilly Media, Inc, 2014

[12] Ben Mariem S, Casas P, Donnet B. Vivisecting blockchain P2P networks: Unveiling the bitcoin IP network ACM CoNEXT Student Workshop. Crete, 2018

[13] 朱岩,张艺,王迪,秦博涵,郭倩,冯荣权,赵章界.网络安全等级保护下的区块链评估方法 [J/OL]. 工程科学学报 :1-20[2020-10-17].https://doi.org/10.13374/j.issn2095-9389.2019.12.17.00

[14] Bowden R, Keeler H P, Krzesinski A E, et al. Block arrivals in the Bitcoin blockchain [J/OL].arXiv preprint(2018-01-23)[2019-12-17]. https://arxiv.org/pdf/1801.07447.pdf

[14] Aniello L, Baldoni R, Gaetani E, et al. A prototype evaluation of a tamper-resistant high performance blockchain-based transaction log for a distributed database // 2017 13th European Dependable Computing Conference (EDCC).Geneva, 2017: 151