基于区块链的农业物联网可信溯源体系

陈锦雯[1], 罗得寸[1], 唐呈俊[1], 唐晨钧[1], 丁 勇[1,2]

([1]广西密码学与信息安全重点实验室, 桂林电子科技大学, 广西桂林 541004;
[2]鹏程实验室网络空间安全研究中心, 广东深圳 518055)

摘要 随着农业信息化的快速发展, 农业物联网逐渐成为智慧农业必要的基础设施。未来智慧农业依赖大量的物联网传感器, 为了解决农业物联网中数据孤立、不可靠、容易被篡改、难以追踪追责的问题, 本文提出了一种基于区块链的农业物联网可信溯源体系。考虑农业物联网数据量大、种类繁多的特点, 本方案基于 Hyperledger fabric 智能合约, 采用分布式 raft 共识协议, 结合物联网设备服务, 精准权限控制, 实现物联网数据存储溯源的可靠可信以及冗余数据过滤与基于属性的访问控制, 大幅减少了冗余数据, 提高了安全性。在本文提供的方案中, 农场管理员授权员工用户绑定所属物联网设备, 物联网设备定期上传的文本数据发送到智能合约进行聚类筛选, 剔除冗余数据后上链存证。利用 IPFS 分布式存储图像、视频, 通过区块链存证其内容哈希、智能合约管理其生命周期。采用链上存储、链下计算的模式对大量本地数据进行分析、统计、可视化展示; 进行链上校验, 针对文本数据采用直接上链存储的方法, 对于图片文件采用本地存储、链上校验的方式确保图片信息不被篡改。仿真和分析结果表明, 本方案在稳定性、安全性上优于传统溯源方案, 性能上满足千级数量的物联网设备并发上链的需求。

关键词 区块链; 农业物联网; 可信溯源; IPFS

中图法分类号 TP311.1 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2022.03.09

The Trusted Traceability System of Agricultural Internet of Things Based on Blockchain

CHEN Jinwen[1], LUO Decun[1], TANG Chengjun[1], TANG Chenjun[1] , DING Yong[1,2]

1. Guangxi Key laboratory of Cryptography and Information Security, Guilin University of Electronic Technology, Guilin, Guangxi 541004.China
2. Pengcheng Laboratory Cyberspace Security Research Center, Shenzhen, Guangdong 518055,China

Abstract With the rapid development of agricultural informatization, Agricultural Internet of Things has gradually become a necessary infrastructure for smart agriculture. In the future, smart agriculture depends on a large number of IoT sensors. In order to solve the problems of isolated, unreliable, easily tampered and difficult to track responsibility in the Agricultural Internet of Things, this paper proposes a trusted traceability system of Agricultural Internet of Things based on blockchain. Considering the characteristics of large amount and wide variety of Agricultural Internet of Things data, this scheme is based on the hyperledger fabric intelligent contract, adopts the distributed raft consensus protocol, combined with Internet of Things equipment services and accurate authority control, it realizes the reliable traceability of Internet of Things data storage, redundant data filtering and attribute based access control, greatly reduces redundant data and improves security. In the scheme provided in this paper, the farm administrator authorizes the employee user to bind the Internet of things device. The text data uploaded by the Internet of things device regularly is sent to the smart contract for clustering and screening. After removing the redundant data, it is linked and stored. IPFS is used to store images and videos, the content hash and smart contract are stored in the blockchain to manage its life cycle. A large number of local data are analyzed, counted and visually displayed in the mode of on chain storage and off chain calculation. For the text data, the method of direct uplink storage is adopted, and the picture files are stored locally and verified in the chain to ensure that the picture information is not tampered with. The simulation and analysis results show that this scheme is superior to the traditional traceability in terms of stability and security, and its performance meets the requirements of thousands of IoT devices on the chain concurrently.

Key words blockchain; agricultural Internet of Things; trusted traceability; IPFS

通讯作者: 丁勇, 博士, 教授, Email**:** stone_dingy@126.com

本课题得到国家重点研发计划项目(No. 2020YFB1006003, No. 2020YFB1006004)、国家自然科学基金项目(No. 61772150, No. 61862012, No. 61962012)、广东省重点领域研发计划项目(No. 2020B0101090002)、广西自然科学基金项目(No. 2018GXNSFDA281054, No. 2018GXNSFAA 281232, No. 2019GXNSFFA245015, No. 2019GXNSFGA245004, No. AD19245048)、鹏城实验室网络空间安全研究中心网络仿真项目(No. PCL2018KP004)的资助。

收稿日期: 2020-11-05; 修改日期: 2021-03-15; 定稿日期: 2022-01-12

1 引言

目前, 随着农业信息化的发展以及人们生活水平的提升, 农产品质量安全问题受到国家的广泛关注。由于农业生产规模大, 监管力度难以到位, 农产品质量问题接连不断, 引发了消费者信任危机。农产品从种植到成为商品出售的过程中, 经历了较大的时间和空间跨度, 在各个环节容易出现各种品质问题。主要表现在以下几个方面: 一是重复人工检测, 由于农产品种植、加工、物流、销售等环节的信息互相孤立, 导致当前的溯源体系很难保证农产品的质量, 农产品质量重复检测的问题较为普遍; 二是信息孤岛, 售卖方对于每一个处理环节都需要对农产品的信息做出检测, 但部分信息并没有出现在其他环节而是封锁在本环节内; 三是监管困难, 在对农产品质量进行监管时, 如对某个环节安全和质量检测不信任, 就会导致信息的重复检测, 没有实现信息的共享, 进而增添了监督的成本。同时, 农产品安全事故一旦出现, 相关信息容易被责任人恶意篡改, 造成事故追责难、取证难的问题。

农业的生产方式目前正在发生着巨大的改变, 从过去的人工参与到现在的机械化、信息化的智慧农业。农业物联网是智慧农业与信息农业的重要基础, 未来农业物联网是大趋所势。农业物联网是通过大量的传感器节点构成监控网络, 以帮助农民及时、准确地发现问题, 这将农业从以人力为中心转向以信息为中心。在进行农产品溯源的过程中, 数据的采集是农产品检测的重要环节。但农产品的生产基地存在位置较偏僻、交通状况不够便利等不利情况, 使得获取数据变得困难。因此, 现代化农业生产开始利用物联网技术, 在农产品生产区域设立若干个数据采集点, 采集点通过物联网设备自动化采集农产品周围环境指标、图像、视频等数据。数据采集以采集点为单位, 通过互联网协议定期上传至服务器, 由服务器保存和备份数据。

做好农产品质量安全工作的关键是要注重农产品源头数据的有效监控, 同时, 各个生产环节必须有效结合。农产品安全问题本质上是市场买卖双方的信任问题, 生产信息的不对称、消息不透明容易造成消费者对产品的恐慌。而如今企业构建的传统溯源系统存在中心化、容易被内部篡改的弊端, 因此这本身是一种“自证”行为, 中心化的数据存储方式容易被删改, 溯源体系自然无信任可言。

区块链作为一种新兴的互联网技术, 底层采用分布式数据存储、共识机制、智能合约、加密算法等核心技术。区块链上的数据具有不可篡改、公开透明、永久保存的特性, 数据被共识确认之后分布在各个节点中, 可以实现安全可靠的数据备份。区块链可以作为分布式网络系统的框架来推动互联网业务可信化。物联网(IoT)的基本特点是数据量大, 传统物联网溯源系统的数据存储在中心化的数据库中, 使得中心数据库压力过大, 安全性、可追溯性都无法得到保证。依托区块链技术可以将农业生产相关的物联网数据和农事管理数据以分布式、多方维护的形式进行持久化保存, 建立农业大数据信任存储模型, 为未来智慧农业的可持续发展创建奠定了基础。具有区块链基础设施的农业系统是可信而不可改变的记录管理系统, 这种特性使得各种资源从源头记录、追踪、使用, 可确保记录和服务的完整性。因此, 本文针对农产品生长阶段各维度的物联网数据, 采用区块链技术设计一种去中心化、安全可信、可维护、去冗余的区块链农业物联网溯源系统。基于可信物联网数据存储与溯源框架, 能够提升消费者的信任, 促进农产品产业的健康发展。

本文主要的研究工作如下。

1) 研究可信溯源体系。设计实现农业物联网可信溯源系统, 提供物联网数据的可信存证, 基于知名 Hyperledger fabric 联盟区块链框架设计, 采用分布式 raft 共识协议, 结合物联网设备服务, 精准权限控制, 实现物联网数据存储溯源的可靠可信。

2) 实现物联网信息多维度存储, 支持多种物联网协议、传感器数据格式, 实现物联网新数据多维度存证与溯源。

3) 海量多样化数据过滤设计。解决海量媒体数据, 存在大量重复信息, 区块链难以承载等问题。

4) 链上信息校验与可视化。本文采用“链上存储、链下计算”的模式, 对大量本地数据进行分析、统计、可视化展示; 进行链上校验, 针对文本数据采用直接上链存储的方法, 对于图片文件采用本地存储、链上校验的方式确保图片信息不被篡改。

2 相关工作

农业是全球经济的重要支柱, 近年来随着食品安全问题频发, 农药、化肥、生长调节剂等农用化学品滥用, 工业废弃物排放增加等原因造成有毒农作物泛滥, 严重危害人们身体健康, 构建农产品溯源系统可以追溯来源, 有效解决质量安全问题[1], 目前在对农产品供应链系统的研究上, 大多通过射频识别[2]、二维码技术[3]、物联网[4]、无线传感网络[5]等建立溯源系统对供应链上的信息进行追踪溯源。农产品供应链具有生命周期长、环节复杂、信息多源

异构等特点[6], 而以上研究并不能完全解决供应链信息保护中存在的问题, 主要有以下两点:一是供应链上各个节点间信息不对称, 信任成本较高, 影响整体效率; 二是传统溯源体系应用中心化数据库, 供应链各个节点数据由企业自主管理, 存在信息丢失和易被篡改的问题。因此, 对于农产品产业, 保证生产的各个环节可追溯、透明化尤为重要。

区块链作为一种新的分布式解决方案, 引起了学术界和工业界的广泛关注。它具有离散化和高度自治的优点, 为用户数据提供了透明的、不可变地存储[7]。区块链技术包括执行智能合约、认证数字身份、增加供应链透明度等[8]。分布式账本技术(DLT)是一种许可链[9]。智能合约是在区块链上运行的一段可信执行代码。智能合约可以根据预先确定的逻辑可靠地在区块链上执行[10]。只有拥有授权证书的用户才能加入许可链。因此, 采用 DLT 实现分布式数据库的访问控制机制可以实现安全灵活的访问控制。区块链技术具有不可篡改、分布式、去中心化、可追溯、高可用等特点[11], 为解决目前传统农产品追溯体系所存在的问题提供可能[12], 近几年国内外学者对区块链技术在供应链追溯领域的应用都进行了探索研究[13], 这些研究推动了区块链技术在农产品溯源领域的应用进展[14], 解决了传统溯源数据保护过程中数据备份繁琐[15]等问题。近几年区块链在农业领域也得到广泛应用, 对于数据的采集, 本文利用传感器获取数据, 传感器在农业方面有大量应用场景[16], 为了实现兰花的高精度栽培, Jiang J A 等人[17]提出了一种在兰花大棚中基于无线传感器网络(WSN)的算法和一种新的动态收敛 cast tree 算法(DCTA)。Qiangyi Y 等人[18]开发了一款基于智能手机的应用程序, 作为人类感知工具来观察土地状况, 如作物覆盖和生长。就物联网和电子农业中使用的众多技术而言, 实现物联网体系结构的全面安全生态系统的主要挑战可以通过区块链技术来实现[19]。基于物联网自动采集的溯源研究在文章[20]中有所体现, 实现了 farm-to-fork 的可追溯性。本文考虑了物联网的特性提出一种基于农业物联网的自动采集与区块链溯源系统, 在智能合约采用 K-means 算法大幅降低冗余物联网数据, 保证农产品可追溯、可视化, 提升消费者的信任, 促进农产品产业健康发展。

3 预备知识

3.1 区块链

(1) 共识机制

在分布式系统中, 不同的主机通过异步通信的方式组成网络集群。为了保证每个主机达成一致的状态共识, 需要在主机之间进行状态复制。异步系统中, 需要在默认不可靠的异步网络中定义容错协议, 以确保各主机达成安全可靠的状态共识。所以, 利用区块链构造基于互联网的去中心化账本, 需要实现不同账本节点上的账本数据的一致性和正确性。

共识机制是指以去中心化的方式就网络的状态达成统一的过程, 也被称为共识算法, 有助于验证和验证信息被添加到分类账本, 确保事务记录在区块链上是真实的, 共识机制负责安全地更新分布式网络中的数据状态。去中心化后, 保证整个系统能有效运行, 各个节点诚实记账, 在没有所谓的中心的情况下, 互相不信任的个体之间就交易的合法性达成共识的共识机制。一个有效的共识机制可以保证各个节点之间按照既定的规则共同维护账本, 本质上是区块链系统中实现不同节点之间建立信任的算法。

Raft 是用于实现分散式系统各个节点的状态达成强一致性的共识机制, 主要用于管理日志复制的一致性。Raft 的核心思想是: 如果在分散式系统中多个数据库的初始状态一致, 只要之后进行的操作顺序一致, 就能保证之后的执行结果一致。

(2) 智能合约

本质上来说, 智能合约是一段可执行程序, 它以计算机指令的方式实现了传统合约的自动化处理。智能合约是双方在区块链资产上交易时, 触发执行的一段代码。智能合约程序不只是一个可以自动执行的计算机程序, 它本身就是一个系统参与者, 对接收到的信息进行回应, 可以接收和储存价值, 也可以向外发送信息和价值。简单来说, 提前规定好合约的内容, 当在满足触发合约条件的时候, 程序就会自动执约内容。

3.2 IPFS 分布式存储

IPFS 是一个分布式的点到点超媒体协议, 能够让互联网传输速度更快, 更加安全, 并且更加开放。IPFS 作为一种分布式存储协议, 将文件分块存储和传输, 不同节点可以共享文件块, 通过分布式路由表协议将节点、文件块索引成一张内容命名的存储网络。

IPFS 具有成本低、可自动去重, 节省存储空间、成本等优点。IPFS 将每个文件分成小块, 每个块通过 merkle DAG 构造成一个文件哈希值, 只有拥有这个哈希值才能索引并获得文件块。IPFS 网络在全国各地都有服务器, 部分服务器失效、篡改之后, 仍然可以通过其他服务器的备份文件块来还原, 因此可

以实现永久保存。

3.3 基于 bloom filter 算法的数据过滤改进算法

为了防止传感器采集的数据冗余, 我们提出一种基于 bloom filter 算法的数据过滤改进算法, 由于传感器较多, 导致尽管在很小的时间范围内容有大量数据传出, 且重复数据较多。bloom filter 算法可以有效实现数据去重, 但其存在误识别率, 有一定可能性的误判, 因此我们利用基于 bloom filter 算法进行相应改进实现数据过滤。

在物联网数据上链过程中, 多条传感器数据会连续传入智能合约, 每条数据都带有时间戳。由于环境指标变化具有稳定性、连续性, 因此相邻一段时间内的数据会发生重复。

首先, 由于数据量大且一定时间范围内数据重复率较高, 因此我们使每个数据携带时间戳, 使得我们可以根据时间进行数据过滤。使用哈希函数将每个元素映射到一个二进制向量的三个位上, 将集合中每个元素对应的三个位记录成 1。当需要判断一个新的元素 w 在不在集合中时, 可以先计算出 w 的三个位, 然后只要发现其中存在任何一个位为 0, 则可以确定 w 不在此集合中。如果该元素在集合中, 则不将其放入集合中。

其次, 由于 bloom filter 算法存在误识别, 数据中仍可能出现重复数据, 下面我们要进行重复数据的删除, 采用 K-means 算法将数据进行离散化, 大幅降低冗余物联网数据, 提高了计算效率, 便于数据处理及分析。K 均值聚类算法(k-means clustering algorithm)是一种迭代求解的聚类分析算法。首先将数据分为 K 组, 随机选取 K 个对象作为初始的聚类中心, 然后计算每个对象与各个聚类中心之间的距离, 把每个对象分配给距离它最近的聚类中心。聚类中心以及分配给它们的对象就代表一个聚类。每分配完一个样本, 聚类的聚类中心会根据现有的对象重新计算。这个过程将不断重复直到满足终止条件。

最后我们得到在一定时间区间内无重复的数据, 为了节省存储空间, 将数据进行压缩, 把出现过的字符串映射到记号上, 这样就可能用较短的编码来表示长的字符串, 实现压缩。压缩后再进行数据的存储, 将其上链。

使用上述算法去冗余首先需要等待数据达到一定量的时候再进行过滤、聚类, 然后在每一类中随机抽取若干个对象进行存储, 并将冗余的数据丢弃, 经过压缩后把过滤后的数据再上链或链下存储, 实现数据的过滤, 以降低区块链存储压力,

4 农业物联网可信溯源体系

4.1 系统模型

基于农业物联网的自动采集与区块链溯源系统的目的是实时监测农产品生长中物联网设备产生的环境数据, 对大量数据进行过滤处理并实现分布式存储, 通过区块链可信体系实现数据存证, 该溯源体系模型如图 1 所示。

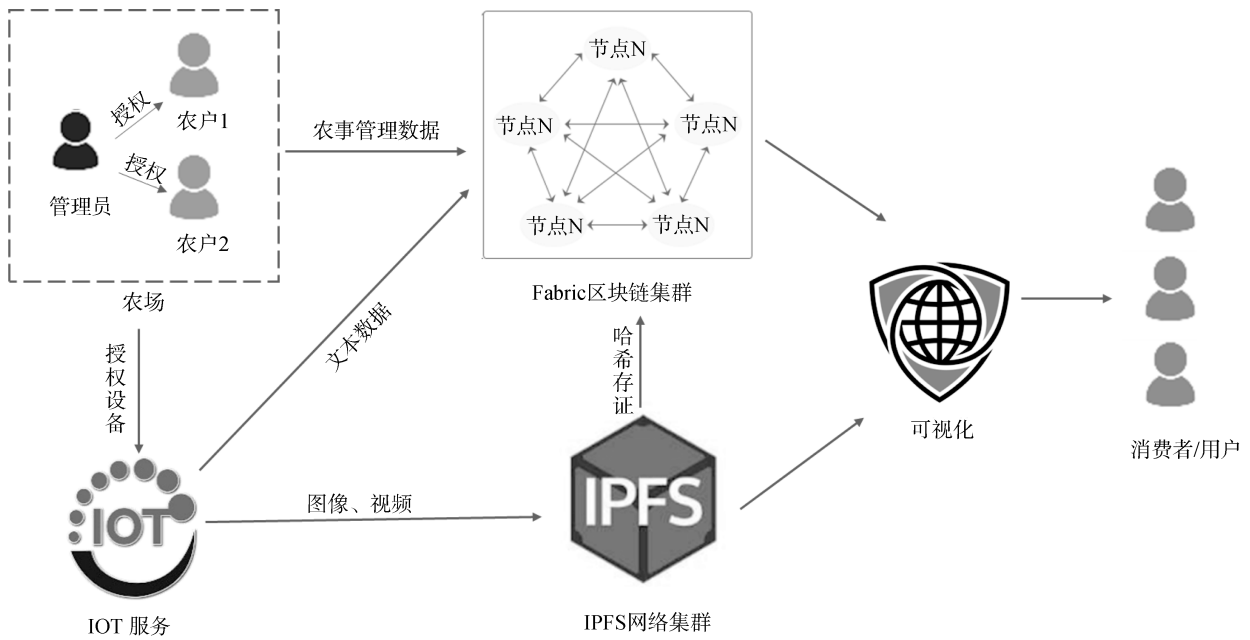


图 1 农业物联网可信溯源体系模型

Figure 1 Agricultural Internet of Things trusted traceability system

农场。一个农场是一个农业生产组织, 组织中的农场管理员可以对各个农户进行授权, 农户获得授权之后可以将农事记录上传至溯源系统, 同时可以添加该农场的物联网设备至系统中。

IOT 服务。物联网(The Internet of Things, IOT)通过信息传感器、感应器等各种装置, 实时采集农作物周围的温度、湿度、土壤水分、位置等各种农作物生长相关信息, 通过各种网络协议接入, 实现对农作物的智能化感知、识别和管理。传感器设备得到农场用户的授权后, 可以将采集到的数据定时上传至区块链。

IPFS 网络集群。物联网采集到的数据中包含的图像和视频通过 IPFS 分布式存储, IPFS 节点将这类大文件进行分块后广播到网络中的其他节点, 改变了传统中心化的存储方式; IPFS 对于文件中相同的文件块只保存一次, 实现自动去重, 大大节省了存储空间、成本; IPFS 中的每个文件都会有对应的哈希值, 只有通过这个哈希值才能索引到文件块, 从而拼成一个完整文件。IPFS 将大文件数据存储后, 通过哈希存证到区块链网络, 实现分布式内容存储与区块链的绑定。

Fabric 区块链集群。物联网采集到的文本数据经过过滤后存入 Fabric 区块链集群。Fabric 区块链网络是分布式点对点架构, 其中的数据保证了去中心化、开放性、不可篡改性、可追溯性。

消费者、用户。消费者通过可视化平台查询指定物联网设备在 IPFS 网络集群和 Fabric 区块链集群中所存储的物联网信息, 根据时间区间进行溯源。同时, 消费者可以根据需要对查询到的数据进行哈希校验, 以确保该信息的真实性和完整性。

本文的系统主要有农场管理员、农户、用户三大角色, 实现用户管理、信息上链、信息溯源功能, 并赋予不同角色不同权限。

农户进行注册并将传感器、摄像头等设备进行授权获取农产品的相应信息, 然后进行定时采集, 最终格式化上链存储, 完成物联网的数据上传。用户登录信息溯源界面, 系统需要根据所输入要查询的时间段和采集点的数据获取到相应的传感器曲线、生长图片和各种农事记录, 并将获取到的信息链上校验, 保证获取信息真实可靠。

4.2 系统实现

基于农业物联网的自动采集与区块链溯源系统按照业务内容进行模块化分类, 主要分为感知层模块、存储层模块、合约层模块、功能层模块以及可视化模块。系统总体框架及其各模块详情如下图 2 所示。

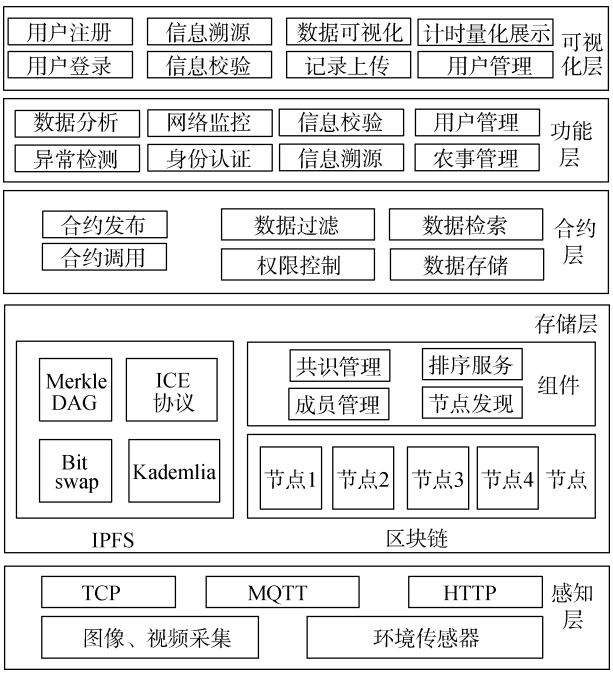


图 2 农业物联网可信溯源系统框架

Figure 2 Agricultural Internet of Things trusted traceability system framework

4.2.1 感知层

感知层基于物联网技术, 由大量的农业物联网环境采集传感器构成, 分别记录了农作物光照、温度、压强、湿度等数据指标。此外, 使用摄像头采集图像、视频的物联网设备也在逐渐增多。物联网设备身份信息包含设备 ID 以及所属农场用户的身份 ID。农场用户将其身份 ID 和设备 ID 通过一对多的关联映射方式, 将所属设备的绑定关系存储到区块链上。设备数据上链时需要进行登记获取临时授权令牌才能将数据上链, 临时令牌超时失效后需要向权限控制合约再次申请。物联网设备采集到的农作物数据通过 TCP 协议、MQTT 协议、HTTP 协议、RMTP 协议等进行传输。采集程序开启协议连接服务来接收传感器数据并推送至数据采集合约, 采集合约按照 K-means 算法的规则进将数据存入区块链, 不符合规则的冗余数据直接丢弃。此外, 采集到的图像、视频数据将推送至 IPFS 存储网络, 同时还通过摄像头进行截图来获取图像数据。图片及视频数据存储在 IPFS 分布式私有存储网络中, IPFS 生成数据的唯一内容索引哈希后再将哈希值保存到区块链存证。感知层物联网设备的数据认证、采集、存储流程如图 3 所示。

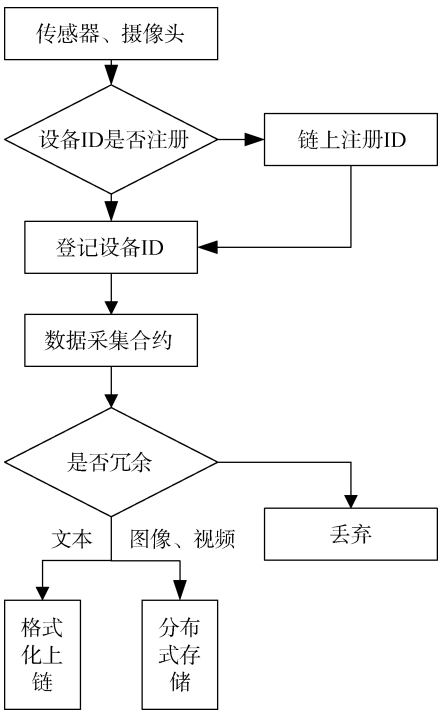


图 3　感知层设备认证以及数据采集、存储流程

Figure 3　Sensing layer equipment authentication and data acquisition and storage process

4.2.2　存储层

存储层由超级账本(Hyperledger Fabric)和 IPFS 构成。超级账本是一种推进可信计算的开源联盟区块链项目，具有点对点网络的特性，分布式账本技术是共享、透明和去中心化的。超级账本基于模块化的框架设计理念，提供了可插拔的共识机制、成员身份管理服务、节点数据库、背书策略以及验证策略等，非常灵活易于扩展。其设计的核心元素是智能合约(链上代码)、数字资产、记录储存库、中心化共识网络、加密安全。超级账本目前提供基于 Raft 的共识机制，通过排序节点分离出区块链网络的共识服务，在性能方面优于大多数联盟链。IPFS 通过梅克尔 DAG、分布式路由、ICE 网络穿透等技术，将数据内容映射成唯一 CID(Content Identity)，通过文件分块技术去除重复内容，在保证了数据完整性的同时大幅降低数据冗余度。

底层超级账本采用 Hyperledger fabric v1.4.6 架设，构建四节点主集群网络。区块链网络拓扑如下图 4 所示，集群可以通过设立 CA 节点的方式扩展其他节点，农场用户向 CA 节点注册 X.509 证书构建记账节点。节点的类型包括 Orderer 排序节点、Peer 记账节点，网络主集群通过 docker 容器部署。IPFS 集群依附在主集群节点，对于图像视频数据，主集群的 IPFS 节点将文件块通过负载均衡机制分发至其他农场节点中，每个节点通过 GC(垃圾回收)机制定时清理过期数据。

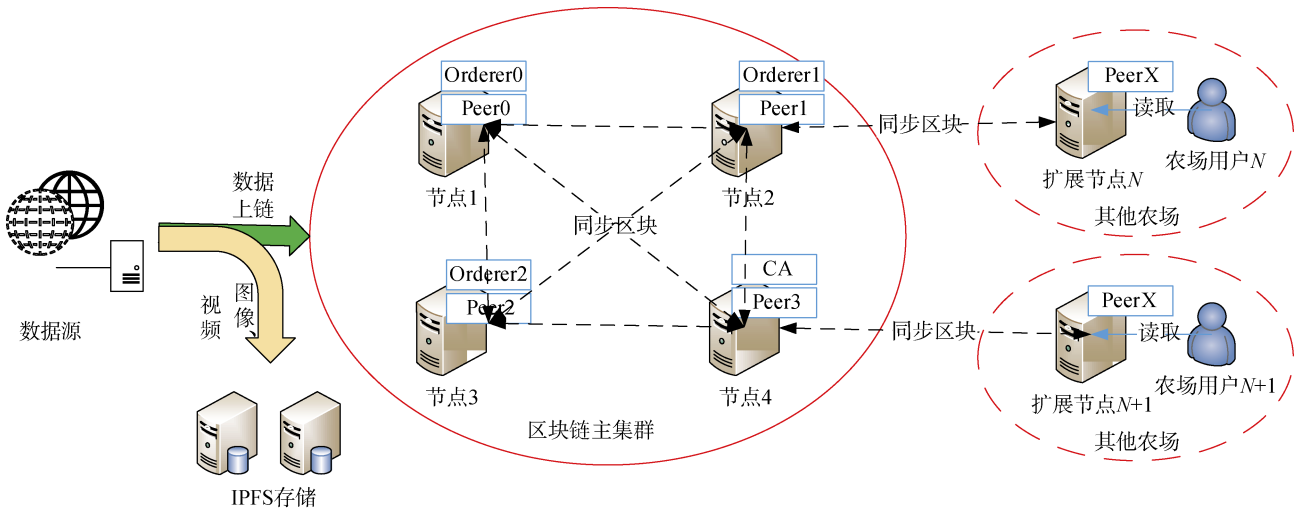


图 4　节点网络拓扑图

Figure 4　Node network topology

由于传感器数据种类较多，数据采集模块将多维数据扁平化后再将数据存储至区块链。上链的数据包括传感器数据信息、图片信息、农事管理信息、用户 ID 信息、设备 ID 信息。传感器数据条、图片以其哈希值作为索引，将其存到区块链上；农事数据条以用户名为索引存到链上；设备 ID 信息以及用户 ID 信息取哈希之后存到链上，便于核验登录。对于农场用户，添加管理员/普通用户两种角色，实现用户在区块链上的身份注册、登陆、吊销。

4.2.3　合约层

智能合约层通过 Go 语言编写实现，主要负责数据存储、数据检索、数据过滤、权限控制等逻辑。

合约由主集群联盟管理方发布, 再获得其他农场节点授权背书的条件下部署到区块链网络中。合约主要包含 4 个算法, 分别是数据存储算法、数据检索算法、数据过滤算法、权限控制算法。

算法 1. 数据存储算法.

定义复合键值对 Mk: <type>~<point>~< time-stamp>:<value>。其中 *name* 为传感器指标名称, *value* 为指标数值, *type* 为传感器类型, *point* 为采集点 ID, *timestamp* 为时间戳。

输入: JSON 格式传感器数据: *data*, 该数据包含 *n* 个环境参数键值对: <name>:<value>;

输出: 链上存储的数据哈希: *tx_id*。

1) 遍历;

2) *raw*←*data* 中的一对<name>:<value>;

3) 生成时间戳 *ts*;

4) 利用 *raw* 和 *ts* 构造复合键值对 Mk;

5) 调用 PutState()接口写区块链获得交易 *id*;

6) 将 *id* 拼接到 *tx_id*;

7) until *data* 遍历完成, 返回 *tx_id*。

算法 2. 数据检索算法.

依据给定的类型、时间区间、采集点 id 进行溯源查询, 并返回链上数据。

输入: 类型 *t*, 开始时间 *st*, 终止时间 *et*, 采集点 *pid*

输出: 溯源结果 *res_list*

1) 构造查询范围:*sk*←*t*~*pid*~*st*, *ek*←*t*~*pid*~*et*;

2) 调用 GetStateByRange()接口获得记录 *res*;

3) 对 *res* 进行分页, 获得 *res_list*;

4) 返回 *res_list*。

算法 3. 数据过滤算法.

在智能合约内, 对于传感器 *s* 在一段时间内发送的数据 *data*, 通过聚类随机选择代表样本存入区块链。

输入: 一段时间内的数据 *data*;

输出: 过滤后的数据 *res*。

1) 使用 *k* 个相互独立的哈希函数, 将集合中的元素映射, 被映射的位置为 *l*;

2) 若元素所映射的位置均为 *l*, 则将该元素删除;

3) 选择 *m* 个点作为初始质心;

4) 将每个点指派到最近的质心, 形成 *m* 个聚类;

5) 重新计算每个聚类的质心;

6) until 聚类不发生变化或达到最大迭代次数;

7) 随机抽取一个聚类中的几个点作为过滤后的数据;

8) 读入字符 *c*, 与 *p* 合成并形成字符串 *p*+*c*;

9) 在字典中查找 *p*+*c*, 若 *p*+*c* 在字典中, *p*=*p*+*c*, 否则将 *p* 的记号输出, 在字典中为 *p*+*c* 建立记号映射, 更新 *p*=*c*;

10) 返回步骤 8 重复, 直至读完所有数据, 实现数据的过滤压缩, 返回 *res*。

算法 4. 权限控制算法.

用户、物联网设备调用 trace 智能合约存储数据需要经过 user 合约进行权限控制。

输入: 证书信息 *cert*;

输出: 授权状态 *status*。

1) 检查 *cert* 的属性条目 *role* 是否为采集点;

2) 查询 *cert.id* 所属的用户 *user*;

3) 检查 *user* 是否正常;

4) 返回 *status*←*ok*。

4.2.4 功能层

功能层基于Hyperledger fabric sdk连接智能合约实现数据存储、检索, 同时连接 IPFS 分布式网关将图片、视频进行存储, 并且为可视化层服务提供功能接口。

采集的图片、视频在 IPFS 分布式存储集群中存储, 功能层连接 IPFS 的锚节点, 通过锚节点的 DHT 服务发现功能定时获取网络集群中存在的节点。在获取的节点列表中选择网络连接质量最好的 *N* 个节点作为文件块分发节点, 依次将文件块分发到这些节点中。功能层为可视化层提供接口服务, 包括提供快速的图片访问接口, 该接口提供查询本地数据库图片索引的服务; 客户端查询接口, 该接口可以获取链上哈希值返回图片数据。

功能层还能实现数据分析、异常检测、网络监控等功能。其中数据分析统计模块主要包括以下两个内容: 溯源数据异常检测, 对查询出来的溯源数据进行异常检测, 计算农产品的健康值并存入数据库, 编写健康值查询接口, 向用户展示; 溯源数据统计, 能够将不同类型的溯源数据按照年、月、周统计, 制作图表能够展示指定地区的所有传感器的数据记录数存入数据库, 最后通过溯源数据接口向用户展示。

4.2.5 可视化层

可视化层主要实现区块链网络状态监控、传感器的状态监控, 包括对区块高度、活跃节点数、采集设备列表的可视化展示。此外, 该层还实现用户注册登录、信息溯源校验、数据分析、数据可视化界面的展示。

4.3 安全需求

农产品可信溯源系统主要针对农产品生长过程

中产生的关键数据进行安全可靠的存证及溯源, 主要安全需求包含如下内容:

1) 系统运行安全

系统运行安全包括系统风险管理、审计跟踪、备份与恢复、应急处理, 保证在系统运行过程中出现程序崩溃、节点异常等的情况下, 系统能够具备一定的容错性, 保证数据存证与查询服务等基本操作正常运行。

2) 数据库存储安全

数据库用户访问时应具备密码访问安全性。真实性、不可伪造性尤为重要, 判断数据库系统中的非法行为, 减少数据库系统的安全风险。

3) 操作数据可追溯

对于敏感数据的用户操作需要在系统后台进行日志上链, 确保能够追溯该类数据的历史情况, 保证数据的可追溯性, 并且要进行用户隐私保护, 用户与其所属的物联网设备通过哈希映射的方式存储在区块链上, 因此他人无法获得任何与用户有关的隐私信息。

4) 系统防御安全性

采用分布式系统架构、Raft 共识机制, 能够在分布式节点总数的三分之一都被恶意控制的条件下, 系统数据账本不会被篡改。

5 实验分析

5.1 实验环境

本文对实现的农业物联网可信溯源体系进行了实验分析。实验通过 docker 容器技术在虚拟机内部署了 4 个虚拟节点, 虚拟机操作系统为 ubuntu16.04, 硬件资源配置为 CPU: AMD Ryzen 5 2600 4 核心, 内存: 8GB, 磁盘: 256GB SSD。

5.2 实验方法

本实验针对农产品数据上链、溯源查询以及用户认证三个接口进行性能测试, 从延迟和吞吐率来评估溯源系统的高可用性和稳定性。测试实验室 Jmeter 接口压力测试软件来模拟 2000 名用户同时发送请求, 每个用户重复发送 5、10 次作为本次系统的测试条件。其中, 延迟的平均值表示响应时间, 吞吐率表示每秒的数据交易数, 可以评估该系统的性能。

5.3 实验结果

1) 数据上链

如表 1、图 5、图 6 所示分别表示了数据上链过程的数据报告、延迟和吞吐率。可以看出上链过程响应时间较长, 每秒钟的转移数较低, 但失败的转移数非常低。这是因为当大量用户同时上链时, 要执行共识机制, 系统无法及时处理, 导致请求堆积, 在内存中停留时间较长。

表 1 数据上链汇总报告

Table 1 The summary report of up data to blockchain

Label	样本(个)	平均值(ms)	最小值(ms)	最大值(ms)
HTTP 请求	20000	26367	227	84458
总体	20000	26367	227	84458

标准偏差	异常(%)	吞吐量(s^{-1})	接收 KB/s	发送 KB/s	平均字节数
18443.77	4.74	50.6	12.77	26.78	258.3
18443.77	4.74	50.6	12.77	26.78	258.3

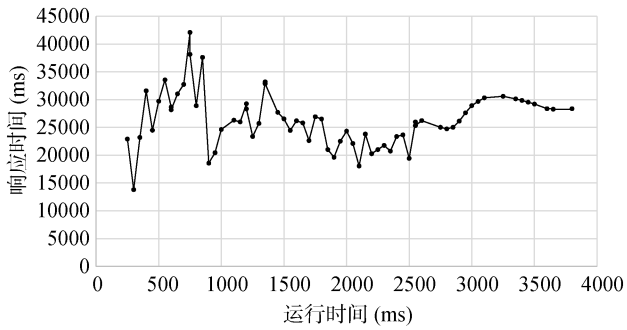


图 5 数据上链的延迟

Figure 5 The delay of up data to blockchain

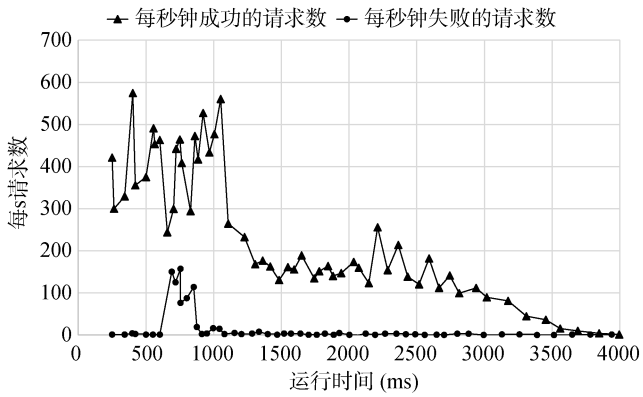


图 6 数据上链的吞吐率

Figure 6 The throughput of up data to blockchain

2) 溯源查询

如表 2、图 7、图 8 所示分别表示了溯源查询过程的数据报告、延迟和吞吐率。可以看出查询过程响应时间很短, 每秒钟的转移数很高, 且失败的转移数为零。区块链的查询只需要从一个节点查询数据即可, 所需时间较短。

表 2 溯源查询汇总报告

Table 2 The summary report of traceability query

Label	样本(个)	平均值(ms)	最小值(ms)	最大值(ms)
HTTP 请求	20000	1623	7	5290
总体	20000	1623	7	5290

标准偏差	异常(%)	吞吐量(s^{-1})	接收 KB/s	发送 KB/s	平均字节数
594.54	0.00	1088.7	2768.32	356.16	2603.9
594.54	0.00	1088.7	2768.32	356.16	2603.9

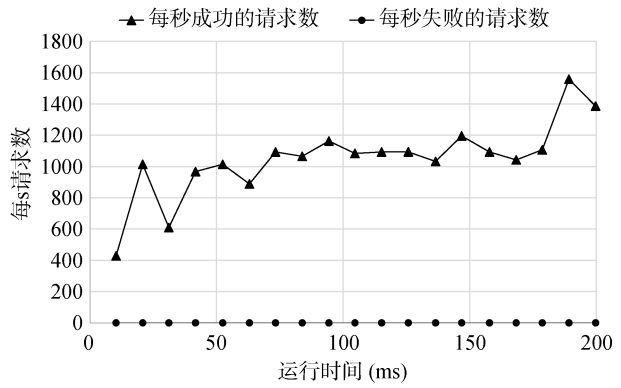


图 7 溯源查询的延迟

Figure 7 The delay of traceability query

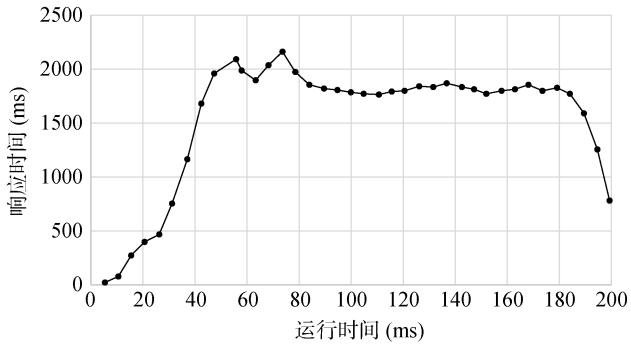


图 8 溯源查询的吞吐率

Figure 8 The throughput of traceability query

3) 节点、用户认证

如表 3、图 9、图 10 所示分别表示了物联网设备节点、农场用户登录认证过程的数据报告、延迟和吞吐率。可以看出登录过程响应时间很长, 每秒钟的转移数很低, 且失败的转移数也较高, 但是整体成功率依然高于失败率。用户、设备节点登录过程中, 为了保证用户身份的安全性, 需要进行多方面权限检查, 所以导致响应时间较长。

6 结束语

农业物联网海量数据的可信存储与溯源是未来智慧农业面临的一个难题。针对上述问题, 本文提出了一种基于区块链的农业物联网可信溯源体系, 解决了海量农业物联网数据的冗余过滤与链上存储溯源的问题。在本文提供的方案中, 农场管理员授权员工用户绑定所属物联网设备, 物联网设备定期上传的文本数据发送到智能合约进行聚类筛选, 剔除冗余数据后上链存证。图片、视频数据通过 IPFS 分布式网络进行存储, 获得唯一 CID 之后再保存至区块链, 在保证数据安全的前提下, 大大降低了区块链的存储压力。本方案为用户设计了溯源查询与数据校验机制, 能够实现按时间区间的数据溯源功能。最

表 3 节点\用户认证的汇总报告

Table 3 The summary report of node\User authentication

Label	样本(个)	平均值(ms)	最小值(ms)	最大值(ms)
HTTP 请求	10000	67665	128	414695
总体	10000	67665	128	414695

标准偏差	异常(%)	吞吐量(s^{-1})	接收 KB/s	发送 KB/s	平均字节数
52051.36	44.91	24.0	27.46	3.13	1174.2
52051.36	44.91	24.0	27.46	3.13	1174.2

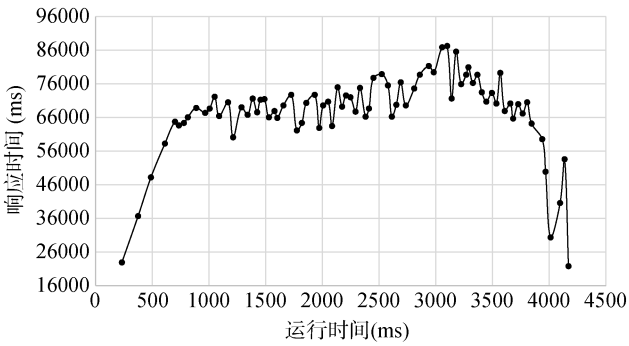


图 9 节点、用户认证的延迟

Figure 9 The delay of node\user authentication

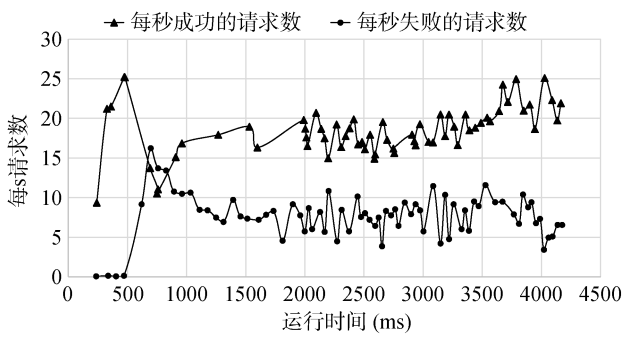


图 10 节点、用户认证的吞吐率

Figure 10 The throughput of node\user authentication

后, 本文利用 Hyperlegder fabric 联盟区块链框架实现了本方案, 安全分析和性能分析表明, 本方案在确保体系可靠、数据安全的同时, 能满足实际应用的性能需求, 同时大幅度降低了区块链的存储压力。

致　谢　本课题得到国家自然科学基金[61772150, 61862012,61962012]、广西重点研发计划[AB171950 25]、广西自然科学基金[2018GXNSFDA281054, 2018 GXNSFAA281232, 2019GXNSFFA245015, 2019GXN SFGA245004, AD19245048]、国家密码发展基金[MMJJ20170217]、鹏城实验室网络空间安全研究中心网络仿真项目[PCL2018KP004]的资助。

参考文献

[1] Sun S N, Wang X P. Promoting Traceability for Food Supply Chain with Certification[J]. *Journal of Cleaner Production*, 2019, 217: 658-665.

[2] Fan B L, Qian J P, Wu X M, et al. Improving Continuous Traceability of Food Stuff by Using Barcode-RFID Bidirectional Transformation Equipment: Two Field Experiments[J]. *Food Control*, 2019, 98: 449-456.

[3] Chen T B, Ding K F, Hao S K, et al. Batch-Based Traceability for Pork: A Mobile Solution with 2D Barcode Technology[J]. *Food Control*, 2020, 107: 106770.

[4] Alfian G, Syafrudin M, Farooq U, et al. Improving Efficiency of RFID-Based Traceability System for Perishable Food by Utilizing IoT Sensors and Machine Learning Model[J]. *Food Control*, 2020, 110: 107016.

[5] Xiao X Q, Li Z G, Matetic M, et al. Energy-Efficient Sensing Method for Table Grapes Cold Chain Management[J]. *Journal of Cleaner Production*, 2017, 152: 77-87.

[6] Cheraghalipour A, Paydar M M, Hajiaghaei-Keshteli M. Designing and Solving a Bi-Level Model for Rice Supply Chain Using the Evolutionary Algorithms[J]. *Computers and Electronics in Agriculture*, 2019, 162: 651-668.

[7] Dorri A, Steger M, Kanhere S S, et al. BlockChain: A Distributed Solution to Automotive Security and Privacy[J]. *IEEE Communications Magazine*, 2017, 55(12): 119-125.

[8] PILKINGTON M. Blockchain technology:Principles and applications[J]. *Research Handbook on Digital Transformations*, 2016:225-253.

[9] Geissler S, Prantl T, Lange S, et al. Discrete-Time Analysis of the Blockchain Distributed Ledger Technology[C]. *2019 31st International Teletraffic Congress*, 2019: 130-137.

[10] Watanabe H, Fujimura S, Nakadaira A, et al. Blockchain Contract: A Complete Consensus Using Blockchain[C]. *2015 IEEE 4th Global Conference on Consumer Electronics*, 2015: 577-578.

[11] Lu Y. The Blockchain: State-of-the-Art and Research Challenges[J]. *Journal of Industrial Information Integration*, 2019, 15: 80-90.

[12] Zhang M, Li C Y, Zhu M T. Application of blockchain technology in supply chain management [J]. *China Storage & Transport*, 2019(7): 103-104.
(张盟, 李成玉, 朱明桐. 区块链技术在供应链管理的应用[J]. *中国储运*, 2019(7): 103-104.)

[13] Zhao G Q, Liu S F, Lopez C, et al. Blockchain Technology in Agri-Food Value Chain Management: A Synthesis of Applications, Challenges and Future Research Directions[J]. *Computers in Industry*, 2019, 109: 83-99.

[14] Feng H H, Wang W S, Chen B Q, et al. Evaluation on Frozen Shellfish Quality by Blockchain Based Multi-Sensors Monitoring and SVM Algorithm during Cold Storage[J]. *IEEE Access*, 2020, 8: 54361-54370.

[15] Sun W J. Research on Computer Network Information Security and Protection Strategy [J]. *Information & Communications*, 2018, 31(6): 109-110.
(孙伟俊. 关于计算机网络信息安全及防护策略探究[J]. *信息通信*, 2018, 31(6): 109-110.)

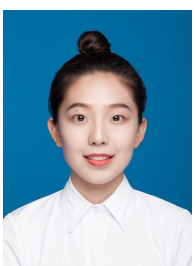
[16] Aqeel-ur-Rehman, Abbasi A Z, Islam N, et al. A Review of Wireless Sensors and Networks' Applications in Agriculture[J]. *Computer Standards & Interfaces*, 2014, 36(2): 263-270.

[17] Jiang J A, Wang C H, Liao M S, et al. A Wireless Sensor Network-Based Monitoring System with Dynamic Convergecast Tree Algorithm for Precision Cultivation Management in Orchid Greenhouses[J]. *Precision Agriculture*, 2016, 17(6): 766-785.

[18] Yu Q Y, Shi Y, Tang H J, et al. EFarm: A Tool for Better Observing Agricultural Land Systems[J]. *Sensors* (*Basel, Switzerland*), 2017, 17(3): 453.

[19] Yang Q, Lu R X, Rong C M, et al. Guest Editorial the Convergence of Blockchain and IoT: Opportunities, Challenges and Solutions[J]. *IEEE Internet of Things Journal*, 2019, 6(3): 4556-4560.

[20] Pan C T, Lee M J, Huang N F, et al. Agriculture Blockchain Service Platform for Farm-to-Fork Traceability with IoT Sensors[C]. *2020 International Conference on Information Networking*, 2020: 158-163.



陈锦雯　(1997−), 女, 广东南海人, 在读硕士研究生, 主要研究方向为区块链、人工智能安全等

唐呈俊　(1997−), 男, 汉族, 湖南湘潭人, 硕士, 在读硕士研究生, 主要研究方向: 人工智能安全和分布式系统安全等。

罗得寸 (1994−), 男, 广西来宾, 在读研究生, 学生, 主要研究方向为区块链、隐私保护。

丁勇 (1975−), 男, 广西桂林人, 博士, 桂林电子技术大学计算机科学与信息安全学院教授, 主要研究方向为密码学和信息安全等。

唐晨钧 (1997−), 男, 桂林永福人, 硕士研究生, 主要研究方向为工业控制网络入侵检测。