doi: 10.14132/j.cnki.1673-5439.2019.02.014

一种基于区块链的去中心化数据溯源方法

张国英,毛燕琴

(南京邮电大学 计算机学院,江苏 南京 210023)

摘要: 为了避免数据欺诈(如数据被篡改)等情况发生,必须保持数据创建、修改和转移的历史记录,即溯源。传统的数据溯源系统大多采用中心化的存储模式,存在易遭受内部、外部攻击,且有单点故障等弊端。区块链是一种随着比特币系统发展起来的,基于互联网的去中心化信任管理机制,其难以被篡改、可追溯等特性为安全的数据溯源提供了新的解决途径。文中提出了一种基于区块链的去中心化数据溯源方法,其中包括建立 PROV 溯源数据模型描述溯源数据;设计了一套溯源数据管理的合约,通过智能合约将溯源数据存储到区块链上,确保用户获得的溯源数据真实可靠;搭建了一个以太坊的私有区块链网络,通过设计一个基于 React 的 JavaScript Web 应用程序,仿真实现和测试了所提出的方法,测试结果证明所提出方案的正确性和可行性。

关键词: 区块链;数据溯源;智能合约

中图分类号: TP309 **文献标志码**: A **文章编号**: 1673-5439(2019)02-0091-08

Blockchain-based decentralized data provenance method

ZHANG Guoying, MAO Yanqin

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: To avoid data fraud, such as data tampering, the history of data creation, modification, and transfer, i.e., provenance must be maintained. Most of the existing data provenance systems adopt a centralized storage mode, which is vulnerable to internal and external attacks, and has the disadvantages of single point of failure. Blockchain is an Internet-based decentralized trust management mechanism developed with the Bitcoin system. Its hard-to-tamper and traceable features provide a new solution for secure data provenance. A blockchain-based decentralized data provenance method is proposed, including establishing a provenance data model to describe provenance data, and designing a set of provenance data management contracts to store provenance data on blockchain through smart contracts, to ensure that the provenance data obtained by the user is authentic and reliable. A private blockchain network of Ethereum is established, and a JavaScript Web application is designed based on React. Implementing and testing the method is conducted by the simulation. The testing results proves the correctness and the feasibility of the method.

Keywords: blockchain; data provenance; smart contract

随着计算机和移动互联网的飞速发展,数据呈爆炸式的增长,人类步入了一个大数据时代。在大数据背景下,数据本身具有的或潜在的价值使其成为重要的资产,这些数据在产生积极影响的同时也

收稿日期: 2019-03-07 **本刊网址**: http://nyzr.njupt.edu.cn

基金项目: 江苏省未来网络前瞻性研究项目(BY2013095-1-08)资助项目

作者简介: 张国英,女,硕士研究生;毛燕琴(通讯作者),女,副教授,yqmao@njupt.edu.cn

引用本文: 张国英,毛燕琴. 一种基于区块链的去中心化数据溯源方法[J]. 南京邮电大学学报(自然科学版),2019,39(2):91-98.

带来了一系列问题:生活中常见的数据一般都经过一系列的处理,简单的如复制,复杂的如各种转换和修改等,由于其中间过程不可知,人们在需要某特定数据时很难判断其来源和可靠性。事实也的确如此,派生的数据可能存在纰漏或由于利益驱使而被非法篡改等。数据溯源,是解决这些问题的有效措施之一[1]。

数据溯源,顾名思义是一种溯本追源的技术,根据记录的溯源数据追踪数据的源头及产生过程。通过数据溯源,人们可以查找要追溯的对象的出处,判断其真假以便维护自己的权利,或者更深入地了解所追溯对象的全面信息等。然而,溯源数据本身也是数据,如果没有适当的保护措施,随着数据经历不同应用层或不可信的环境,溯源数据有可能会遭到意外破坏甚至恶意篡改、删除[2]。迄今为止,对数据溯源的研究主要集中在溯源模型、溯源存储以及溯源应用等工作上,对确保溯源数据安全方面的研究较少。而数据溯源能够充分发挥作用是建立在溯源数据本身的可靠性和安全性上,所以安全的数据溯源研究非常重要,如何实现有效的数据溯源成为行业研究重点。

2014 年开始,比特币背后的区块链技术受到广泛关注并得到快速发展,其防伪溯源的特性被认为是最具有应用前景的区块链落地性能之一。目前,区块链技术已被应用于一些领域来实现数据追溯。例如,Provenance 项目[3]首次讨论了利用区块链(基于比特币区块链实现)进行数据溯源。在这项工作中,区块链交易用于存储从食品生产到被消费的溯源细节。Tian[4]建立了一个基于 RFID 技术、物联网和区块链技术的农产品溯源系统,并针对区块链需要存储大量数据而存在的可扩展性问题引入了 BigchainDB[5]的概念。ProvChain 系统[6]通过采用区块链技术解决云环境中的数据溯源问题。

上述利用区块链技术进行溯源的系统大多是针对特定应用场景,采用自己搭建的区块链来存储产品的溯源数据,其系统的安全性必然弱于比特币[7]或以太坊(Ethereum)[8]等区块链。本文基于以太坊区块链平台,主要进行了如下 2 个问题的研究:首先,数据溯源涉及的参与主体较多,包括数据源、数据传递方、数据审核方、数据使用方。这些参与方将数据的关键信息以及当前状态记录在区块链上。如何标识、验证各方的身份并确保在整个过程中数据没有被篡改。其次,一个数据对象的溯源数据通常要比数据本身多,各方怎样有效处理溯源数据并进行数据“上链”从而实现有效、可信的溯源。

1 相关技术

1.1 数据溯源技术

现有的溯源系统大多采用中心化存储方式,数据库中存储、维护的是数据的当前状态,数据的历史信息和处理过程通常存储在数据库日志中,用于故障恢复,并不直接提供查询服务(在系统无故障正常运行的情况下也不参与查询的处理)[9]。这种存储方式的缺陷是:(1) 这类系统内生性地受制于“基于信用的模式”的弱点,存在利益驱使导致的溯源数据造假的问题;(2) 如果中央服务器受到威胁,整个数据溯源系统可能会瘫痪,即单点故障;(3) 数据的历史信息存储在数据库日志中,难以实现追溯。在基于分布式架构的溯源系统中,各方分散孤立地记录和保存相关数据,形成信息孤岛,存在信息不对称,数据易被篡改以及追溯效率低的问题。

要从根本上解决上述问题,必须建立去中心化、可信的溯源机制,同时要求系统在通信故障甚至在被蓄意攻击时仍能确保数据存储的可靠性和正确性。不同于传统的分布式存储,区块链网络具有如下特点:首先,各参与节点拥有完整的数据存储过程,并且各个节点是独立、对等的,避免了单点故障;其次,每一个区块都包含前一区块的哈希,形成链式结构,数据的历史可追溯;再者,其分布式共识机制保证存储的最终一致性,存储数据的可信度与安全性。区块链技术的这些特性天然适合解决传统溯源的痛点。

1.2 区块链技术

公认的最早关于区块链的描述性文献是中本聪于 2008 年所撰写的《比特币:一种点对点的电子现金系统》[7],其中,区块用于记录比特币系统一段时间内的交易信息,而每一个区块都包含前一个区块的哈希,形成链式结构,这就是“区块链”名字的由来。区块链,其本质可以理解为一种由多方共同维护,利用块链式数据结构存储数据,利用分布式共识算法实现数据一致,利用密码学保证传输和访问安全从而具备数据难以被篡改、历史可追溯等特性的记录技术,也称为公用账本技术(Public Ledger Technology)[10]。

按照访问和管理权限,区块链目前从实现角度分为公有链、私有链和联盟链[11-12]。公有链,顾名思义,向所有人公开,每个人都能成为系统中的一个节点参与记账,依靠激励机制和密码学技术来维护数据安全,但面临高度去中心化和低吞吐量的两难,

以比特币为典型代表。私有链是指不对外开放,仅在特定组织内部使用,需要权限认证的区块链。联盟链由联盟内成员节点共同维护,节点通过授权后才能加入联盟网络,记账权由联盟成员协商确定,以超级账本 Hyperledger[13] 为典型代表。私有链和联盟链都属于专用的区块链,即专有链。公有链和专有链的区别在于读写权限以及去中心化的程度。一般而言,去中心化的程度越高,可信度越高。

无论是以太坊还是超级账本,均能较好地支持数据溯源,通常需要根据具体的应用场景选择适合的区块链。考虑到联盟链和私有链去中心化、开放程度低,而以太坊除了具备可信存储、可信验证等特性,还提供了灵活通用的智能合约,能够建立各种去中心化的服务,且便于部署和二次开发,因此本文基于以太坊平台实现数据溯源。

以太坊是当前被广泛采用的基于区块链的开源智能合约应用平台[8]。智能合约是区块链的核心构成要素,可以理解为具有状态的,由事件驱动的,部署在区块链中的,当满足特定的条件时能够按照程序设定自动执行,预设的条件为区块链所能验证且合约执行记录到区块链上的计算机程序。可以通过合约制定实现一些需要的功能,和一般的程序不同,智能合约一经发布便无法篡改。合约有 2 种数据存储方式: 一种是利用账户存储,另一种是通过事件,也称日志存储。这 2 种方法都在合约代码中实施[14-15]。考虑到日志存储后续不便于查询,本文采用账户存储的方式设计了一套溯源数据管理的合约,以实现数据实体溯源的存储和查询(即验证) 。

1.3　数据溯源模型

建立一个有效的数据模型是数据溯源技术的关键所在,根据模型可以初步确定溯源信息记录了哪些内容,及其存储和后期使用等操作。2012 年,第 4 届 IPAW (International Provenance and Annotation Workshop) 会议提出 PROV 数据模型 (PROV-DM) [16],其发展自 OPM[17] (Open Provenance Model) ,是目前溯源数据的最新标准。本文基于 PROV 数据模型建立了溯源数据模型来进行溯源数据的描述。

在 PROV-DM 中,溯源数据被定义为一条记录,描述了 3 个核心要素即 Entity(实体) 、Activity(活动) 和 Agent(代理) 及它们之间的关系。

实体是对要记录溯源信息的事物的描述,包括电子对象如文件或网页,物理对象如书籍、车辆以及抽象的概念。活动是发生或作用于实体的某种形式的操作,描述实体如何成为目前的状态以及实体的属性如何变化,包括实体的产生、转换或修改等。

活动可能会产生新的实体,以文档实体为例,修改文档会生成新的文本,用 wasGeneratedBy 表示; 活动也利用实体,用 used 表示。

代理为一个实体的存在、一个活动的发生或另一个代理的活动承担的某种形式的责任,可以是人、软件或其他可以赋予责任的角色的实体。当代理对活动负有责任时,PROV 数据模型用 associatedwith 表示代理与活动有关,实体属于一个代理,则用 attributedto 来表示。在不同版本之间,实体的一个或多个属性可能会发生变化,每个新的版本就是一个新的实体,wasDerivedFrom 表示衍生关系。图 1 说明了 PROV-DM 中包含的类型和关系。

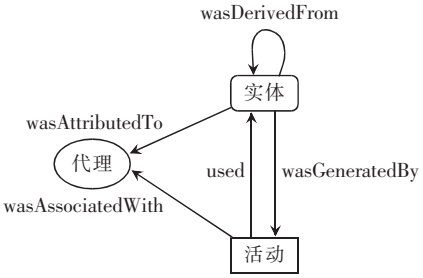

图 1　PROV 核心结构图

溯源信息的表示主要是通过溯源元数据来体现,本文将溯源数据模型表示为一个元组,描述要溯源的对象,对溯源对象所做的操作,用户即创建者、编辑者的信息,即溯源记录可以表示为(数据实体,对数据所做的操作,用户名) 。

在实际应用中,要处理的溯源数据更为复杂,但其本质仍然是元数据,描述人(或组织机构) 、数据实体和产生、影响并发送数据或其他活动的记录,可以根据具体的数据溯源应用场景进行扩展。

2　基于区块链的数据溯源方法

针对传统溯源系统存在的中心化存储、数据易被篡改等问题,本文基于区块链的分布式、难以篡改及可追溯等特性提出基于区块链的数据溯源方案。溯源系统的主要挑战是溯源数据的可信收集、可信存储和可信验证。本文假设已获取相关数据,不考虑其具体如何收集,重点关注溯源数据的存储和验证,即上述第二个研究问题。本文基于 PROV 数据模型建立了溯源数据模型,并设计了一套溯源数据管理的合约,通过智能合约将溯源数据存储在区块链上。对于第一个问题,即数据溯源参与各方的身份真实性验证问题,依靠密码学技术来解决,涉及哈

希算法[18]和数字签名[19]等。

2.1 参与方的身份真实性验证

可信的身份真实性验证是安全数据溯源的第一道屏障。数据溯源涉及众多参与主体,包括数据源、数据传递方、数据审核方、数据使用方,他们共同参与、维护区块链。各方首先注册为区块链中的用户,注册后每个用户会生成一个公钥/私钥对。公钥用于标识系统内用户的身份,私钥用于数字签名,确保用户身份的真实性,允许数据的接收者用以确认数据的来源,防止被人伪造。该实现基于 secp256k1 椭圆曲线数学的数字签名算法(Elliptic Curve Digital Signature Algorithm,ECDSA)。公钥通过私钥计算得出,但由公钥不能推导出私钥。数据是否被篡改的问题利用 keccak 256 哈希算法解决,与比特币、区块链采用的 SHA-256 算法相比安全性高,且效率也有提升。假设用户 A 需要将数据 M 传递给用户 B,A 首先采用 keccak 256 哈希算法计算出数据的哈希值,使用私钥对数据进行签名,之后将数据和签名一起发送。B 接收到信息后,使用 A 的公钥来解密签名,得到提供的数据哈希值,与对数据重新进行哈希运算得到的结果进行比较。若两个值一致,说明该信息确实是 A 所发出的,并且数据内容未被篡改。

2.2 溯源数据的可信存储

本文基于 PROV 数据模型(PROV-DM)建立了溯源数据模型来描述溯源记录,以便跟踪数据的变化并识别导致变化的实体,通过智能合约将溯源数据存储到区块链上。

这里,先简单考虑数据源将数据上传到区块链的情况,以上述场景为例,根据 PROV-DM 的定义,我们想要溯源的数据对象 M 即为实体,需要计算数据的哈希值 $H(M)$;用户 A 是数据 M 的所有者或创建者,定义为代理,之后需要对数据进行数字签名操作;由于 A 并未对数据做进一步处理,活动指数据源获取数据的时间。因此,溯源记录(Provenance Record,PR)至少包含:要溯源的数据的哈希值 $H(M)$,时间戳 timestamp,所有者签名 sig。即 $PR=(H(M),\text{timestamp},\text{sig})$。除此之外,本文定义了溯源对象的类别、格式和属性,活动的类型及代理的类型等,溯源数据模型如图 2 所示。

数据溯源的其他参与方(如用户 B)还可以对数据进行编辑等操作,但必须将对数据所做的操作记录在链上以供将来验证,则溯源记录需包含:之前版本的数据哈希值 $H(M)$、时间戳和签名的哈希值,当前版本(对数据做修改后)的数据哈希值 $H(M')$,所做的操作 activity,操作者签名 sig_k,即 $PR=(H(M'),H(H(M),\text{sig},\text{timestamp}),\text{activity},\text{sig}_k)$。上述示例的溯源关系图如图 3 所示。实体、活动以及代理分别用圆角矩形、矩形和椭圆形表示。这里可以找到溯源记录的细节。

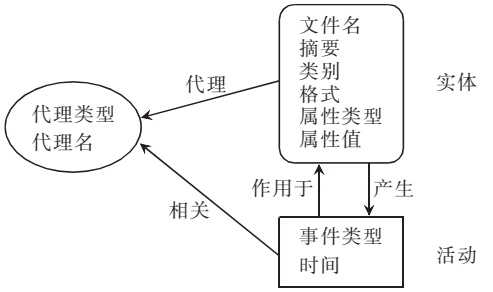

图 2 溯源数据模型

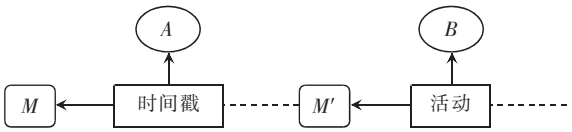

图 3 溯源链图

2.3 智能合约

基于 PROV 数据模型建立的溯源数据模型可以描述溯源信息,然而,未涉及如何以可靠和永久的方式记录必要的元数据和数据以进行日后的验证。本文基于以太坊平台,通过智能合约将溯源记录存储在区块链上,合约可以看作区块链中的自治代理,在满足预设条件时自动执行,减少了人工干预,从而实现溯源记录的可信存储和可信验证。

智能合约的构建和执行可分为如下 3 步:(1)多方用户通过约定,共同参与制定一份智能合约;(2)智能合约通过 P2P 网络传播,存储到区块链中;(3)当合约被触发后,能按照设定条件自动执行合约内容。

智能合约的主要组成部分包括状态和逻辑。本文基于溯源数据模型设计了一套智能合约,定义了溯源数据的结构,参与方读取和写入溯源数据的逻辑,分别为:Object. sol,Agent. sol,Event. sol 和对应的 ProvObject. sol,ProvAgent. sol 和 ProvEvent. sol,如图 4 所示。

在图 4 中,最顶层为合约名、中间层为属性(若存在),最底层为方法。以 ProvAgent. sol 合约为例,最顶层 ProvAgent 为合约名,中间层属性有 name、agentType 等,最底层函数有 setAgent、setObject 等。还可以看出合约之间的关系,有关联关系(图 4 中用实线箭头表示)和继承关系(图 4 中用带空心三角形的实线箭头表示)。

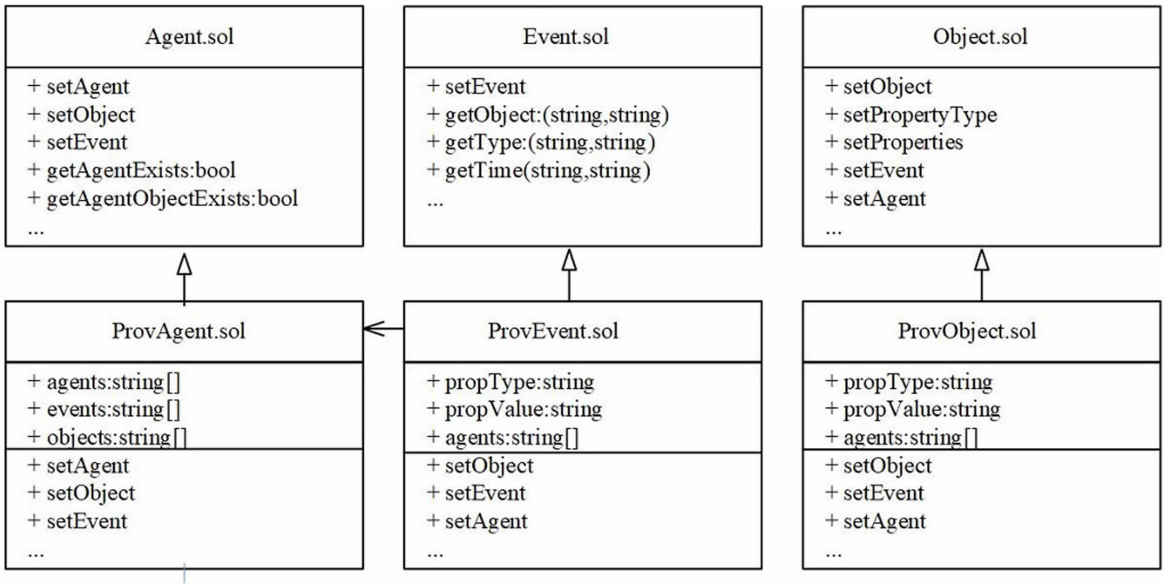

图 4　智能合约类图

合约通过 P2P 的方式在网络中传播,每个节点都会收到一份。区块链中的验证节点会对合约进行验证,验证的内容主要是合约参与者的私钥签名是否与账户匹配,验证通过的合约才会最终被写入区块链中。

合约部署到区块链后,区块链会返回合约的应用二进制接口(Application Binary Interface,ABI) ,包括变量、事件和可以调用的方法。与合约交互可以实现溯源数据的存储和查询。智能合约为数据的传输和存储提供了安全可靠的机制,保证了数据的难以篡改及可靠性。

3　基于区块链的数据溯源的实现

3.1　功能架构

本文提出的基于区块链的数据溯源方案旨在通过解决数据溯源参与各方的身份真实性验证,溯源数据应该记录什么内容以及溯源数据的分布式存储等技术问题来实现可信的数据溯源。主要实现了 2 个功能: 溯源数据的存储功能和溯源数据的验证功能。另外,本文还进行了可视化模块的设计,便于溯源数据的查看。因此,设计了如图 5 所示的基于区块链的数据溯源的功能模块图。

图 5 中的部件分为 2 类,一类是用矩形表示的实体部件,即构成区块链网络的区块链节点。另一类是用矩形表示的功能部件,包括身份真实性验证功能,溯源数据的存储功能和溯源数据的验证功能及可视化模块。

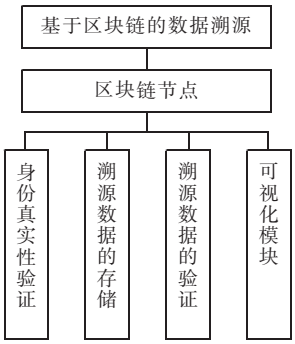

图 5　功能模块图

基于上述功能模块的划分,设计了基于区块链的数据溯源的功能实现架构,如图 6 所示。该架构由 3 部分组成,分别是区块链节点搭建的区块链网络、智能合约和应用程序的前端。

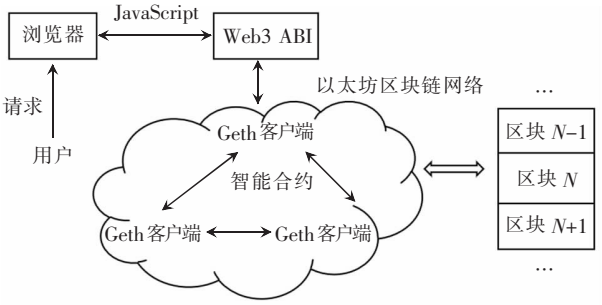

图 6　实现架构图

下面,分别对各部分进行阐述。

(1) 区块链网络由区块链节点共同搭建(本文采用 Geth(Go-ethereum) 客户端,在后面详细介绍) ,主要负责溯源数据的存储、验证以及区块链网络的维护。

(2) 智能合约部分,通过 Solidity 语言可以灵活编写适用于应用的、全网节点都需要严格执行的智能合约脚本,合约描述基于区块链的数据方案所实现的业务逻辑,包括溯源数据存储和验证的逻辑。

(3) 前端页面展示部分,在区块链网络正常运行,合约部署成功的基础上,参与方可以基于前端页面实现溯源数据的存储和溯源记录验证功能。

3.2 仿真实现

根据图 6 的功能实现架构,分析基于区块链的数据溯源的实现流程。首先,数据溯源的各参与方必须注册为区块链的用户,区块链返回给用户一对公钥和私钥,公钥作为用户的身份标识,私钥用于数字签名。在应用的前端,对于参与方上传的数据,系统会计算出数据的摘要便于之后核对,溯源数据包含了图 2 所示溯源数据模型中的信息。参与方提交溯源数据后会产生相应的事务,需要对事务进行签名。当满足预先设定的条件时,触发智能合约执行,将溯源数据存储到区块链中。

合约一经发布,便不可篡改。溯源数据的验证是基于溯源数据存储的,数据存储成功后,可以通过将重新计算得到的溯源对象的摘要与从区块链上获取的摘要值相比较来验证溯源数据是否被篡改。

为了评估本文提出的方案的正确性和可行性,本文基于以太坊平台,建立了一个私有区块链测试环境模拟区块链应用环境,并设计了基于 React 的 JavaScript web 应用程序对上述方案进行仿真实现。溯源数据管理合约编写好之后,需要以太坊客户端来部署和运行智能合约。客户端选择的是目前比较主流的 Geth 客户端,所有 Geth 客户端共同参与、维护区块链。

4 实验与结果分析

4.1 实验环境搭建

实验的运行系统是 Ubuntu16.04,采用目前比较流行的 Truffle 开发框架来进行智能合约的编译、部署和测试。智能合约必须要部署到链上进行测试,Truffle 官方推荐 Ganache 客户端来进行智能合约的测试,待测试成功后将合约部署到 Geth 客户端上,应用前端基于 React 框架开发,通过与合约进行交互来实现溯源数据的存储和查询。

4.2 实验过程与结果分析

首先,采用 Solidity 语言编写智能合约,Solidity 源码文件通常以.sol 作为扩展名。然后,部署合约到区块链上,如图 7 所示,将合约编译,生成字节码文件,编译过程会产生智能合约的 ABI,之后可以调用智能合约中的函数来实现数据的存储。部署到 Ganache 客户端如图 8 所示。

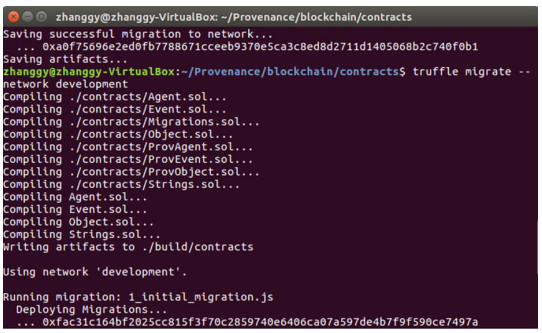

图 7 部署智能合约

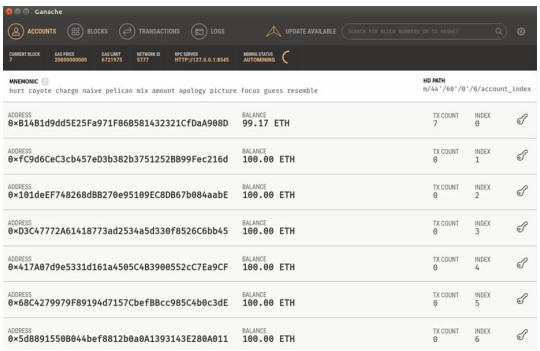

图 8 部署合约到 Ganache 客户端

对于溯源数据的存储,用户上传文件后,系统会计算出文件的摘要,然后,用户根据 2.2 节定义的溯源数据模型提交溯源记录,提交后,智能合约被触发执行,执行结果记录到区块链上。如图 9 所示,将示例 test1 文件上传,可得到其摘要为 da5af2a27cee31b29485e621509b7695a7e2eef61caeff7eb04da2e220b5c442。

图 9 溯源数据的存储

数据上链成功后,可以从区块链上获取之前上

传的溯源数据信息,如图 10 所示。

通过重新计算上传数据的摘要并与区块链上的摘要值进行比较可以验证溯源数据是否被篡改。由图 9 和图 10 分别显示的溯源数据的存储和查询信息可见,前后的数据哈希值相同,证明数据未被篡改以及数据上传方的身份真实,说明本文所提出的方法的正确性。

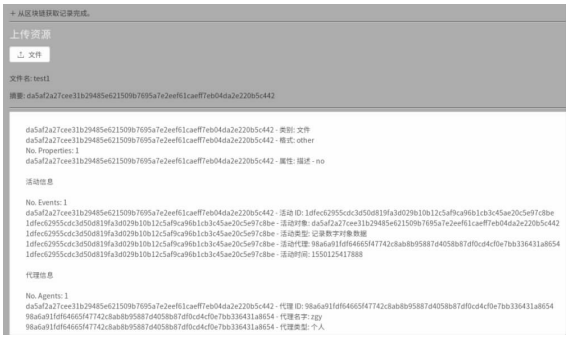

图 10　溯源数据的查询/验证

考虑数据可能被篡改的情况,如将示例文件 test1 的内容稍作修改或者完全替换成另一个文件如 test2,可以通过从区块链上查询 test2 的溯源数据进行验证,如图 11 所示,test2 文件的哈希值为 fd678e53ca219e76e79f1feef0d5ca2ea3c8f2afe915ae9d7182cc226acf030f,前后哈希值不一致,且溯源信息为空,说明此数据不是真实有效的。

图 11　被篡改的数据的验证

上述实验说明了本文所提出的方案能够确保溯源数据的安全存储,并防止恶意篡改,实现可信的数据溯源。由于本文的实验是在私有区块链测试环境中进行,在真实环境中可能还需要根据具体应用进行适当的调整。

5　结束语

随着信息技术的繁荣发展,人类进入了大数据时代,保持数据创建、修改和转移的历史(即溯源),确保数据的质量,特别是确保数据未被篡改,越来越重要。针对传统数据溯源存在的中心化存储,数据易被篡改等问题,本文利用区块链分布式,可信存储以及可信验证的特性,提出基于区块链的数据溯源方案。但从存储角度看,区块链链上只存储溯源数据的哈希值,不适合存储大量的数据。本文基于 PROV 数据模型建立了溯源数据模型来描述溯源信息,并设计了一套溯源数据管理的智能合约,将溯源记录存储在区块链上,确保用户获得的溯源信息真实可靠。但区块链只能保证上链数据不被篡改,并不能保证链下数据的真实性,物联网技术能一定程度上解决这个问题,未来,这两种技术的结合有望实现链下和链上整个数据链的可信存储和可信溯源。

参考文献:

[1] MING H, ZHANG Y, FU X. Survey of data provenance [J]. Journal of Chinese Computer Systems, 2012, 33(9): 1917 – 1923.

[2] HASAN R, SION R, WINSLETT M. The case of the fake picasso: preventing history forgery with secure provenance [C]//7th USENIX Conference on File and Storage Technologies. 2009.

[3] Blockchain: the solution for transparency in product supply chains [EB/OL]. [2019-01-05]. https://www.provenance.org/whitepaper.

[4] TIAN Feng. A supply chain traceability system for food safety based on HACCP, blockchain & Internet of Things [C]// International Conference on Service Systems and Service Management (ICSSSM). 2017: 1 – 6.

[5] MCCONAGHY T, MARQUES R, MÜLLER A, et al. BigchainDB: a scalable blockchain database (DRAFT) [EB/OL]. [2019-01-05]. http://www.blockchain.jetzt/wp-content/uploads/2016/02/bigchaindb-whitepaper.pdf.

[6] LIANG X, SHETTY S, TOSH D, et al. ProvChain: a blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability [C]// 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID). 2017: 468 – 477.

[7] NAKAMOTO S. Bitcoin: a peer-to-peer electronic cash system [EB/OL]. [2018-12-20]. http://www.cryptovest.co.uk/resources/Bitcoin%20paper%20Original.pdf.

[8] Ethereum project [EB/OL]. [2019-01-20]. https://www.ethereum.org/.

[9] 钱卫宁,邵奇峰,朱燕超,等. 区块链与可信数据管理: 问题与方法 [J]. 软件学报,2018,29(1): 150 – 159.
QIAN Weining, SHAO Qifeng, ZHU Yanchao, et al. Research problems and methods in blockchain and trusted data management [J]. Journal of Software, 2018, 29(1): 150 – 159. (in Chinese)

[10] 姚国章,吴春虎,余星. 区块链驱动的金融业发展变革研究 [J]. 南京邮电大学学报(自然科学版),2016,36

(5):1-9.
YAO Guozhang, WU Chunhu, YU Xing. Research on the development and reform of financial industry driven by blockchain [J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2016, 36(5):1-9. (in Chinese)

[11] BUTERIN V. On public and private blockchains [EB/OL]. [2019-01-20]. https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/.

[12] 王化群,吴涛. 区块链中的密码学技术[J]. 南京邮电大学学报(自然科学版),2017,37(6):61-67.
WANG Huaqun, WU Tao. Cryptography in blockchain [J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2017, 37(6): 61-67. (in Chinese)

[13] Hyperledger project [EB/OL]. [2019-01-24]. https://www.hyperledger.org/.

[14] WOOD G. Ethereum: a secure decentralised generalised transaction ledger [EB/OL]. [2018-11-10]. https://ethereum.github.io/yellowpaper/paper.pdf.

[15] RAMACHANDRAN A, KANTARCIOGLU M. SmartProvenance: a distributed, blockchain based dataProvenance system [C]// Proceedings of the 8th ACM Conference on Data and Application Security and Privacy. 2018: 35-42.

[16] GROTH P, MOREAU L. PROV-overview: an overview of the PROV family of documents [EB/OL]. [2019-01-10]. https://www.w3.org/TR/prov-overview/.

[17] MOREAU L, CLIFFORD B, FREIRE J, et al. The open provenance model core specification (v1.1) [J]. Future Generation Computer Systems, 2011, 27(6): 743-756.

[18] GUNTER D V, SHACHAF L M. System and method of transmitting encrypted packets through a network access point: US 6751728 [P]. 2004-06-15.

[19] RIVEST R L, SHAMIR A, ADLEMAN L. A method for obtaining digital signatures and public-key cryptosystems [J]. Communications of the ACM, 1978, 21(2): 120-126.