基于区块链的信用信息可信共享模型研究 *

王继林, 陈英姿, 冯小青

(浙江财经大学 信息管理与工程学院,浙江 杭州 310000)
通讯作者: 冯小青, E-mail: fenglinda@zufe.edu.cn

摘 要: 由征信行业现状可知,信用信息缺乏可信的共享环境是造成信用信息共享程度不高,形成信用信息孤岛的重要因素.为了鼓励信用信息的共享,提出基于区块链的信用信息共享模型,为信用信息共享平台建立和可信共享服务提供可行思路.首先提出了改进的实用拜占庭容错算法,然后阐述了信用信息共享模型的各个阶段及其实现原理.提出的共享模型主要功能有:保护信息主体权益,解决共享过程中数据、第三方数据存储等不可信问题,共享数据在链下存储与访问.针对共享数据通过云服务器进行存储和访问时的隐私保护问题,分别探讨了两种存储和访问方式.最后总结了本文提出方案的优势和影响.

关键词: 信用信息;可信共享;PBFT;区块链;隐私保护

A Blockchain-based Trusted Credit Information Sharing Scheme

Abstract: According to the situation of credit reporting industry at present, the lack of trusted credit information sharing environment is the most important reason for low sharing willingness, and thus form information islands. In order to encourage the sharing of credit information and provide feasible way for credit information sharing platform and trusted shared services, we propose a blockchain-based credit information sharing model. We first introduced improved Practical Byzantine Fault Tolerance algorithm, then introduced the stages of credit information sharing model and their concrete realizations. The model proposed has the following functions: protect the rights and interests of information subjects, sharing incentive, solve the untrusted problems of data storage in a third-party in the process of sharing data, store and access the data outside the blockchain. Two ways of storing and accessing are discussed respectively for privacy protection when using the cloud server. Finally, conclude the advantages and impacts of the proposed scheme.

Key words: credit information; trusted sharing; PBFT;blockchain; privacy-preserving

信用信息是能够反映个人、企业等信用状况的信息,是个人和企业在社会活动中产生的各种各样的信用记录.从全球信用体系来看,由于各国经济、文化、历史等发展不同,目前除少数美国、欧洲等发达国家拥有完善的征信体系外,很多国家仍然缺少有效的信用信息共享机制.新时代变化下信用体系建设迎来更大的需求,如大数据融入传统征信已经成为一种趋势[1],个人征信拥有广阔的发展前景,小微企业融资难、融资贵问题还很严重,信用信息共享机制的缺乏将严重阻碍征信业的发展.有效的信用信息共享机制的建立可以促进市场主体间进行信用信息共享,解决市场主体间信用信息不全面、不真实、不准确的问题,从而促进征信行业的发展.

然而,有效的信用信息共享机制的建立面临着许多挑战.从数据采集方面看,征信机构的信用信息来源存在局限性,一方面各个数据源间没有有效的机制激励信息共享,企业数据壁垒严重,如互联网征信机构所使用的征信数据往往是自身所依赖的平台数据,外部合作数据较少,形成数据孤岛,另一方面市场上流通的数据面临数据污染、非法采集等问题,真实性和合法性得不到保证,而传统的征信机构数据除了一些来自传统金融行业的数据外还需要通过购买、抓取等方式获得.从数据使用方面看,可能存在数据的二次使用和隐私泄露的问题,大数据时代对数据隐私保护提出了新的挑战.从数据管理方面看,传统的数据交易中心出于利益原因,不仅存在篡改或者删除数据的可能性,还可能存在隐私信息泄露和转卖等现象.有效的信用信息共享机制需要保证信用信息

* 基金项目: 国家自然科学基金(NSFC-61202197);教育部规划基金项目(12YJAZH136);浙江省公益技术应用研究项目(No.2016C31081)

Foundation item: Natural Science Fund of China (NSFC-61202197); Planning Fund Project of Ministry of Education (12YJAZH136); Public Welfare Technology and Industry Project of Zhejiang Provincial Science Technology Department (No.2016C31081)

的真实性、合法性,提供激励信用信息拥有者参与共享的机制,确保信用信息提供者对共享资源的控制、交易透明、隐私保护等需求.区块链技术具有去信任、不可篡改、数据溯源等特点,区块链技术的出现为解决信用信息共享难问题提供了一个可行的方向.

区块链技术[2]被认为能够解决征信领域的很多问题[3].区块链通过全局账本这种特殊的数据存储结构来存储所有交易等信息,和共识机制、密码学算法等技术使得区块上所有的信息可以追溯和不可篡改;通过共识机制保证交易的合法性以及所有合法节点维持相同的全局账本;通过数字签名技术来保证信息的不可抵赖等.目前区块链在征信领域的应用已经受到了不少关注,出现各种基于区块链的征信项目,如云棱镜[4]通过区块链技术为用户形成个人信用凭证,把基于区块链的真实信用数据应用于消费金融风控等领域;AFC 信用链[5]将区块链技术和人工智能技术结合以提供征信服务;LinkEye 构建基于区块链的征信联盟,在联盟成员间共享失信人名单等服务[6].

本文针对信用信息共享难等问题提出了一种基于区块链的信用信息共享模型,以提供可信的信用信息共享服务.本文主要通过引入授权、异议处理等环节解决数据确权问题和保护用户权益,利用区块链技术来消除由于第三方存储数据带来的安全和隐私问题,以及保证共享交易透明、数据不可篡改,从而允许共享参与方对共享交易情况和共享数据等进行溯源和审计.本文提出的方案中共享数据将上传到云服务器以减少区块链上数据存储的压力.针对共享数据链下存储与访问的问题,本文探讨了两种存储和访问方式.提出采用代理重加密技术对云服务器上存储的共享数据进行多次共享,云服务器存储信用信息提供者加密的共享数据,因此云服务器以及其他没有相应密钥的用户无法得到共享数据的具体内容;提出使用门限秘密共享技术和多云服务商来解决共享数据集中存储和隐私保护的问题,共享数据经过处理后被分散存储到各个云服务器中,同时拥有单独份额数据的云服务商不能恢复出完整数据内容.

本文接下来结构安排如下:第 1 节介绍与基于区块链的数据共享有关的研究工作,第 2 节介绍部分与本文提出研究内容相关的技术. 第 3 节介绍了本文采取的改进的 PBFT 共识机制.第 4 节介绍信用信息可信共享的基本框架和流程,并详细描述该共享机制的实现过程.第 5 节分别探讨了代理重加密和门限秘密共享技术在本文中的应用.最后对本文提出的方案进行了总结.

1 相关工作

区块链因其去信任、不可篡改、数据溯源等特点受到越来越广泛的关注,但目前将其与信用信息共享建设相结合的研究还相对较少.本章节将首先重点介绍围绕基于区块链的数据共享的相关研究,目前关于数据共享的研究场景主要是物联网数据共享和电子医疗数据共享.

Missier[7]等人利用区块链技术和智能合约建立一个分散的、可信的、透明的和开放的体系结构,用于物联网流量测量和合约履行.该框架建立在物联网数据代理和发布/订阅模型上,通过区块链技术实现代理的去信任化,通过智能合约实现透明的价格交易.Ouaddah[8]提出基于区块链的去中心化的访问控制系统 FairAccess 来直接访问物联网设备数据.其访问控制主要由两个阶段构成,一是授予访问权限,被请求者根据请求者的请求内容将其访问策略以锁定脚本的形式封装在交易上,并输出代币作为未花费交易输出,二是请求者完成访问条件通过解锁脚本的形式发送交易,交易验证成功后请求者就可以凭借代币进行访问.

Xue[9]等人结合医疗行业现状,采用改进的股份授权证明机制提出一个基于区块链的医疗数据共享模型,以解决医疗数据共享中权限审查和数据校验问题.Xia [10]等人提出了一个基于区块链的数据共享框架,用以解决云环境中电子医疗记录访问控制的问题.该系统主要通过联盟链和密码学密钥的设置来实现用户对共享数据的访问请求.Xia[11]等人在相同医疗数据共享场景下,提出 MeDShare 系统使用智能合约和访问控制策略跟踪数据行为和撤销恶意的访问,以实现医疗数据共享系统数据溯源、审计和控制的功能.

以上研究都针对专门领域的共享与访问进行了探讨,利用区块链技术解决了各自领域内面临的一些难题.其他共享模型如 Enigma[12]的计算模型致力于多方共享数据存储和计算时的隐私保护,利用区块链技术管理访问控制、身份和用作防止篡改的事件日志.与本文研究内容直接相关的是[13],Roman 等人讨论了欧洲地区阻碍

征信创新的核心数据问题,并以征信的视角提出了一个可信的数据交易市场框架,讨论新兴技术如区块链、智能合约等对征信受益人之间数据共享的价值作用.国内相关研究中王强等人[14]提出构建基于区块链的联盟链来搭建征信数据共享交易平台,给出征信机构与征信机构共享部分用户信用数据和征信机构从其他机构获取用户信用数据两种模式的思路.郭树行等人[15]从监管层面给出了基于区块链的数据交易模式.两篇论文都为区块链技术在征信行业应用提供了建议和构建思路.本文介绍和设计了基于区块链的信用信息可信共享方案,研究和探讨构建过程中区块链共识机制选择、信用信息存储与共享隐私保护等问题,以期解决征信领域信用信息共享难,以及共享过程中数据、费用结算、第三方数据存储不可信等问题.

2 相关技术简介

2.1 区块链和共识机制

区块链作为比特币的底层技术,最早来源于中本聪 2009 年发表的论文[16],但目前还未形成统一的定义.区块链将交易信息以及交易 hash 值、时间戳等参数打包成区块,通过一定的机制将这些区块以密码学方式形成一种按时间顺序排列的链式结构.区块链具有不可篡改、透明、可溯源等特点,并且重新定义了信任,支持安全、快速、可信和透明的解决方案,被认为是一种革命性的技术[17].

区块链技术有时也被称为分布式账本技术,它通过共识机制来保证各个节点维护的全局账本的一致性.区块链所在的系统是分布式系统,各个节点通过 P2P 网络进行有效通信.在分布式系统中,网络节点可能会因为技术故障或者任意恶意行为(拜占庭故障)失去响应或者出现错误,比特币系统采用的是公有链形式,使用了工作量证明来解决区块链的共识问题,但这种挖矿的形式会耗费大量的计算机资源.本文提出的方案主要是帮助征信机构从其他行业机构获取信用信息以实现可信共享,信用信息一般来自于特定的行业机构,而联盟链是一种只针对某个特定群体成员的区块链类型,联盟链相比公有链在效率和灵活性上更具优势,因此本文研究采用联盟链构建信用信息共享平台.

2.2 实用拜占庭容错算法

实用拜占庭容错(Practical Byzantine Fault Tolerance,PBFT)算法[18]是 Miguel Castro 和 Barbara Liskov 于 1999 年提出的, 解决了原始拜占庭容错算法效率不高的问题.PBFT 算法能够解决客户端/服务器框架的一致性的问题,且无须消耗大量资源,效率更高,被认为是能够应用于联盟链的一种有效的算法.

PBFT 算法主要包括一致性协议(agreement)、视图转换协议(view change)和检查点协议(checkpoint).PBFT 算法中主要包含两种类型节点:主节点(primary)和从节点(backups).所有的服务器节点都在相同的配置信息下工作,这个配置信息被称作视图(view),当主节点失效时就需要启动视图更换过程.检查点协议用于定期维护服务器状态,清理日志来节约资源和提升系统性能.主节点负责接收客户端的请求并将其排序,然后主要采用三阶段协议将请求广播给各个从节点以实现一致性,三阶段协议的基本思路如下:

(1) 预准备(pre-prepare)阶段:该阶段主节点为从客户端收到的请求分配序号,然后广播 PRE-PREPARE 消息给从节点.从节点验证 PRE-PREPARE 消息,如果从节点接受了这条 PRE-PREPARE 消息,它就会进入 prepare 阶段.

(2) 准备(prepare)阶段:各个从节点进入到自己的 prepare 阶段后,向全网广播一条 PREPARE 消息,并且将 PRE-PREPARE 消息和 PREPARE 消息写入自己的消息日志.同时,如果一个节点收到 $2f$ (f 为可容忍的拜占庭节点数)个从其它节点发来的与 PRE-PREPARE 消息一致的 PREPARE 消息,以及对消息验证通过后,则将 PREPARE 消息写入日志中,准备阶段完成.预准备阶段和准备阶段确保了所有正常节点对同一个视图中的请求序号达成一致.

(3) 确认(commit)阶段:各个节点在准备阶段完成后,广播一条 COMMIT 消息给其它所有节点,如果一个节点验证 PREPARE 消息为真,且收到 $2f+1$ 条与 PRE-PREPARE 消息一致的 COMMIT 消息,则执行相应请求,请求完成后各个节点发送回复给客户端.prepare 阶段和 commit 阶段用来确保那些已经达

到 commit 状态的请求即使在发生 view change 后在新的 view 里依然保持原有的序列不变.

2.3 默克尔树

默克尔树(Merkle Tree)也称为 hash 树,是一种树形数据结构,其叶节点是数据块的哈希值,非叶节点是其子节点的哈希值,最后得到一个根哈希值.梅克尔树最简单和常见的形式是二进制梅克尔树,如图所示,每个非叶子节点是其两个子节点的 hash 值,比特币系统使用梅克尔证明(Merkle proofs)来确定包含的一笔交易.Merkle Patricia Tree(MPT 树)是更复杂的 Merkle Tree,它实际上是一种 trie 前缀树, 用来存储所有的(key,value)对.由于状态树需要经常地进行更新,以太坊使用 Merkle Patricia Tree 来维护账户数据及其他状态数据.本模型区块中所有的交易记录通过二进制 Merkle Tree 得到一个 Transaction Root 存储在区块头上,检验交易的存在;使用二进制 Merkle Tree 记录各个信用信息提供者上传的数据的哈希值,以快速验证云数据库中数据上传/下载过程中数据的完整性;使用 Merkle Patricia Trees 来记录征信机构和信息信息提供者的账户余额信息.

3 改进的 PBFT 共识算法

本文研究采用联盟链构建信用信息共享模型,使用资源使用更少、效率更高的 PBFT 算法[24]来建立模型.直接将 PBFT 算法应用在模型中存在些问题,如 PBFT 算法是基于 C/S 架构设计的,不能直接应用在区块链模型的 P2P 网络架构.PBFT 算法的协议中需要大量的两两交互,网络通信量和网络带宽消耗较大,且当节点数量增大时效率明显下降,不适合应用在 P2P 网络存在大量节点和区块消息较大的情况.

因此,模型对 PBFT 算法进行了适当改进.首先将原来 PBFT 算法的客户端和服务器角色重新进行设计,其次由于 PBFT 算法的特点,设计区块链共识算法时主节点需要执行所有的验证、排序等工作,主节点工作量很大,因此引入验证节点对客户端请求进行验证并背书,以此减轻主节点任务量,提升模型性能.最后模型参考 DPoS 共识机制对原有 PoS 共识机制的改进思想,授权部分节点作为共识节点参与共识.初始化共识节点设置根据现有企业信用评级进行授权,后续根据企业信用变化进行授权.其他节点作为普通节点不参与共识,只保存或查询共识节点维护的账本记录.改进后的 PBFT 算法降低了系统通信量和提高了系统吞吐量.改进后的 PBFT 算法流程如下:

Step 1: 客户端发起交易并广播给验证节点

Step 2: 验证节点验证交易并背书,将签名信息返回给客户端

Step 3: 客户端收集足够背书信息后向共识节点发送交易请求

Step 4: 各节点将接收到的交易请求暂时存储在自己的交易池中

Step 5: 每隔一段时间主节点将自己交易池中的交易打包成区块并发送 PRE-PREPARE 消息<<PRE-PREPARE,*v*,*n*,*d*>,*m*>给从节点,其中 *v* 是视图编号,*m* 是背书后的请求消息,*d* 是请求消息的摘要,*n* 为新区块高度

Step 6: 从节点验证 PRE-PREPARE 消息,若接受该消息则发送 PREPARE 消息<PREPARE,*v*,*n*,*d*,*i*>给其他共识节点,*i* 为节点信息

Step 7: 共识节点验证收到的 PREPARE 消息,若消息为真,并且收到 2*f* 个从其它节点发来的与 PRE-PREPARE 消息一致的 PREPARE 消息,则再广播一条 COMMIT 消息<COMMIT,*v*,*n*,*d*,*i*>给其它共识节点

Step 8: 若共识节点收到了 2*f*+1 条与 PRE-PREPARE 消息一致的 COMMIT 消息,则共识完成,按请求执行交易,并将请求交易以区块形式记录在本地区块链账本上.

系统每隔固定时间后根据企业信用排名变化重新设置共识节点,在基于改进的 PBFT 算法设计的区块链共识机制中,验证节点负责对请求进行验证,主节点负责将客户端的请求以新区块的形式广播,和从节点交互实现区块链账本的一致性.

4 基于区块链的信用信息可信共享模型

4.1 模型概述

征信大数据包含赊销、借贷等活动的历史记录信息,这些信息通常来自于商业银行、信托公司、租赁公司等,随着大数据征信的发展,越来越多的非传统数据也被用作征信数据,如政府部门公共信息,电费、通讯费、水费等信息,甚至是电商支付数据、社交数据等.征信活动通常包括征信机构主动去调查被征信人的信用状况和依靠授信机构或其他机构批量报送被征信人的信用状况.本文研究征信机构通过信用信息共享平台来采集征信数据以提供征信服务,设计基于区块链的信用信息可信共享模型.根据研究内容,用户客户端使用人群一般为个人征信机构(credit bureau)、信贷登记系统(credit registry)和企业征信机构(commercial credit reporting company),为了保证征信的真实性原则,征信机构应给予被征信人一定的知情权,因此模型还设计了信息主体角色.模型的主要参与主体还有信用信息提供者、云服务提供商以及监管机构,如图 1 所示,他们在模型中担任角色如下:

(1) 信息主体

信息主体是信用信息的归属对象, 对其信用信息应当拥有控制权.他们的信用信息通常存储在一些金融机构如信贷机构、电商金融机构和非金融机构如电信运营商、政府部门,其信用信息的使用应当被充分告知并授权使用.

(2) 征信机构

征信机构系统地收集、整理、加工和分析关于企业或个人信用信息为其提供信用报告,并提供多样化的征信服务,其自身一般不产生征信数据.收集和整理全面、真实地信用信息是提供信用报告的关键和前提.

(3) 信用信息提供者

信用信息提供者是拥有赊销、借贷信息的银行、证券等金融机构和其他电信运营商、政府部门等非金融机构.这些拥有企业或个人信用信息的机构共享信用信息,由征信机构生成信用报告和提供征信服务.信用信息提供者在模型中对应区块链中的共识节点和普通节点.

(4) 云服务提供商

云服务提供商提供云存储和云计算等服务,当信用信息提供者向征信机构共享信用信息时,先将信用信息上传到云服务器中,征信机构从云服务器中下载共享数据.

(5) 监管机构

为了保障信用信息共享中各个主体的权利和平台的健康运行,设立监管机构负责对共享过程中各方的异议进行处理,如信息主体对信用信息共享者上传的信用信息存在异议等.

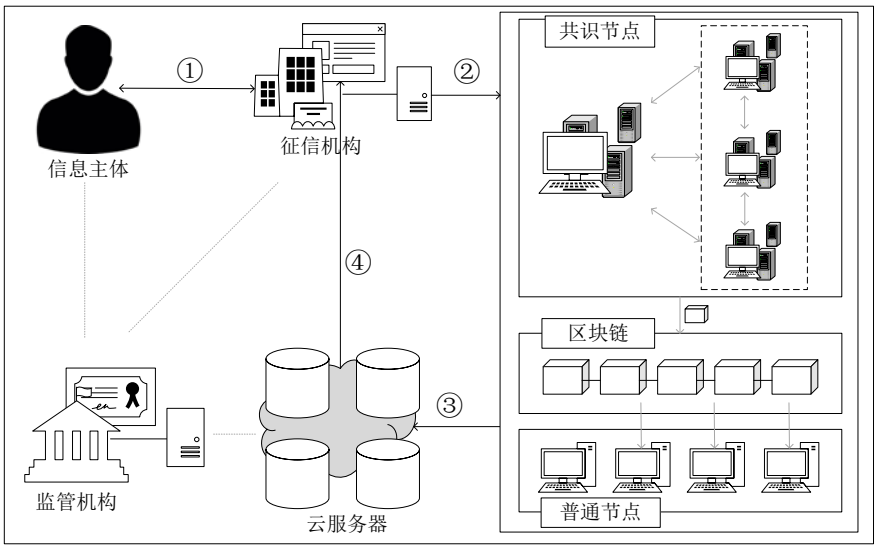

Figure 1 信用信息共享模型结构

基于区块链的信用信息共享流程(图 2)如下:1)征信机构在获取用户信用前须获得用户对相应数据的授权. 2)征信机构凭借授权信息通过客户端向信用信息提供者发布共享请求任务以获得相应数据. 3)信用信息提供者在收到任务后按照一定的数据格式要求将数据以数据交易形式上传到区块链网络(信息主体不介意数据公开和数据较小情况下),或将数据加密上传到共享云服务器后将数据摘要信息等上传到区块链网络. 4)验证节点执行校验和签名背书操作,共识节点维护区块链账本信息,普通节点更新数据. 5)征信机构凭借私钥访问共享数据,若征信机构或信息主体对共享信息内容存在异议,则由监管机构介入进行审计.用户授权和征信机构的请求记录,以及信用信息提供者上传的数据信息保留在区块链上,区块链具有不可篡改、可追溯的特点,由多个节点共同维护,因此可以作为各种矛盾解决的依据,提供一个公平的信用信息共享环境.

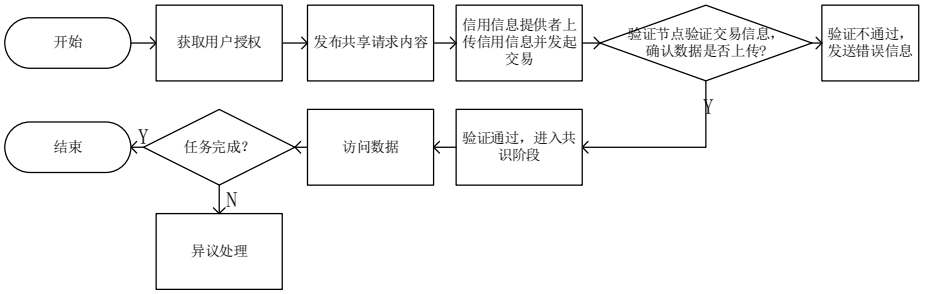

Figure 2 基本工作流程图

4.2 模型设计

本文设计的信用信息可信共享模型主要包括四个阶段:获取用户授权、发布请求、信息可信共享和异议处理.本节将描述各个阶段的具体实现过程.

4.2.1 获取用户授权

征信机构(Credit Bureau, C)需要得到信息主体(Information Subject, S)的授权后才能发布与信息主体相关的信用信息请求任务.在该阶段中征信机构首先发送请求数据的地址(*CompanyName*)、内容(*Resource*)、使用目的(*Purpose*)和范围(*Scope*)以及包含自己公钥的数字证书(Digital Certificate,*DC*)给信息主体,向其告知将要采集的信息,信息主体授权后发送自己签名后的授权信息 $Sig_{SI}\left(hash(m)\right)$ 返回给征信机构.

C→S: *m*, 其中 *m*=(*CompanyName,Resource,Purpose,Scope,DC*)

S→C: $Sig_S\left(hash(m)\right)$

4.2.2 发布请求

在取得信息主体的授权信息后,征信机构通过客户端向全网发布共享请求.发布的请求应当遵从一定的格式,包括请求类型 *V*,时间戳 *T*,锁定时间 *t*,锁定时间表示该请求的有效时间,还需要附上获取用户授权阶段的采集消息 *m* 和授权信息 $Sig_S\left(hash(m)\right)$,客户端验证签名后输出签名验证结果 *Veried_m*.验证通过后该请求被客户端广播到网络各节点.请求(*Request*)内容如下所示:

Request= (*V, T, t, index,* Pk_{SI} *, input* (), *output* ())

其中 $input=\left(m, Sig_{SI}\left(hash(m)\right), Sig_{CB}(m,p)\right)$, *output*=(*Veried_m*)

4.2.3 信息可信共享

信息可信共享阶段包含上传数据和广播交易、交易上链和共享访问三个步骤:

(1) 上传信息和广播交易

信用信息提供者根据共享请求,将信用信息按一定的数据格式以数据交易形式上传到区块链网络,或者按照要求加密信用信息上传到云服务器中,将数据摘要信息广播到区块链网络.

(2) 交易上链

信用信息提供者向区块链网络发起数据交易(Transaction, Tx)应当满足一定格式,包括交易类型 V、接收的请求(*Request*)、信用信息提供者的公钥(Pk_{CIP}) 、交易序号(*Nonce*)等内容,同时输入信用信息或自己上传到云服务器中的加密数据的摘要信息.首先由验证节点验证数据是否上传和数据交易格式是否正确,若验证节点验证成功则对交易签名背书,当该交易获得一定数量的签名背书后则再转发给区块链共识节点.主节点打包一段时间内验证完成的交易,以区块的形式广播给从节点,与从节点进入三阶段共识阶段,三阶算结束后各个节点将心区块加入到自己维护的区块链账本上.一旦交易出现在区块链上,则表明该交易所请求的任务已经上传完成,征信机构直接或者凭借私钥访问共享数据.

主节点打包形成的区块结构包括区块头和区块体,具体区块结构如图所示,其中 *previous block hash* 表示上一个区块的区块头 hash 值,*Nonce* 值表示区块高度,交易根表示区块内所有交易树的根哈希值,data 根表示上传至云服务器中数据树的根哈希值,这两颗 merkle 树的存在方便了数据的验证查询.

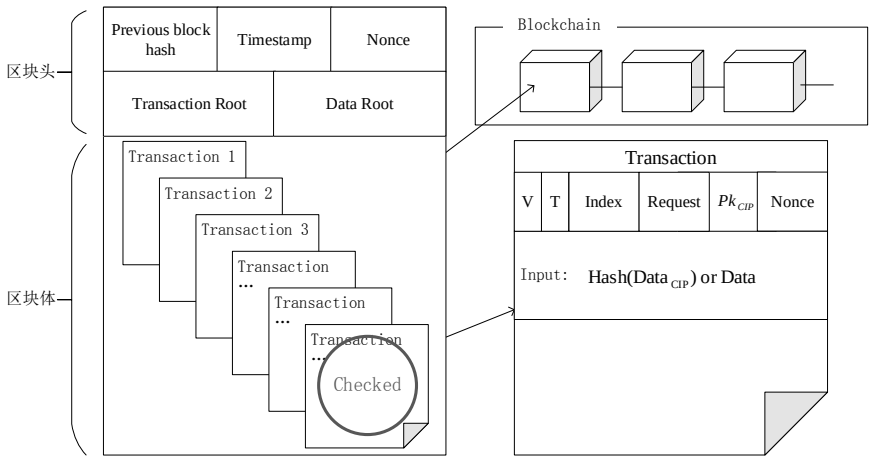

Figure 3 区块结构图

4.2.4 异议处理

在信息共享过程中,若征信机构或信息主体对共享信息内容存在异议,则由监管机构进行审计.监管机构根据区块链上用户授权信息、交易信息,以及云服务器中数据信息等对异议进行处理.为了控制区块链上交易上链的时间,异议处理过程应当在链下进行,由监管部门制定相应规则进行处罚和处理.

5 数据存储与共享访问

在以上本文提出的方案中使用了云服务商来提供云存储服务,为了实现数据存储时的隐私保护,信用信息提供者将数据加密后上传至该云服务器中,与该数据上传有关的交易出现在区块链账本上时,征信机构就可凭借其私钥从云服务器中下载数据.然而使用单一云服务商使得共享数据过于集中,增加了数据存储和安全风险.当某些数据需求量大时,为了保护隐私信用信息提供者需要频繁上传相同数据的加密信息.本文针对这两个问题分别探讨了基于门限秘密共享和基于代理重加密的区块链云存储与共享方案.

5.1 基于门限秘密共享的区块链云存储与共享方案

在基于门限秘密共享的区块链云存储与共享方案中,秘密分发阶段的公开参数信息由信用信息提供者以交易的方式记录在区块链上,并将共享数据的各个秘密份额存储到各个云服务提供商,该种方案下单个云服务商无法恢复数据内容,只有当征信机构根据区块链上的信息从一定数量的云服务器获取数据才能恢复最终共享信息.本文介绍采用经典的 Shamir 门限秘密共享方案[19]设计协议方案.

(1) 参数设置:选择大素数 $p(p>n, n$ 为云服务商数量)和门限值 t.

(2) 秘密分发:

a) 从有限域 GF(p)中选取 n 个互不相同的非零元素 $x_1, x_2, \cdots, x_n$ 分别分发给 n 个云服务提供商,并将其已交易的形式公开在区块链账本上.

b) 从有限域 GF(p)中选取 t-1 个元素 $a_1, a_2, \cdots, a_{t-1}$,构造 t-1 次多项式:

$$h(x)=\left(S+a_1x+a_2x^2+...+a_{t-1}x^{t-1}\right)\bmod p$$

其中 S 为共享原始数据,根据该多项式计算 $s_i=h(x_i),(1\le i\le n)$,将其作为秘密份额分别分发给 n 个云服务提供商.

(3) 秘密恢复:

征信机构从 t 个或 t 个以上的云服务商下载数据恢复密钥,首先构造

$$h(x)=\sum_{i=1}^{t}s_i\prod_{j=1,j\ne i}^{t}\frac{x-x_j}{x_i-x_j},$$

然后计算 $S=h(0)=\sum_{i=1}^{t}s_i\prod_{j=1,j\ne i}^{t}\frac{x_j}{x_j-x_i}$ 得到原始数据.

5.2 基于代理重加密的区块链云存储与共享方案

信用信息提供者在收到请求后,将信用信息按一定的数据格式加密后上传到云服务器中,当收到相同数据请求时,信用信息提供者可以使用代理重加密技术来实现对云服务器中已上传数据的共享访问,避免频繁上传数据.代理重加密允许一个代理者将由授权人加密的密文转换为由被授权人公钥加密的密文,其概念由 Blaze 等人于 1998 年在欧密会提出[20].在这种机制中代理者不能得到任何关于授权人的明文信息,实现了授权人与被授权人之间的数据共享.本文将在[21]给出的代理重加密算法基础上设计协议方案,其中将由主节点充当代理角色进行重加密操作.

(1) 密钥生成 $KeyGen(1^k)\to(pk_i,sk_i)$:根据安全参数 1^k,输出用户 i 的一对公私钥 (pk_i,sk_i)

(2) 重加密密钥生成 $ReKeyGen(pk_A,sk_A,pk_B,sk_B)\to r_{A\to B}$:根据授权人(信用信息提供者)A 的公私钥和被授权人(征信机构)B 的公钥,输出代理重加密密钥 $r_{A\to B}$

(3) 加密 $Encrypt(pk_A,m)\to c_A$:信用信息提供者 A 根据其公钥和信用信息 m,输出 m 的密文,并将其上传到云服务器中.

(4) 重加密 $ReEncrypt(rk_{A\to B},c_A)\to c_B$:信用信息提供者将自己到征信机构的代理重加密密钥 $r_{A\to B}$ 以交易形式发送到区块链上,共识节点验证交易,由主节点进行重加密操作, 根据代理重加密密钥 $rk_{A\to B}$ 和 A 的密文 c_A,输出针对 B 的重加密密文 c_B.

(5) 解密 $Decrypt(sk_B,c_B)\to m$:征信机构凭借其私钥 sk_B 和密文 c_B,输出原始明文 m.

6 结论

本文提出了基于区块链的信用信息可信共享模型,解决了征信行业面临的一些难点、痛点,为信用信息共享提供了公平、透明、可信的共享环境.信用信息的每一次共享记录都被记录在区块链中,以公开、不可篡改的分布式账本形式被公共查询,也便于监管机构进行审计和异议处理.本文构建方案采用的是基于 PBFT 共识机制的联盟链形式,避免了公有链资源浪费和效率低下等缺点.共享信息以加密形式存储在云数据库,并分别设计了基于门限秘密共享和基于代理重加密的区块链云存储与共享方案,避免了数据的集中存储和资源提供方频繁上传相同数据.本方案可以解决大数据时代下的数据确权问题,以及对共享数据的监管和追责,为信用信息可信共享的实现提供思路.

References:

[1] Ruiz S, Gomes P, Rodrigues L, et al. Credit Scoring in Microfinance Using Non-traditional Data[M]// Progress in Artificial Intelligence. 2017:447-458.

[2] Zheng Z, Xie S, Dai H, et al. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends[C]// IEEE International Congress on Big Data. IEEE, 2017.

[3] LIU Cailin. Research on the Application of Block Chain Technology in the Construction of Social Credit System in China [J]. Credit Reference, 2017, 35(8):28-32.

[4] http://www.yintongzhengxin.com,2018/831

[5] http://www.antifraudchain.com/,

[6] https://www.linkeye.com/

[7] Missier P, Bajoudah S, Capossele A, et al. Mind My Value: a decentralized infrastructure for fair and trusted IoT data trading[C]// Iot Conference. 2018.

[8] Ouaddah A, Elkalam A A, Ouahman A A. Towards a Novel Privacy-Preserving Access Control Model Based on Blockchain Technology in IoT[M]// Europe and MENA Cooperation Advances in Information and Communication Technologies. Springer International Publishing, 2017.

[9] Xue T F, Fu Q C, Wang C, et al. A Medical Data Sharing Model via Blockchain[J]. Acta Automatica Sinica, 2017, 43(9):1555-1562.

[10] Xia Q, Sifah E, Smahi A, et al. BBDS: Blockchain-Based Data Sharing for Electronic Medical Records in Cloud Environments[J]. Information, 2017, 8(2):44.

[11] Xia Q, Sifah E B, Asamoah K O, et al. MeDShare: Trust-Less Medical Data Sharing Among Cloud Service Providers via Blockchain[J]. IEEE Access, 2017, 5(99):14757-14767.

[12] Zyskind G, Nathan O, Pentland A. Enigma: Decentralized Computation Platform with Guaranteed Privacy[J]. Computer Science, 2015.

[13] Roman D, Stefano G. Towards a Reference Architecture for Trusted Data Marketplaces: The Credit Scoring Perspective[C]// International Conference on Open and Big Data. IEEE, 2016:95-101.

[14] WANGQiang,QINGSude,BAJieru. Discussion on block chain applicationin creditinvestigationindustry [J]. Telecommunication network technology, 2017(6).

[15] GUO Shuohang, SONG Ziqi. Design and application of blockchain pattern for credit information industry [J]. Chinese Journal of Network and Information Security, 2018(4).

[16] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[J]. Consulted, 2008.

[17] Underwood, S.: Blockchain beyond Bitcoin. Communications of the ACM 59(11) (2016) 15-17

[18] Castro M, Liskov B. Practical Byzantine fault tolerance[C]// Symposium on Operating Systems Design & Implementation. ACM, 1999:173-186.

[19] Shamir A. How to Share a Secret[J]. Communications of the Acm, 2011, 22(22):612-613.

[20] Blaze M, Bleumer G, Strauss M. Divertible protocols and atomic proxy cryptography[J]. Lecture Notes in Computer Science, 1998, 1403:127-144.

[21] Ateniese G, Fu K, Green M, et al. Improved proxy re-encryption schemes with applications to secure distributed storage[J]. Acm Transactions on Information & System Security, 2006, 9(1):1-30.

[22] Kiayias A, Russell A, David B, et al. Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol[C]// International Cryptology Conference. Springer, Cham, 2017:357-388.

[23] https://bitshares.org/technology/delegated-proof-of-stake-consensus/

[24] Castro M. Practical byzantine fault tolerance and proactive recovery[J]. Acm Transactions on Computer Systems, 2002, 20(4):398-461.

附中文参考文献:

[3] 刘财林. 区块链技术在我国社会信用体系建设中的应用研究[J]. 征信, 2017, 35(8):28-32.

[9] 薛腾飞, 傅群超, 王枞,等. 基于区块链的医疗数据共享模型研究[J]. 自动化学报, 2017, 43(9):1555-1562.

[14] 王强, 卿苏德, 巴洁如. 区块链在征信业应用的探讨[J]. 电信网技术, 2017(6).

[15] 郭树行, 宋子琦. 面向征信的区块链模式设计与应用研究[J]. 网络与信息安全学报, 2018(4).