软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2019,30(6):1614−1631 [doi: 10.13328/j.cnki.jos.005739]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于联盟链的物联网动态数据溯源机制*

乔 蕊[1,2], 曹 琰[1,2], 王清贤[1,2]

[1](信息工程大学,河南 郑州 450001)

[2](数学工程与先进计算国家重点实验室,河南 郑州 450001)

通讯作者: 乔蕊, E-mail: 18033023@qq.com

摘 要: 物联网动态数据安全保护的重点是拒绝非授权用户的篡改,实现对物联网动态数据操作的过程留痕和追踪溯源.为解决大量物联网设备产生的动态数据安全存储与共享问题,建立了物联网动态数据存储安全问题的数学模型,提出了用于实现操作实体多维授权与动态数据存储的双联盟链结构,设计了基于验证节点列表的共识算法,给出了一种基于联盟链的动态数据溯源机制优化方案.进行了分析及实验,物联网操作实体个数小于 10^6,操作实体授权特征值位数取 64 时,攻击者篡改授权的概率几乎为 0,验证了所提方案具有较强的抵抗双重输出攻击、重放攻击及隐藏攻击的能力,能够有效杜绝攻击者对物联网动态数据的篡改、伪造等非授权访问操作,具有较好的应用价值.

关键词: 物联网;动态数据;溯源机制;联盟链;多维授权

中图法分类号: TP309

中文引用格式: 乔蕊,曹琰,王清贤.基于联盟链的物联网动态数据溯源机制.软件学报,2019,30(6):1614−1631. http://www.jos.org.cn/1000-9825/5739.htm

英文引用格式: Qiao R, Cao Y, Wang QX. Traceability mechanism of dynamic data in Internet of things based on consortium blockchain. Ruan Jian Xue Bao/Journal of Software, 2019,30(6):1614−1631 (in Chinese). http://www.jos.org.cn/1000-9825/5739.htm

Traceability Mechanism of Dynamic Data in Internet of Things Based on Consortium Blockchain

QIAO Rui[1,2], CAO Yan[1,2], WANG Qing-Xian[1,2]

[1](Information Engineering University, Zhengzhou 450001, China)

[2](State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China)

Abstract: The focus of dynamic data security protection of IoT (Internet of things) is to reject tampering of unauthorized users, meanwhile, to realize the process in evidence and track tracing of the dynamic data operation of IoT. In order to solve the problems such as secure storage and sharing of dynamic data generated by a large number of IoT devices, firstly, a mathematical model for the security of dynamic data storage was established, as well as dual consortium chain structure is proposed to realize multidimensional authorization and dynamic data storage of operational entities. Then, a consensus algorithm based on VNL (verification nodes list) was proposed. After that, an optimization scheme of dynamic data traceability mechanism based on consortium block chain was put forward. Finally, through open experiments and performance analysis, it shows that when the eigenvalue of the operation entity authorization is 64 and the number of IoT operation entities is less than 10^6, the succeed probability of the attacker is almost 0. Meanwhile the proposed scheme can effectively avoid potential attacks on dynamic data, such as double output attacks, replay attacks and hidden attacks and so on. Thus it can

* 基金项目: 国家重点研发计划(2016YFB0800203); 河南省高校科技创新团队支持计划(17IRTSTHN009)

Foundation item: National Key R&D Project of China (2016YFB0800203); Program for Innovative Research Team (in Science and Technology) in University of Henan Province (17IRTSTHN009)

本文由区块链与数字货币技术专题特约编辑斯雪明教授和陈文光教授推荐.

收稿时间: 2018-06-25; 修改时间: 2018-10-12; 采用时间: 2018-12-18; jos 在线出版时间: 2019-03-27

CNKI 网络优先出版: 2019-03-27 16:40:21, http://kns.cnki.net/kcms/detail/11.2560.TP.20190327.1640.005.html

effectively prevent the attackers from unauthorized manipulation of the IoT, such as tampering or counterfeiting under approved accession mode. The scheme has good application value while ensuring the dynamic data storage security.

Key words: Internet of things; dynamic data; traceability mechanism; consortium chain; multidimensional authorization

物联网广泛应用于工业、医疗、教育、供应链等众多领域,在多方授权实体的参与下,以时间为基本维度产生新的数据,本文称为动态数据(dynamic data,简称 DD),这些数据的操作要求安全、可追溯,以便用于各种取证及决策等[1,2].动态数据具有以下特征.

(1) 持续性.伴随着时间的推移,在多方实体参与下持续产生新的动态数据;

(2) 时间敏感性.动态数据是对产生时间和应用时间敏感的数据,例如取某段时间内产生的动态数据对未来进行预测等;

(3) 多维度.在不同应用场景中,动态数据存在除时间维度外的多种维度数据,如供应链系统中参与交易的实体地址、交易额等,工业控制系统中工程文件的操作、权限的设置等[3];

(4) 可用性.动态数据应具备可用性,支持用户尤其是企业用户的安全管理需求,如分析查看日志信息、了解数据使用情况以及展开违法操作调查等.

可追溯性是确保动态数据完整性和可靠性的重要前提,是物联网系统中动态数据可用性的重要体现.因此,保证动态数据的可追溯性具有重大意义.

动态数据的可追溯性包括动态数据本身及对动态数据的历史操作的可追溯,其目的是确保动态数据的完整性和可靠性,即保证在存储及转移的过程中未发生篡改或伪造.近年来,网络犯罪已经从个人行为转变为有组织的行为,攻击在数据篡改、伪造等方面越来越专业[4].而现有的数据基础设施最初设计为应用于合法的数据存储场景,通常采用中心数据库与访问控制、接入认证、信息加密、数字水印等传统密码学方法结合的安全手段,将动态数据集中存储和处理[5−8],或采用以云计算为基础的数据存储,将各种数据资源抽象成资源池[9,10],供用户使用.上述设计方案存在安全隐患,例如,高价值数据集中存储极易被攻击、算法复杂度高等.因此,防止动态数据被篡改和被伪造成为具有挑战性的任务.为了提高动态数据存储的安全性,必须从两方面对数据进行保护:一方面验证动态数据的正确性,避免被篡改、伪造;另一方面,实现对动态数据操作历史的可追溯,提供数据恢复能力.

本文提出一种动态数据溯源机制:采用区块链方式记录网络动态数据流转的全生命周期,对动态数据进行记录、追溯、确权,以从源头保证该数字资产以及所代表信息的真实性,减少甚至阻止篡改攻击的可能性.分析共识终端最大化自身收益的局部行为与保障动态数据存储安全性和有效性整体目标的一致性,提出适用于动态数据存储的共识机制,减少算力浪费.采用密钥分发机制,分层传递并验证各级动态数据存储平台信息,相邻层次间通信采用二次散列迭代的方式,利用公钥加密正反向不对称性,增加系统被攻破的难度.

本文的主要工作如下:建立了物联网动态数据存储安全问题的数学模型,提出了用于实现操作实体多维授权与动态数据存储的双链结构;分析群体博弈过程中,单个节点进行决策的诚实行为动机及特定行业背景下分布式节点合作的本质,提出了一种适用于动态数据存储的共识机制,以共识机制保障动态数据存储的安全性;提出了一种动态数据溯源信息在物联网多方实体间动态流转的分层溯源机制,通过公钥机制构建通信通道,完成动态数据在通信系统中端到端加密安全传输,利用加密运算正反向不对称性,有效防止动态数据被篡改、伪造.

本文第 1 节介绍相关工作.第 2 节抽象出对物联网动态数据操作具有通用性的方式和过程,提出动态数据存储问题模型.第 3 节介绍区块链相关概念,分析联盟链解决物联网动态数据储存的适用性,基于博弈论理论分析物联网环境下各节点达成共识的边界条件,提出基于验证节点列表的联盟链共识算法,进一步提出基于上述共识算法的物联网动态数据存储体系结构.第 4 节通过理论分析及实验部署证明本方案对于抵御常见攻击,实现动态数据操作溯源的有效性.第 5 节总结全文.

1 相关工作

在物联网飞速发展的同时,物联网动态数据安全面临严峻的挑战,许多研究人员开展了对中心数据库和云

存储服务安全性的研究[11−18].文献[12]审查了 29 个不同的基于 USB 对数据库的攻击,并将它们分为 4 个主要类别.提出了一种方法来识别每个攻击的相关和脆弱的 USB 外围设备和硬件,但这意味着该方法允许攻击发生,存在数据库被破坏的可能性.文献[13]提出了一种数据库入侵检测机制,通过在网站上使用 SQL 注入,记录入侵者的所有活动来增强数据库的安全性.管理员可以查看详细信息,阻止攻击者向数据库注入恶意代码,窃取、销毁或修改数据库.但单方信任机制无法控制拥有高级访问权限的工作人员对动态数据进行恶意篡改或伪造[14].此外,文献[15]指出,动态数据多由智能处理终端或现场采样设备采集、编码和存储,这些设备的处理和存储性能有限,加之动态数据的持续性特征导致其随时间增长的数据量较大,因此,复杂度较高的安全算法不适用解决动态数据的防篡改、防伪造问题.文献[16]指出:由于云端数据允许多授权用户访问,无法对数据信息的去向和各级主体的操作历史提供充分的证据,因此无法满足某些特殊领域(如工业控制系统、溯源系统等)对系统动态数据的整个访问过程进行审计的需求,一旦出现问题难以定责.文献[17]为解决云平台下不受控制的恶意修改可能破坏共享数据的可用性问题,提出了一种公共审计解决方案,可以同时保护群体成员的身份隐私和身份可追溯性.但在云平台下,用户无法与云服务提供商建立信任,并确保服务协议仅使用 Web 前端接口[18].为了避免敏感信息被窃取、篡改和伪造,系统需要一个可靠的云平台服务提供商.

鉴于传统数据库和云存储服务存在安全隐患,且不可避免,实现动态数据的可追溯性是保障物联网动态数据安全应用的关键.通过分析近几年溯源领域的论文,发现目前许多现代可追溯系统是基于射频识别(radio frequency identification devices,简称 RFID)技术[19−21].文献[19]提出了一种电子谱系的食品可追溯系统,利用射频识别技术跟踪、定位物品在被无线传感器网络收集储存和运输过程中的温度和湿度.但针对传感器数据损坏或丢失的问题,文中仅采用预测的方式,无法从根本上解决.文献[20]提出一种基于公钥加密技术的高级数据保护方案,该方案能够实现 RFID 数据的可追溯性和链性活动.与传统的 RFID 安全方案相比,该方案适用于没有任何加密功能的标准 RFID 标签,并且不需要中心数据库.但操作的复杂度较高,对标签性能要求较高.文献[21]对 RFID 标签的加密能力进行研究,提出了能够执行加密操作的 RFID 标签体系结构.但是加密功能增加了标签的成本,并且涉及昂贵的身份验证计算.

根据物联网系统的应用特点和要求,需要采取安全措施对系统产生的动态数据进行存储和共享,实现动态数据的可追溯.RFID 在溯源系统中应用的研究仅适用于对有形资产的追溯,不适用于对物联网动态数据的追溯.而云计算等中心化数据库仅仅实现了动态数据的存储,在抵抗恶意用户(包括具有高级权限的内部人员)篡改、伪造动态数据方面具有天然的缺陷[22,23].区块链在不引入第三方中介机构的前提下,可以提供去中心化、不可篡改、安全可靠等特性保证[24].目前,已有研究来创建更具可扩展性的区块链,文献[25]提出 Bitcoin-NG 区块链协议,它是拜占庭容错的,共享相同的信任模型,具有较强的鲁棒性.文献[26]提出的 GHOST 规则解决了提高块创建速度的问题,这是对比特币节点构建和重新组织区块链的一种改进.文献[27]提出可以重组链,构建区块的有向非循环图,并降低允许交易的授权规则.虽然这些模型显著提高了运算速度,但它们可扩展性较差,并且需要复杂的数据结构或共识机制.文献[28]介绍了一种基于区块链技术的去信任物联网设备匿名共享方法,对于解决物联网动态数据的溯源问题有一定借鉴.本文通过分析动态数据面向多机构的区块链应用场景,在联盟链的基础上,提出一种全新的去中心化基础架构与分布式计算模型,将区块的共识及可见性限制在联盟链内部,有效地降低了参与记账节点的数量,实现快速共识验证,可以很好地解决动态数据的存储及溯源问题.

2 动态数据安全问题建模

定义 1. 操作实体(operation entity,简称 OE)是指动态数据 D 生命周期数据流动过程中的所有参与实体,用四元组$\langle ID,FOE,ROE,D\rangle$表示.

- ID 是操作实体的标识符,用以对操作实体进行唯一标识;
- $FOE_j=\{(ID_i,X_t)\mid ID_i\xrightarrow{X_t}ID_j,i,t\in N,i\neq j\}$ 表示对操作实体 OE_j 授权的操作实体标识 ID_i 及授权类型 X_t 的集合;
- $ROE_j=\{(ID_i,X_t)\mid ID_j\xrightarrow{X_t}ID_i,i,t\in N,j\neq i\}$ 表示操作实体 OE_j 授权的操作实体标识 ID_i 及授权类型

X_t的集合;

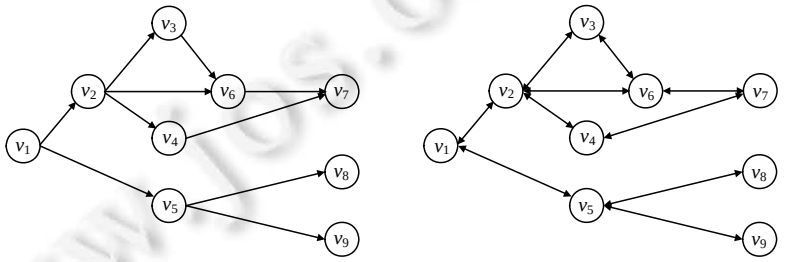
- $D_j=\{D_{(ID_i,X_t)}\mid(ID_i,X_t)\in ROE_j\}$表示操作实体$OE_j$的授权操作执行后得到的动态数据集合.

定义 2. 实体OE_i的授权属性特征值用集合l_i表示,H:$\{0,1\}^*\to\{0,1\}^\omega$,$l_{ij}\in\{0,1\}^\omega$,$H$是理想化的哈希函数,$\omega$为哈希后得到的特征值长度,$\delta$是授权操作 DAG 图中节点最大入度:

$$l_1=\varepsilon,l_i=\{l_{ij}\mid l_{ij}=H(SK_{p_j},l_{p_j},ID_i),where(p_j)=parents(i),j<\delta,i>1\}.$$

由定义 2,若操作实体的授权属性集合 l_i 中存在多个元素,表示存在多个父节点对其授权,SK_{p_j} 为实体 OE_i 父节点的私钥,从集合 l_i 中删除某个元素表示某个父节点取消授权;反之亦成立.$l_i=\varepsilon$表示其拥有根权限,$l_i=\varnothing$表示该实体授权为空.

与图 1 对应的动态数据操作轨迹用多维 DAG 图表示,如图 2 所示,节点表示动态数据文件授权,其中,圆形节点表示对应操作实体仅持有一项操作授权,将产生一份动态数据文件;柱状节点表示对应操作实体持有多项操作授权即多维授权,将产生多份动态数据文件;有向边表示动态数据文件在新操作下的演进轨迹.



(a) 操作授权 (b) 授权追溯

Fig.1 Schematic diagram of dynamic data operation authorization

图 1 动态数据操作授权示意图

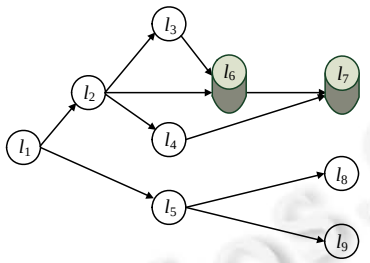


Fig.2 Strong DAG of dynamic data operation track

图 2 动态数据操作授权多维 DAG 图

定义 3. 对于 DAG 图中节点 i,入度为δ_i,若$\delta_i\leqslant 1$,保持节点不变化;$\delta_i>1$,节点为每个入度授权的集合,其元素个数$|l_i|=\delta_i$,这样得到的图称为多维 DAG 图.$\delta_i>1$ 时,存在多个父节点对节点 i 的授权,称为父节点对节点 i 的多维授权.

定义 4. 原子操作(atomic operation,简称 AO)是指操作实体拥有的某项授权操作对动态数据文件的一次实施过程,$AO_i\cap AO_j=\varnothing,i\neq j$.

定义 5. 操作实体及其后继节点执行原子操作产生的动态数据文件的集合,称为该操作实体的域(domain),域中的动态数据文件称为该域的对象(object);操作实体的操作授权范围覆盖自身的域;操作实体的域可以包含其后继操作实体的域,被包含的域称为子域(sub domain,简称 SD);域所包含的子域和对象统称为该域的成员(member),子域所包含的成员,称为间接成员.

定义 6. 操作实体受到攻击导致其授权操作的动态数据均不可靠,则该操作实体域中的对象均不可靠,称为

该操作实体的失效覆盖集(failure coverage set,简称FCS).除创始节点外的各节点均处于操作实体的多重失效覆盖集.

实际情况下,操作实体均存在由恶意攻击导致的数据篡改或伪造的可能性,本文对动态数据存储面临的安全威胁问题进行如下假设.

(1) 每个攻击者单独到来,相互独立;

(2) 在时间[0,t]内,系统受到的攻击数量{$N(t),t\geqslant 0$}满足参数为λ的泊松分布;

(3) 操作实体第 i 次受到攻击的损失为 L_i,且损失随时间按负指数衰减,损失可累加;

(4) 每次攻击到达的时间间隔和造成的破坏相互独立.

t=0时,损失为 L;$t=t_i(t_i>0)$时,损失为 $Le^{-\alpha t_i}$.设{$L_i,i\geqslant 1$}独立同分布,且与{$N(t),t\geqslant 0$}独立,那么时间[0,t]内不考虑失效覆盖的损失表示为

$$\xi(t)=\sum_{i=1}^{N(t)}L_i\mathrm{e}^{-\alpha(t-t_i)} \tag{1}$$

条件期望为

$$\begin{aligned}E\{\xi(t)\mid N(t)=n\}&=E\left\{\sum_{i=1}^{N(t)}L_i\mathrm{e}^{-\alpha(t-t_i)}\middle| N(t)=n\right\}\\&=E\left\{\sum_{i=1}^{n}L_i\mathrm{e}^{-\alpha(t-t_i)}\middle| N(t)=n\right\}\\&=\sum_{i=1}^{n}E\{L_i\mid N(t)=n\}E\{\mathrm{e}^{-\alpha(t-t_i)}\mid N(t)=n\}\\&=EL\mathrm{e}^{-\alpha t}\sum_{i=1}^{n}E\{\mathrm{e}^{\alpha t_i}\mid N(t)=n\}\end{aligned} \tag{2}$$

记 $Y_1,\ldots,Y_n$ 为[0,t]上独立同均匀分布的随机变量,有:

$$E\left\{\sum_{i=1}^{n}\mathrm{e}^{\alpha t_i}\middle| N(t)=n\right\}=E\left\{\sum_{i=1}^{n}\mathrm{e}^{\alpha Y_i}\right\}=n\int_0^t\mathrm{e}^{\alpha x}\frac{\mathrm{d}x}{t}=\left(\frac{n}{\alpha t}\right)\cdot(\mathrm{e}^{\alpha t}-1).$$

所以:

$$E\{\xi(t)\mid N(t)=n\}=\left(\frac{n}{\alpha t}\right)\cdot(1-\mathrm{e}^{-\alpha t})EL.$$

即:

$$E\{\xi(t)\mid N(t)\}=\left(\frac{N(t)}{\alpha t}\right)\cdot(1-\mathrm{e}^{-\alpha t})EL.$$

因此:

$$E\{\xi(t)\}=E\{E(\xi(t)\mid N(t))\}=\left(\frac{\lambda EL}{\alpha}\right)\cdot(1-\mathrm{e}^{-\alpha t}) \tag{3}$$

考虑失效覆盖的情况,用 $FCS(i)$表示节点 i 的失效覆盖范围,本文优化目标为

$$\min E\{\xi'(t)\}==\left(\frac{\sum_{i=1}^{N(t)}FCS(i)\cdot\lambda EL}{\alpha\cdot N(t)}\right)\cdot(1-\mathrm{e}^{-\alpha t}),t\in[t_1,t_2] \tag{4}$$

3 基于联盟链的改进型动态数据授权操作机制

3.1 区块链

学术界对区块链技术并没有统一的定义,但一般认为,区块链是一种按照时间顺序将数据区块以链条的方

式组合形成的特定数据结构,并以密码学方式保证其不可篡改和不可伪造的去中心化、去信任的分布式共享总账系统[29].区块链的提出是计算机科学的一个突破,它有望降低个人和组织建立、维护信任的成本[30],让没有信任关系的人们在无中心化信任机构的情况下合作[31].自 2008 年 Nakamoto 发表奠基性论文[32]以来,经过近年的快速发展,区块链技术越来越受到政府、银行及相关研究人员的重视.世界经济论坛(world economic forum,简称 WEF)于 2016 年 8 月发布了研究报告[33],区块链成为当前技术研究的热点,包括对区块链协议的分析[34−36]、区块链技术在某些领域的应用等[37−39].

区块链可以分为 3 类:公共链、联盟链和私有链.公共链对外公开,用户不用注册就能匿名自由出入网络,无需授权即可访问网络和区块链,如比特网[32]和以太坊[40];联盟链仅限于联盟成员参与,区块链上的读写权限、参与记账权按联盟规则来制订,如由多家银行参与的区块链联盟 R3[41]、Linux 基金会支持的超级账本项目[42]都属于联盟链架构;私有链仅在私有组织使用,区块链上的读写权限、参与记账权限按私有组织规则来制订.

联盟链可以根据应用场景来决定对公众的开放程度,其网络由成员机构共同维护,节点通过成员机构的网关节点接入,因此适用于物联网行业背景下多成员机构对动态数据的存储、管理、授权、监控和审计.在实际物联网应用背景下,用户、资源、服务、终端存在泛在接入与授权操作的特点,参与的多方实体存在一定的信任前提和利益约束,实体间数据操作共识激励机制和分布式账本记账权确定等问题还需要进一步研究.

3.2 达成共识的边界条件

分布式共识是构建基于区块链技术零信任动态数据溯源机制必须解决的关键问题,而达成共识的条件在公开匿名场景下和带权限管理的场景下需求差异较大[43,44].例如,比特币等金融系统在决策权高度分散的去中心化系统中采用经济激励机制,使各节点高效地针对区块数据的有效性达成共识,该方式面向公有链中的任意节点的自由加入简单有效.而动态数据常常是与特定工作过程联系紧密的行业内部数据,对动态数据的管理更适合采用联盟链方式,仅允许核准的节点加入,货币体系背景下共识激励显然不适用于联盟链方式下对动态数据的管理.在联盟链方式下,参与多方存在一定的信任前提和利益约束,本节通过分析该群体博弈过程中单个节点进行决策的诚实行为动机,提出特定行业背景下分布式节点合作的本质——使各节点在与环境的交互与分布式计算过程中获得最大的累积效用,进一步分析动态数据可追溯系统中各节点达成共识的边界条件,优化共识算法设计.

设动态数据可追溯系统中参与区块信息验证的节点集合为有限集,对每个参与区块信息验证的节点 i 有策略空间及收益函数 U_i,即每个参与节点 i 在策略空间 $S_i=(s_1,s_2,\ldots,s_n)$下的冯·诺依曼-摩根斯坦效用为 $U(S_i)$,本文将策略空间 S_i 下节点预期效用 $U(S_i)$作为评价各动作的价值函数.

每个参与节点的目标是最大化自己的收益,因此为了简化问题,除节点 i 以外的所有其他节点标记为“$-i$”.通过分析节点 i 和$-i$ 相互作用并达成具有约束力协议的共识过程,得到节点 i 和$-i$ 收益矩阵见表 1.

Table 1 Yield matrix of nodes i and $-i$

表 1 节点 i 和$-i$ 收益矩阵

Node i	Node $-i$	
	Cooperative(C)	Betray(B)
Cooperative (C)	$P_iCC,P_{-i}CC$	$P_iCB,P_{-i}CB$
Betray (B)	$P_iBC,P_{-i}BC$	$P_iBB,P_{-i}BB$

表 1 中,C 表示某节点合作(cooperative),B 表示背叛(betray),收益函数表达式中第 1 项为对应策略下节点 i 的收益(分别为 P_iCC,P_iBC,P_iCB,P_iBB),第 2 项为对应策略下节点$-i$ 的收益(分别为 $P_{-i}CC,P_{-i}BC,P_{-i}CB,P_{-i}BB$).溯源系统各节点共识模型的构建基于以下前提条件.

(1) 对于节点 i,在各种策略组合下的收益满足:

$$P_iBC>P_iCC>P_iBB>P_iCB \tag{5}$$

式(5)表明,在节点行为不一致的情况下,采取背叛策略的一方可以从牺牲其余节点的合作行为中得到比所有节点均合作时更高的收益;在所有节点均合作,即达成共识能够获得比都背叛更高的收益;一方合作,其余节

点均背叛将会给合作方带来很大损失,或者说导致最低收益.

(2) 节点 i 估计节点–i 背叛的概率为 λ,即节点 i 对节点–i 的信任度为 $1-\lambda$.本文采用联盟链核准加入的方式,节点由相关溯源系统监管机构、社会团体及志愿者构成,在参与溯源信息验证的 n 个节点中,诚实节点占多数(比例相当大),节点–i 发生背叛是指除节点 i 以外的其余节点产生错误共识的情况.由上述分析可知,这种可能性非常小,即 $\lambda\to 0^+$.

(3) 根据条件(2),节点–i 发生背叛的可能性很小,在其采取合作的前提下,若节点 i 采取合作将获得收益 P_iCC;若节点 i 基于投机主义采取不合作策略,其将获得表 1 中短期最大自身收益 P_iBC,但这将导致系统在时间 $[t,t+\Delta t]$ 内识别出节点 i 的背叛,并对其进行惩罚,惩罚代价函数 $P(S_i)$ 用节点信誉 AR_i 表示,将在本文第 3.3 节详细描述.因此,节点 i 发生背叛的总体收益为

$$U_{BC}(S_i)=P_iBC-P(S_i) \tag{6}$$

根据上面的条件进一步分析得出如下结论:节点 i 采取合作或背叛策略取决于当前时刻 t 节点 i 与节点–i 合作所带来的收益期望值 $E[U(S_i)]$ 与节点 i 背叛所带来的收益期望值 $E[U_{BC}(S_i)]$ 的比较,分两种情况:

$$E[U(S_i)]\geqslant E[U_{BC}(S_i)] \tag{7}$$

$$E[U(S_i)]<E[U_{BC}(S_i)] \tag{8}$$

其中,若公式(7)成立,节点 i 采取合作策略;若公式(8)成立,节点 i 将采取背叛策略.

令 $\theta(0<\theta<1)$ 为节点 i 的折扣因子,用来调节当前收益对长期收益的影响.λ 为节点 i 估计节点–i 在一轮验证过程中采取非合作策略的概率,节点 i 采取合作策略时收益的期望值可表示为

$$\left.\begin{aligned}E[U(S_i)]=&\lambda\left(P_iCB+\frac{\theta}{1-\theta}P_iBB\right)+\lambda(1-\lambda)\left(P_iCC+\theta P_iCB+\frac{\theta^2}{1-\theta}P_iBB\right)+\\&\lambda(1-\lambda)^2\left(P_iCC+\theta P_iCC+\theta^2P_iCB+\frac{\theta^3}{1-\theta}P_iBB\right)+...\\\approx&(1-\lambda)[1+(1-\lambda)\theta+(1-\lambda)^2\theta^2+...+(1-\lambda)^n\theta^n]P_iCC+\\&\lambda[1+(1-\lambda)\theta+(1-\lambda)^2\theta^2+...+(1-\lambda)^n\theta^n]P_iCB+\\&\frac{\lambda\theta}{1-\theta}[1+(1-\lambda)\theta+(1-\lambda)^2\theta^2+...+(1-\lambda)^n\theta^n]P_iBB\end{aligned}\right\} \tag{9}$$

由公式(9)得:

$$E[U(S_i)]=\frac{\lambda}{1-(1-\lambda)\theta}\left(\frac{1-\lambda}{\lambda}P_iCC+P_iCB+\frac{\theta}{1-\theta}P_iBB\right) \tag{10}$$

推导过程与上文类似,节点 i 采取背叛策略时收益的期望值可表示为

$$E[U_{BC}(S_i)]=(1-\lambda)P_iBC+\frac{1}{1-\theta}[(1-\lambda)\theta+\lambda]P_iBB \tag{11}$$

由公式(7)、公式(10)、公式(11)得节点 i 采取合作策略的条件为

$$\frac{\lambda}{1-(1-\lambda)\theta}\left(\frac{1-\lambda}{\lambda}P_iCC+P_iCB+\frac{\theta}{1-\theta}P_iBB\right)\geqslant(1-\lambda)P_iBC+\frac{1}{1-\theta}[(1-\lambda)\theta+\lambda]P_iBB \tag{12}$$

由公式(12)得:

$$\frac{P_iBC-P_iCC}{(P_iBC-P_iBB)(1-\lambda)}+\frac{\lambda(P_iBB-P_iCB)}{(P_iBC-P_iBB)(1-\lambda)^2}<\theta \tag{13}$$

公式(13)即为动态数据可追溯系统中各节点达成共识的边界条件.上式中,θ 是节点 i 长期收益的折扣因子.θ 越大,表明相较当前收益,长期收益对节点 i 的影响越大;θ 越小,表明长期收益对节点 i 的影响越小.在给定节点行为收益的前提下,不等式左边的值取决于系数 $\alpha_1=1/(1-\lambda)$ 和 $\alpha_2=\lambda/(1-\lambda)^2$,只有当不等式右边折扣因子 θ 超过一定值时,公式(13)才成立.假设 θ 为常数,选择非合作策略将导致公式(13)左边的值变大;反之亦成立.因此,为了使节点选择合作行为,必须降低不等式左边的值,即降低系数 $\alpha_1=1/(1-\lambda)$ 和 $\alpha_2=\lambda/(1-\lambda)^2$.得出如下结论.

(1) 节点间相互信任是进行合作的必要非充分条件.公式(13)中,若$\lambda\to1^-$,即节点间几乎不存在信任,不等式左边的取值 $F(\lambda)\to+\infty$,节点间不可能产生合作,因此,节点间相互信任是进行合作的必要条件;若$\lambda=0$,不等式左边的值 $F(\lambda)=(P_iBC-P_iCC)/(P_iBC-P_iBB),\forall\theta\in(0,1)$,公式(13)为非重言式的可满足式,因此,节点间信任不是产生合作的充分条件;

(2) 在本文第 3.3 节提出的共识算法中,机构优先选择估计节点–i 背叛概率λ较低的节点 i 作为验证节点列表(verification nodes list,简称 VNL)中的节点.随着λ的减少,节点 i 对节点–i 的信任度将增大,公式(13)中系数α_1和α_2将减少,进而不等式左边的值 $F(\lambda)$下降,节点 i 选择合作策略的可能性增大.

3.3 基于验证节点列表的共识算法

通过研究使得共识终端最大化自身收益的局部行为与保障动态数据存储安全性和有效性整体目标的关系得出:当所有终端都持有待提交验证的区块,为了让自己的收益最大,任何一方都不会(或者无法)改变自己对其他区块的验证结果.

根据本文第 3.2 节达成共识的边界条件,提出基于信誉的共识激励机制:通过授权一部分信任节点组成一个验证节点列表 VNL,$T_{VNL}=\{T_1,T_2,\dots,T_n\}$,$\forall T_i\in T_{VNL}$,初始状态下,节点信誉 $AR_i=1$,每个节点通过为其他节点服务保持信誉,每轮共识选取最佳区块打包验证节点的同时,以系数γ降低最坏区块打包验证节点的信誉,即 $AR_i=\gamma AR_i(0<\gamma<1)$.为了阻止自私行为并鼓励节点保持其信誉,当 VNL 列表中验证节点信誉低于某一阈值 w 时,将该节点移出 VNL 列表,当超过 1/3 验证节点被移出,则必须由授权机构重新授权组成新的 VNL 列表.假设信誉阈值是全局的,即所有节点使用相同的值,关于特殊情况下某些节点定义局部阈值方面的问题,还有待进一步研究.

每个参与验证的节点会获取在共识开始之前未被记录的所有有效操作,并且以候选集的形式公开他们.然后,每个参与验证的节点合并 VNL 中所有其他验证节点的候选集合,并对所有操作的真实性进行比对投票.对动态数据的有效操作分为两种情况:一是新数据的发布;二是动态数据在不同实体间的流转.上述两种操作都必须由通过机构授权的节点来实现,且均看做一次交易:新数据的发布可以没有输入,但必须有输出,拥有与输出地址公钥对应私钥的节点即为可对该地址数据进行有效操作的授权节点;动态数据在不同实体间的流转既要有输入,又要有输出,其输入需要通过上一笔输出地址所对应的私钥进行签名,验证当前节点是否为授权节点.通过行业顶层管理机构预先颁发根 CA 证书(certificate authority),构建基于根 CA 及中间层 CA 到最底层实体 CA 的完整的证书信任链来实现上述信任基础.系统中全节点服务器负责维护 VNL 列表,验证节点在达成共识时只考虑 VNL 中成员的验证结果完成区块生成,这种共识算法在保证安全性的同时,大幅提高了系统达成共识的效率.同时,由于验证节点是机构授权的节点,一旦其中出现背叛节点便于系统核实身份并追究责任.

区块共识过程的数学形式描述如下.

在动态数据存储系统中,$T_{VNL}=\{T_1,T_2,\dots,T_n\}$为系统中验证节点集合.验证节点 T_i 获取的待验证有效操作候选集记为$\chi(T_i)$,合并后的待验证候选集为

$$\chi(T_{VNL})=\sum_{i=1}^{n}\chi(T_i) \tag{14}$$

某终端 $T_i\in T_{VNL}$ 提交的打包区块 $B_{newi}=\{tx_1,\dots,tx_m\}$,$tx_j\in\chi(T_{VNL})$,获得其他终端验证组合及其收益用集合 $G_i=\{\eta_{i1},\dots,\eta_{in}:u_i\}$表示.由某个终端 T_i 进行打包的区块 B_{newi} 组成的各终端验证组合$(\eta_{i1},\dots,\eta_{in})$中,任意参与验证方 T_k 对 T_i 提交区块 B_{newi} 的验证结果表示为η_{ik},且满足:

$$\eta_{ik}=\begin{cases}1, & \forall tx_j(tx_j\in B_{newi}\wedge tx_j\in\chi(T_{VNL})),\text{经}T_k\text{验证},T_i\text{提交区块合法}\\ -1, & \exists tx_j(tx_j\notin B_{newi}\vee tx_j\notin\chi(T_{VNL})),\text{经}T_k\text{验证},T_i\text{提交区块不合法}\end{cases} \tag{15}$$

$$T_k\xrightarrow{\eta_{ik}=1}T_i:u_i=u_i+1 \tag{16}$$

$$T_k\xrightarrow{\eta_{ik}=-1}T_i:u_i=u_i-1 \tag{17}$$

根据物联网系统的规模及对吞吐率的要求,为共识过程设置合适的等时间间隔轮,用 r 表示.在一轮时间内,达成共识的步骤为:

步骤 1. *VNL* 列表验证节点数大于 2*n*/3,则执行步骤 2;否则,等待授权机构授权新列表;

步骤 2. 本轮时间未结束,对于最早出现的 $T_i \in T_{VNL}$,且使 $u_i(\eta_{i1},\ldots,\eta_{ij},\ldots,\eta_{in})=n$,则选取 T_i 打包区块为本轮最佳区块,即选取最早通过验证节点列表终端验证的区块,转步骤 5;否则,执行步骤 3;

步骤 3. 本轮时间结束,$\exists T_i,\forall T_j \in T_{VNL}$,使得 $n>u_i(\eta_{i1},\ldots,\eta_{ij},\ldots,\eta_{in})>u_j(\eta_{j1},\ldots,\eta_{ji},\ldots,\eta_{jn})$,则选取 T_i 打包区块为本轮最佳区块,即选取经验证节点列表终端验证获得最大收益的区块,转步骤 5;否则,执行步骤 4;

步骤 4. 本轮时间结束,$\exists T_i,T_j,\forall T_k \in T_{VNL}$,若 $n>u_i(\eta_{i1},\ldots,\eta_{ij},\ldots,\eta_{in})=u_j(\eta_{j1},\ldots,\eta_{ji},\ldots,\eta_{jn})>u_k(S_{k1},\ldots,S_{kj},\ldots,S_{kn})$,则从 T_i,T_j 中选取最早达到 u_i 当前值的验证节点打包区块为最佳区块,即选取最早经验证节点列表终端验证获得最大收益的区块;

步骤 5. 本轮时间结束,$\exists T_i,\forall T_j \in T_{VNL}(j\neq i)$,若 $u_i(\eta_{i1},\ldots,\eta_{ij},\ldots,\eta_{in})<u_j(\eta_{j1},\ldots,\eta_{ji},\ldots,\eta_{jn})$,则选取 T_i 打包区块为本轮最坏区块,并执行 $AR_i=\gamma AR_i(0<\gamma<1)$,以降低该节点的信誉.设置信誉阈值为 w,若 $AR_i<w$,则认为验证节点 T_i 不可靠,将其移出 *VNL* 列表.

3.4 动态数据操作与存储

动态数据操作与存储体系采用静态多维授权关系链和动态数据存储链双联盟链模式,实现数据操作授权关系和动态数据本身的防篡改.对应的所有权的转移过程可看做文献[45]描述的所有权转移.

依据本文第 2 节提出的操作实体间授权方式将各实体对动态数据的授权及操作类型作为交易发布到授权关系联盟链网络上,不同的应用场景下可以选择以明文或密文方式发送.根据整个行业物联网操作实体间的合作关系,操作授权往往只需限定在较小的子系统内,涉及的操作实体节点数目较少;此外,由于物联网系统具有可靠性高和生存期长的特点,操作实体间授权关系相对稳定,变更较少.鉴于上述调研现状,本文采取静态方式为各子系统生成实体授权关系,由定义 2 描述的方法创建实体间多维授权链,并作为一笔交易发布到授权关系联盟链网络上.各操作实体数据交付过程如图 3 所示,若存储一个实体节点的编码需 n_c 位,存储用于溯源路径链接的实体节点特征值需 ω 位,实体间授权边的总数为 E,实体总数为 N,每笔交易的数据量上限为 $E\cdot\omega+N\cdot n_c$.当发生操作实体间授权关系变更时,需将授权链作为新的交易在网络上重新发布并记入区块,以便授权追溯.

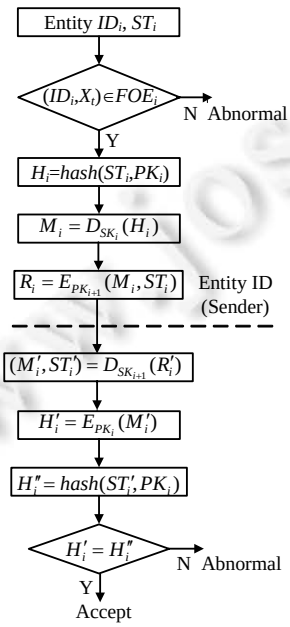


Fig.3 Process of adata delivery between djacent entities

图 3 操作实体间数据交付过程

由密钥分发机构为实例系统中各操作实体 ID_i 生成密钥对(PK_i,SK_i),用于操作实体对动态数据的授权验证.授权节点申请发布新的动态数据或对某个动态数据进行操作时需包含自己的实体证书,经过验证并获得共识的动态数据及相关信息(包括发布方及接收方的地址,动态数据文件的哈希值等)形成发布摘要信息存储到动态数据联盟链网络上.每一轮共识结束后,验证节点会将满足条件的所有交易进行分组哈希运算,将哈希值存储于 Merkle 树状数据结构中,方便实现区块的快速归纳和完整性校验.再利用区块链中的区块生成机制生成数据区块.区块之间利用区块头的哈希指针连接形成链状数据结构.接收方收到动态数据后,在本地计算其哈希值并采用 Merkle 树支持的简化支付验证协议与区块链上的对应数据进行比较,如果不一致,说明文件遭到篡改并向监控中心报警.下面分析图 3 中操作实体对动态数据的授权操作过程.

对于任意操作实体 OE_i,可用 ID_i 唯一标识,为叙述方便,后面也用 ID_i 表示操作实体,简称实体.假设操作实体 ID_i,ID_{i+1} 需要进行的操作 X_t 为:ID_i 采用加密方式发送数据 ST_i,ID_{i+1} 接收数据并对其完整性进行验证.若对完整的动态数据进行签名将导致两方面的缺陷:一方面,存储完整消息对应的数字签名往往需要大量的空间;另一方面,采用非对称加密技术对完整消息进行加密计算开销较大,处理速度较慢.因此,本文在实例系统中相邻层次实体间通信时,采用二次散列迭代的方式,将发送方公钥及消息 ST_i 同时作为哈希函数的输入,得到可作为特征值的哈希运算消息认证码.用发送方的私钥对消息认证码进行签名,由于数据量较少,可保证此运算过程较快.动态数据在授权实体间流转的步骤为:

步骤 1. 判断(ID_i,X_t)$\in FOE_i$,若为 True,证明其为授权节点,执行步骤 2;否则,报错未授权;

步骤 2. 计算 ST_i 和 PK_i 的哈希值 H_i,减少实体 ID_i 签名信息量;

步骤 3. 用实体 ID_i 的私钥 SK_i 对 H_i 进行签名,得到 M_i;

步骤 4. 实体 ID_i 用实体 ID_{i+1} 的公钥对 M_i,ST_i 进行加密并发送;

步骤 5. 实体 ID_{i+1} 用自己的私钥 SK_{i+1} 对接收到的加密信息 R_i' 进行解密,得到 ST_i',M_i';

步骤 6. 实体 ID_{i+1} 用实体 ID_i 的公钥验证签名 M_i',得到 H_i';

步骤 7. 对 ST_i' 和 PK_i 进行哈希运算,得到 H_i'';

步骤 8. 判断 H_i' 与 H_i'' 是否相等:若相等,接收数据;不相等,报错数据异常.

采用与比特币系统类似的方法,将操作实体对动态数据的一次处理过程看做一笔交易,操作实例系统的区块形成过程可描述为:各个操作实体的帐户名为其公钥的哈希值,使用自己的私钥对验证过的信息进行签名.新交易 TX 创建过程由文献[45]定义,TX 通过 P2P 网络进行广播,区块链中各节点都不断地监听网络并收集尚未进入块链的交易 TX 的列表,生成待验证区块,各节点对接收到的区块进行验证,判断区块中是否存在无效交易,即没有正确签名或重复交易.将验证结果再次通过 P2P 网络进行广播,并按照本文第 3.3 节提出的共识机制选出本轮获得共识的区块作为新生区块,通过与前一区块头部链接写入账本.此时,新账本为系统中最长区块链,获得记账权的节点将新创建的区块链向全网广播,其他节点收到后,将其与本地区块链进行比较:若长度大于本地区块链,则将本地区块链更新.

假设创世区块存在且新生区块 B 非空,区块有效性验证算法见算法 1.

算法 1. 区块有效性验证算法.

输入:区块链 C,新生成区块 B;

输出:若创世区块不存在或新区块 B 不存在,返回错误提示 $Error$;若新区块 B 合法,返回加入新区块后的区块链 C;若 B 非法,返回 B.

Function *validate_block*(C,B)

$B\leftarrow V(x_c)$

If $C\wedge(B\neq\varepsilon)$ **then**

do {

$\langle Num,Type,Code,Len,S\rangle\leftarrow(TX_i\in M)$

i++

} **While** ($ValidTX(Type_i,Code_i,Len_i)\wedge(S_i=1)\wedge(i<\mathrm{Max}(B))$)

If ($\Phi^{\Gamma}(n,m)=1)\wedge(i<\mathrm{Max}(B))\wedge(\rho==1)$ **then**

$C^{\perp}\leftarrow B|h(tail(C))$

Else

$B\leftarrow$False

Return (B)

End if

Return (C)

Else

Return (*Error*)

End if

End function

其中,函数 $v(x)$收容当前交易并将其打包成区块 B,若创世区块不存在(C=False)或新区块 B 不存在($B=\varepsilon$),返回错误提示 *Error*;若 B 非空且存在创世区块,则依据五元组$\langle Num,Type,Code,Len,S\rangle$定义的方法对区块 B 中交易进行验证,在区块链诚实节点为多数的前提下,若区块 B 中所有交易 *TX* 均通过验证且获得本轮共识,则将 B 作为新生区块链接到区块链 C 的末尾,返回新生成的当前最长区块链 C;否则,将 B 标记为 False 并返回.

4 分析及实验

本文提出动态数据授权操作机制和基于信任节点列表的区块链共识算法,通过机构授权决定记账节点,提供不可篡改且能够在任何时间点恢复的数据库服务,是一种高效的共识机制.下面对其性能进行分析.

4.1 可靠性分析

假设操作实体总数为 N,其每个实体编码占用 n_c 个比特,根据定义 2,得到各实体的操作授权集合 $\{l_i\}_{i\in\{0,1\}^{\leqslant\omega}}$,对来自诚实或恶意用户的操作请求均需在多维 DAG 图中至少回溯查询 q 次才能获得确认,用 τ 表示从当前提出操作请求的实体向根实体回溯路径上的节点集合,即:

$$\begin{cases}\tau=\{i\parallel parents(i)\}_{i\leqslant p}\\ i=\{i\mid l_{ij}=H(SK_{p_j},l_{p_j},ID_i),where(p_j)=parents(i),j<\delta\}\\ i=1,\ \text{where } l_i=\varepsilon\end{cases} \tag{18}$$

例如,某用户提供自己的授权类型 l 申请对图 1 中 v_7 进行操作(为了简化问题,操作的类型暂不讨论),存在多条回溯路径:$\tau=\{v_7,v_6,v_3,v_2,v_1\}$,$q=5$;或 $\tau=\{v_7,v_6,v_2,v_1\}$,$q=4$;或 $\tau=\{v_7,v_4,v_2,v_1\}$,$q=4$.由此可见,回溯路径 τ 不唯一,下面证明其具备性质 1.

性质 1. 理想化哈希函数表示为 $H:\{0,1\}^{*}\rightarrow\{0,1\}^{\omega}$,$\omega$为哈希后得到的特征值长度,操作实体至少在多维 DAG 图中查询 q 次才能获得确认($q\leqslant N$),攻击者将动态数据操作权限 l 伪造成 l',并获得攻击成功,即 $l\neq l'$,$H(l,ID)=H(l',ID)$的概率上限为 $q^2/2^{\omega+1}$.

证明:第 i 次查询输出时,前 $i-1$ 次查询输出相同的概率至多为$(i-1)/2^{\omega}$,因此,q 次查询输出均相同的概率为

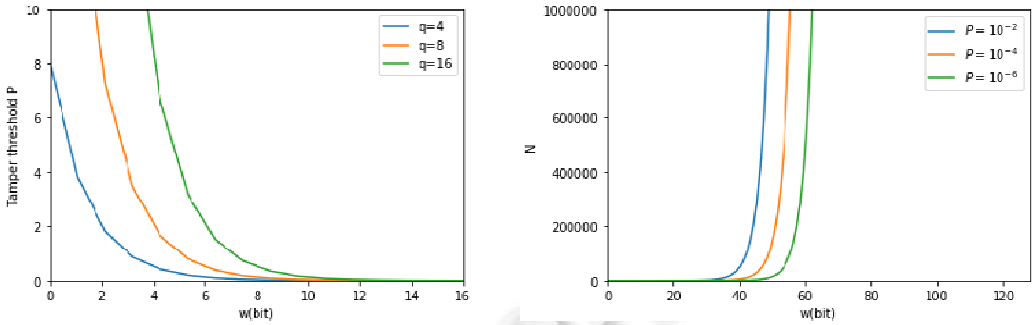
$$\sum_{i=1}^{q}(i-1)\Big/2^{\omega}<q^2/2^{\omega+1}. \qquad \square$$

在多维 DAG 图中,由公式(18)计算 τ,并回溯至 τ 中根节点 r(即该节点无父节点),按照定义 2 重新计算授权路径上各实体授权特征值 l_i',与联盟链上经过 *VNL* 验证的实体授权特征值的保存值 l_i 进行比较,若满足:

$$\{l_i'=l_i\}_{\forall i\in\tau,\text{且}q_i'\geqslant q_i} \tag{19}$$

则该请求合法;否则,拒绝请求.公式(19)中,q_i 表示为满足安全系数,实体 OE_i 设定的查询次数,q_i' 表示实际执行的查询系数.若某中间层实体撤销对其子节点的授权会造成 $q_i'<q_i$,此时应拒绝该实体的操作.

由性质 1,参考比特网,给定取值ω=256,q=6 时,攻击者篡改动态数据操作权限并获取成功的概率非常小,理论上存在,但实际很难做到.图 4 对不同物联网规模下本文方案的系统可靠性进行分析.



(a) 局部可靠性 (b) 全局可靠性

Fig.4 System reliability analysis

图 4 系统可靠性分析

当物联网子系统规模较小,操作实体授权链深度较小,回溯查询的次数 q 在较小范围内变动(如取值 4,8,16),授权特征值位数|ω|=16b 时,攻击者成功篡改授权关系的概率 P 趋近 0,如图 4(a)所示;当物联网规模较大,需要更多的比特位对授权特征值ω进行编码,当|ω|=64b,物联网规模 N=10^6 时,攻击者成功篡改授权关系的概率 P<10^{-6},如图 4(b)所示.

4.2 安全性分析

鉴于比特币网络中经常出现双重输出攻击、重放攻击和隐藏攻击,下面对本文提出的方案在上述 3 种常见攻击类型下的性能进行分析.

(1) 双重输出攻击

双重输出攻击是指操作实体隐藏对其他操作实体授权的动态数据文件及授权关系.与比特币类似,可以通过创建两个不同的交易分支来分割自己的区块链.本文提出的动态数据溯源架构对双重输出攻击具有防御能力,违规者会被发现并失去他人的信任.

图 5 说明了本方案对这种攻击的防御机制,描述了溯源架构中某操作实体执行双重输出攻击的过程(注:符号 C 本段局部定义为操作实体).

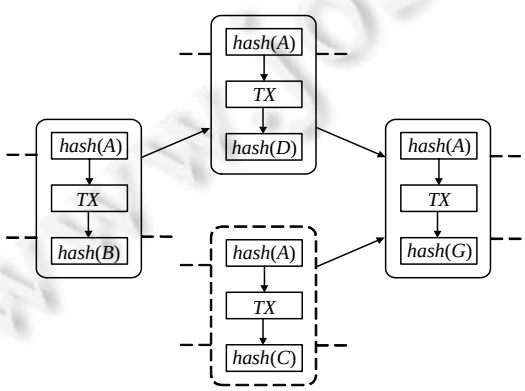


Fig.5 Double output attack

图 5 双重输出攻击

在这种情况下,操作实体*A*希望隐藏虚线块,即:对操作实体*C*授权的动态数据文件及授权关系,只传播关于他对操作实体*D*授权的动态数据文件及授权关系.虽然这种攻击看起来似乎成功,但是当操作实体*C*除*A*以外的授权实体,比如*B*,查看操作实体*C*的历史记录时,会发现在*C*链的验证期间*A*隐藏了一笔对*C*授权的交易.这与*B*关于操作实体*A*所涉及的交易的知识相矛盾.这样,操作实体*A*创建的两个区块形成了一个欺诈证据,并被其他节点在网络中进行广播.其他操作实体可以使用上述方法以较小计算量来验证双重输出攻击行为,将其列入黑名单或拒绝服务.

(2) 重放攻击

重放攻击试图重复使用由某个操作实体创建的动态数据文件及授权关系签名重放已经发生的交易.恶意操作实体将指针重用到另一操作实体的先前块.图 6 说明了本方案对重放攻击的抵御机制.操作实体*A*使用两次相同的事务增长其块链,这种攻击背后的动机是:恶意操作实体隐藏动态数据的时间属性,达到对物联网控制系统的某种破坏.这种攻击相对容易发现:当由另一个操作实体验证操作实体*A*的事务链正确性时,将检测出有两个块具有相同的输出指针.恶意操作实体在重放攻击期间创建的块组成欺诈证据,网络中的任何操作实体都可以通过观察块的输出指针来验证欺诈.

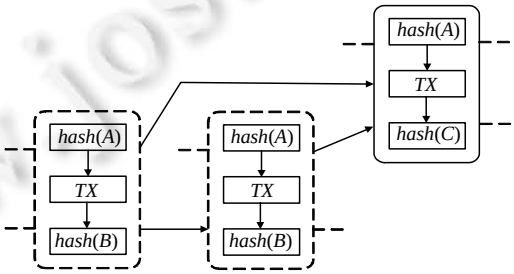


Fig.6 Replay attack

图 6 重放攻击

(3) 隐藏攻击

一旦操作实体对动态数据进行操作,就会在网络中创建记录.一个操作实体可能只想公开对其声誉有正面影响的操作或授权,同时隐藏对其声誉有负面影响的操作或授权.本文提出的溯源链架构能够抵御这种攻击:由于每条记录都包含一个序列号,网络中的任何人都可以请求其他人的特定记录,如果某操作实体无法提供自身的历史记录,则在该实体所要求的记录被提供并被验证之前,其他实体可以选择不与这个操作实体发生交易.值得注意的是,操作实体不能防止其授权对象授权给其他操作实体.

4.3 部署和实验

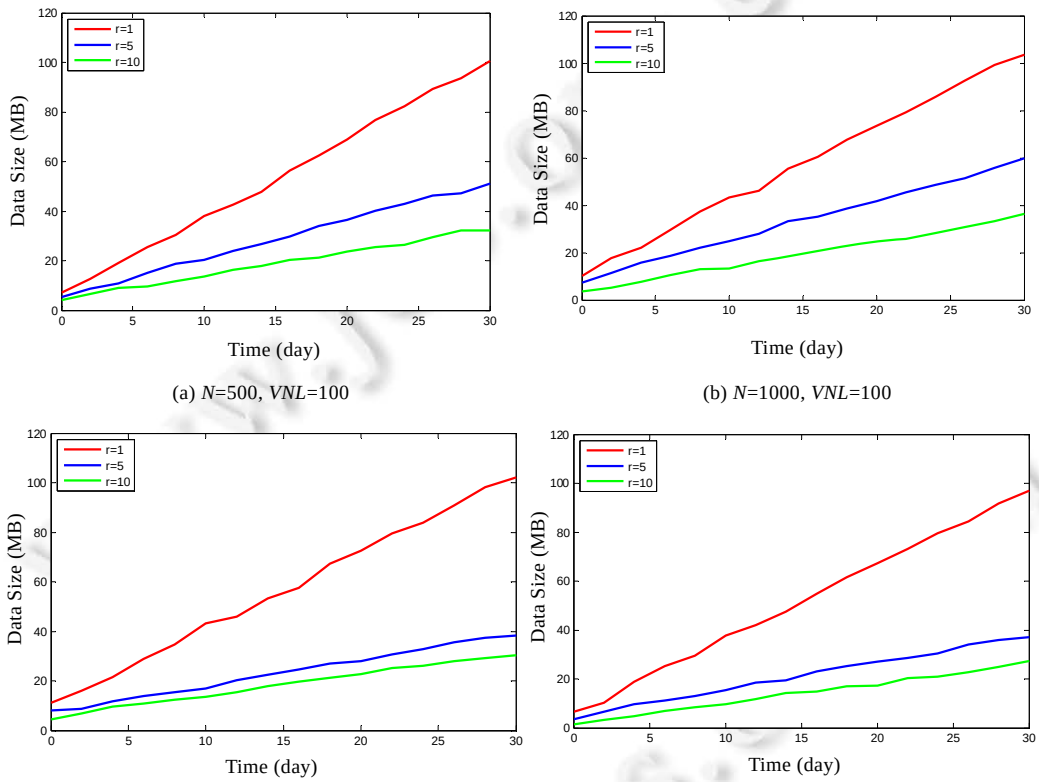
本文进行了一项公开实验,对物联网动态数据溯源机制性能进行评估,从互联网招募的 1210 名志愿者参加了为期一个月的开放式研究.这些志愿者在 Ubuntu 16.04 下安装并使用了 ChainSQL 平台.该平台是在瑞波币的基础上改进的,是我们和众享比特公司长期合作进行的基于区块链的物联网应用安全研究的一部分.修改配置文件,使用不同的配置文件启动应用程序,得到普通节点和验证节点.随机生成初始化测试数据集合,不同轮次的测评实施需基于相同的测试数据以确保测试结果的有效性.测试数据作为新交易相继提交到 ChainSQL 测试网络,采用多组交易并发激励的方式,以测试较高并发交易场景下系统的性能.

图 7 给出了操作实体总数 N,验证节点列表中节点数目为 VNL,ω=256,n_c=16,达成共识的每轮时间 r 取不同值时 ChainSQL 平台溯源数据增长情况.

图 7(a)中取 N=500,VNL=100,r=5s 时,数据量高于 r=10s,但其数据量的增幅不到 r=10s 时数据量的两倍,说明相比 r=5s 的情况,r=10s 时部分轮次在本文第 3.3 节提出的共识机制的步骤 2 完成;同理,r=1s 与 r=5s,r=10s 时相比数据量有显著增加,但增幅低于相应的时间倍值.

图 7(b)中取 N=1000,VNL=100,相比图 7(a),操作实体数目多了 1 倍,但 r 取 3 种不同值时数据量同比均有所下降,r=5s 和 r=10s 时下降程度较 r=1s 时大.这说明随着系统规模的增大,达成共识的速度减慢,且在共识机制步骤 2 完成的轮次受影响较大,在共识机制步骤 3 和步骤 4 完成的轮次受影响较小.

图 7(d)中取 N=1000,VNL=200,相比图 7(b),机构授权的验证列表节点数目多了 1 倍,r 取 3 种不同值时达成共识的速度均有所减慢,因此数据量同比均有所下降;相比图 7(c),操作实体数目多了 1 倍,数据量增长同比变化差异不大,这说明机构授权的验证列表节点数目增多将导致共识时间增加,使得大部分轮次均在共识机制步骤 3 或步骤 4 完成,此时,系统溯源链中数据量的增幅对操作实体数目的增多呈现一定程度的鲁棒性.



(a) N=500, VNL=100　　(b) N=1000, VNL=100

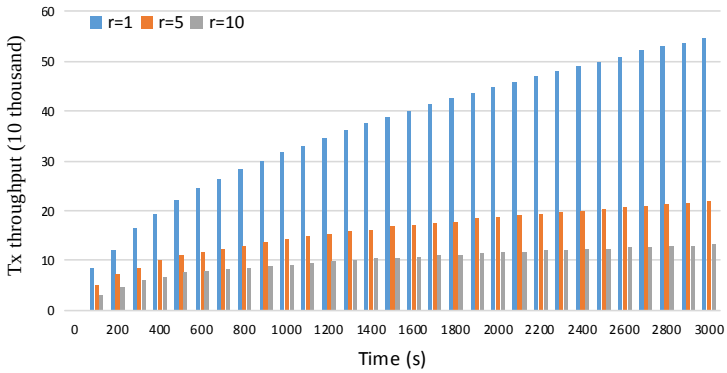
(c) N=500, VNL=200　　(d) N=1000, VNL=200

Fig.7　Traceability data growth comparison on ChainSQL platform

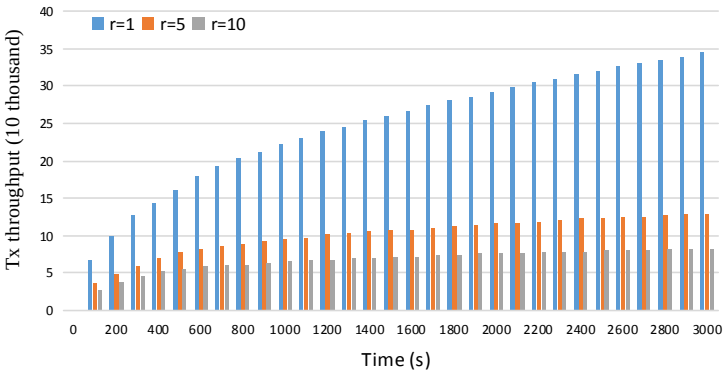
图 7　ChainSQL 平台溯源数据增长对比

图 8 显示了 N=1000,VNL=100,r 取不同值时,在联想 T480S 和 Y450A-TSI(W)型号移动终端上,3000s 时间区间内,系统吞吐量随时间的变化情况.从总体上看,r 取值越小,系统吞吐量越大.r=1s 时,在性能较高的移动设备联想 T480S 上,3000s 内吞吐量均值为 182tps;在性能较低的移动设备 Y450A-TSI(W)上,3000s 内吞吐量均值为 115tps.通过观察图 8(a)和图 8(b)在不同性能的移动设备上同一时间区间和网络环境下的吞吐量数据对比,可以看出系统吞吐量受硬件性能的影响.结果表明:即使在性能较差的移动终端上,也可以实现对大量轻型交易的创建和处理,这些测试数据为我们进一步开发嵌入式系统 ChainSQL 接口提供了参考.假定不论 r 取何值,均在每轮时间用完后形成共识并产生新区块,显然,r=5s 和 r=10s 时,在 3000s 时间内产生新区块的数目将分别是 r=1s 时的 1/5 和 1/10.图 8(a)和图 8(b)的数据表明:r=5s 和 r=10s 时,对应的吞吐量均值均大于 r=1s 时相应的倍值.这是因为 r=1s 时,各节点将接收到并存储在本地的交易广播出去,大部分在接近或等于 r 时形成共识;r=5s 和 r=10s 时,会有一部分节点提交的区块在该轮时间尚未用完就获得共识.通过分析图 8 测试数据发现,几种情况下系统

吞吐量的变化存在共同点:测试初始阶段,吞吐量较大;随着测试时间增加,吞吐量变小.对于这一现象的合理解释是:测试初始阶段,节点本地存储器为空,生成的新交易区块获得共识后,在 ChainSQL 数据库中快速插入;随着数据库的增长,每次插入新的区块都需要查询和同步数据库来获取最新联盟链区块的信息,插入开销有所增加,系统吞吐量随着数据库大小的增加而减少.



(a) 移动终端 T480S 下系统吞吐量



(b) 移动终端 Y450A-TSI(W)下系统吞吐量

Fig.8 TX throughput on diffenent devices

图 8 不同移动终端交易吞吐量对比

通过为期一个月的由志愿者参与的实验,证明了本文提出的溯源机制的实际适用性和成熟度水平.结果表明:本文提出的物联网动态数据溯源机制在没有任何中心数据库的情况下,能够在网络上以共识方式保证各操作实体数据的完整性和可靠性;且随着系统中机构授权的验证列表节点数目和操作实体数目增多,系统达成共识的速度对这两项参数呈现鲁棒性,而主要受预先设置的每轮时间影响.

5 总 结

本文针对动态数据安全问题,设计了动态数据安全问题模型,分析达成共识的边界条件,提出了联盟链方式下基于节点信誉的共识激励机制和基于验证节点列表的共识算法,实现了对动态数据授权操作和安全管理.分析了该机制下授权机制的可靠性及系统抵抗双重输出攻击、重放攻击及隐藏攻击的能力等安全性能,通过在 ChainSQL 平台下的公开实验,分析了系统规模、验证节点列表规模、每轮时间与实际达成共识的速度、系统吞吐量的关系,得出随着系统中机构授权的操作实体数目和验证节点列表节点数目增多,系统达成共识的速度对这两项参数呈现鲁棒性,而主要受预先设置的每轮时间影响的结论,证明了核准加入方式下该方案能够有效防御攻击者对动态数据的篡改、伪造等潜在安全威胁,在保证动态数据安全存储的同时,提高了系统达成共识

的效率.目前的研究已经取得了阶段性的成果,但仍存在一定的局限性:首先,本文所讨论的节点信誉阈值是全局的,所有节点被赋予相同的初始阈值,关于特殊情况下某些节点定义局部阈值方面的问题,还有待进一步研究;其次,目前,改进后共识机制的部署和调用都需要专业区块链研发人员进行,在易用性和对应用的支持上还需要进一步验证.

结合物联网行业应用需求和安全技术发展热点,课题组拟进一步开展以下几方面的研究工作.

(1) 智能合约形式化证明与部署.与传统授权协议相比,如基于角色的访问管理协议OAuth 2.0,OpenID等,智能合约能够为物联网设备提供基于单方或多方身份验证和业务逻辑的高效授权访问规则.但智能合约一经部署就不能修改,其漏洞将给物联网系统应用带来巨大损失.因此,部署智能合约前必须对其正确性进行完备的形式化证明;

(2) 轻量级安全通信协议.常见的物联网通信协议 MQTT,CoAP,RPL,6LoWPAN 等,无法提供通信安全性.此类协议必须嵌入在其他安全协议中,如 DTLS,TLS,IPSec 等,以提供安全通信.然而,DTLS,TLS,IPSec 甚至轻量级 TinyTLS 协议对计算和存储资源的性能要求都超出物联网设备的承受能力,因此,迫切需要开发适用于物联网设备的轻量级安全协议,用以保障链下动态数据安全.

References:

[1] Zhu LH, Gao F, Shen M, Li YD, Zheng BK, Mao HL, Wu Z. Survey on privacy preserving techniques for blockchain technology. Journal of Computer Research and Development, 2017,54(10):2170−2186 (in Chinese with English abstract). [doi: 10.7544/issn1000-1239.2017.20170471]

[2] Jongkuk L, Udatta SP, William Q. Supply chain efficiency and security: Coordination for collaborative investment in technology. European Journal of Operational Research, 2011,210(3):568−578. [doi: 10.1016/j.ejor.2010.10.015]

[3] Ning HS, Xu QY. Research on Global Internet of things' developments and it's construction in China. Acta Electronica Sinica, 2010,2(11):2590−2599 (in Chinese with English abstract).

[4] Faheem U, Edwards M, Ramdhany R, Babar MA, Rashid A. Data exfiltration: A review of external attack vectors and countermeasures. Journal of Network and Computer Applications, 2018,101(1):18−54. [doi: 10.1016/j.jnca.2017.10.016]

[5] Wagner J, Rasin A, Glavic B, Heart K, Furst J, Bressan L, Grier J. Carving database storage to detect and trace security breaches. Digital Investigation, 2017,22(8):127−136. [doi: 10.1016/j.diin.2017.06.006]

[6] Khanduja V. Database watermarking, a technological protective measure: Perspective, security analysis and future directions. Journal of Information Security and Applications, 2017,37(12):38−49. [doi: 10.1016/j.jisa.2017.10.001]

[7] Trivedi D, Zavarsky P, Butakov S. Enhancing relational database security by metadata segregation. Procedia Computer Science, 2016,94(1):453−458. [doi: 10.1016/j.procs.2016.08.070]

[8] Wassim EH, Ghassen BB, Hazem H, Haidar S, Ralph A. Security by construction in Web applications development via database annotations. Computers & Security, 2016,59(3):151−165. [doi: 10.1016/j.cose.2015.12.004]

[9] Feng DG, Zhang M, Zhang Y, Xu Z. Study on cloud computing security. Ruan Jian Xue Bao/Journal of Software, 2011,22(1):71−83 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3958.htm [doi: 10.3724/ SP.J.1001.2011.03958]

[10] Wei LF, Zhu HJ, Cao ZF, Dong XL, Jia WW, Chen YL, Vasilakos AV. Security and privacy for storage and computation in cloud computing. Information Sciences, 2014,258(3):371−386. [doi: 10.1016/j.ins.2013.04.028]

[11] Zhang YQ, Zhou W, Peng AN. Survey of Internet of things security. Journal of Computer Research and Development, 2017,54(10):2130−2143 (in Chinese with English abstract). [doi: 10.7544/issn1000-1239.2017.20170470]

[12] Nissim N, Yahalom R, Elovici Y. USB-based attacks. Computers & Security, 2017,70(9):675−688. [doi: 10.1016/j.cose.2017.08.002]

[13] Yashashree D, Manasi N, Sumedha V, Nikita Z. Database security using intrusion detection system. Int'l Journal of Scientific & Engineering Research, 2017,8(2):30−35.

[14] Yang BH, Chen C. Block Chain Principle, Design and Application. Beijing: China Machine Press, 2017. 5−19 (in Chinese).

[15] Huh S, Cho S, Kim S. Managing IoT devices using blockchain platform. In: Proc. of the 19th Int'l Conf. on Advanced Communication Technology (ICACT). IEEE, 2017. 464−467. [doi: 10.23919/ICACT.2017.7890132]

[16] Zhang YQ, Wang XF, Liu XF, Liu L. Survey on cloud computing security. Ruan Jian Xue Bao/Journal of Software, 2016,27(6):1328−1348 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5004.htm [doi: 10.13328/j.cnki.jos.005004]

[17] Yang GY, Yu J, Shen WT, Su QQ, Fu ZJ, Hao R. Enabling public auditing for shared data in cloud storage supporting identity privacy and traceability. Journal of Systems and Software, 2016,113(3):130−139. [doi: 10.1016/j.jss.2015.11.044]

[18] Helpnetsecurity. Top 12 cloud computing threats in 2016. https://www.helpnetsecurity.com/2016/03/01/top-12-cloud-computing-threats-in-2016/

[19] Alfian G, Rhee J, Ahn H, Lee J, Farooq U, Fazalljaz M, Syaekhoni MA. Integration of RFID, wireless sensor networks, and data mining in an e-pedigree food traceability system. Journal of Food Engineering, 2017,212(11):65−75. [doi: 10.1016/j.jfoodeng.2017.05.008]

[20] Gandino F, Montrucchio B, Rebaudengo M. A security protocol for RFID traceability. Int'l Journal of Communication Systems, 2016,30(6):1−14. [doi: 10.1002/dac.3109]

[21] Choi S, Kim H, Lee S, Lee K, Lee H. A fully integrated CMOS security-enhanced passive RFID tag. ETRI Journal, 2014,36(1):141−150. [doi: 10.4218/etrij.14.0112.0674]

[22] Liu TT. Research on key technologies of data security towards cloud computing [Ph.D. Thesis]. Zhengzhou: PLA Information Engineering University, 2013 (in Chinese with English abstract).

[23] He M, Chen GH, Liang WH, Lai HG, Ling C. Cloud data storage security and privacy protection policies under IoT environment. Computer Science, 2012,39(5):62−65,90 (in Chinese with English abstract). [doi: 10.3969/j.issn.1002-137X.2012.05.013]

[24] Li WW, You WX, Wang XP. Survey of cyber security research in power system. Power System Protection and Control, 2011, 39(10):140−147 (in Chinese with English abstract). [doi: 10.1080/17415993.2010.547197]

[25] Eyal I, Gencer AE, Sirer EG, Renesse RV. Bitcoin-NG: A scalable blockchain protocol. In: Proc. of the 13th USENIX Symp. on Networked Systems Design and Implementation. Berkeley: USENIX Association, 2016. 45−59.

[26] Sompolinsky Y, Zohar A. Secure high-rate transaction processing in bitcoin. In: Proc. of the Int'l Conf. on Financial Cryptography and Data Security. Berlin: Springer-Verlag, 2015. 507−527. [doi: 10.1007/978-3-662-47854-7_32]

[27] Lewenberg Y, Sompolinsky Y, Zohar A. Inclusive block chain protocols. In: Proc. of the Int'l Conf. on Financial Cryptography and Data Security. Berlin: Springer-Verlag, 2015. 528−547. [doi: 10.1007/978-3-662-47854-7_33]

[28] Hardjono T, Smith N. Cloud-based commissioning of constrained devices using permissioned blockchains. In: Proc. of the ACM Int'l Workshop on IoT Privacy, Trust, and Security. New York: ACM Press, 2016. 29−36. [doi: 10.1145/2899007.2899012]

[29] Liu AD, Du XH, Wang N, Li SZ. Research progress of blockchain technology and its application in information security. Ruan Jian Xue Bao/Journal of Software, 2018,29(6):2092−2115 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5589.htm [doi: 10.13328/j.cnki.jos.005589]

[30] Iansiti M, Lakhani KR. The truth about blockchain. Harvard Business Review, 2017,95(1):118−127.

[31] Yuan Y, Wang FY. Blockchain: The state of the art and future trends. Acta Automatica Sinica, 2016,42(4):481−494 (in Chinese with English abstract). [doi: 10.16383/j.aas.2016.c160158]

[32] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. In: Proc. of the Consulted. 2008. https://bitcoin.org /bitcoin.pdf

[33] McWaters R, Galaski R, Bruno G, Chatterjee S. The future of financial infrastructure: An ambitious look at how blockchain can reshape financial services. In: Proc. of the World Economic Forum. 2016. 8.

[34] Fan J, Yi LT, Shu JW. Research on the technologies of Byzantine system. Ruan Jian Xue Bao/Journal of Software, 2013,24(6):1346−1360 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4395.htm [doi: 10.3724/SP.J.1001.2013.04395]

[35] Pass R, Seeman L, Shelat A. Analysis of the blockchain protocol in asynchronous networks. In: Proc. of the Advances in Cryptology—EUROCRYPT. Berlin: Springer-Verlag, 2017. 643−673. [doi: 10.1007/978-3-319-56614-6_22]

[36] Garay JA, Kiayias A, Leonardos N. The bitcoin backbone protocol: Analysis and applications. In: Proc. of the Advances in Cryptology—EUROCRYPT (2). Berlin: Springer-Verlag, 2015. 281−310. [doi: 10.1007/978-3-662-46803-6_10]

[37] Nakasumi M. Information sharing for supply chain management based on block chain technology. In: Proc. of the 2017 IEEE 19th Conf. on Business Informatics (CBI). IEEE Computer Society, 2017. 140−149. [doi: 10.1109/CBI.2017.56]

[38] Pinzón C, Rocha C. Double-spend attack models with time advantage for bitcoin. Electronic Notes in Theoretical Computer Science, 2016,329(9):79−103. [doi: 10.1016/j.entcs.2016.12.006]

[39] Sikorski JJ, Haughton J, Kraft M. Blockchain technology in the chemical industry: Machine-to-machine electricity market. Applied Energy, 2017,195(2):234−246. [doi: 10.1016/j.apenergy.2017.03.039]

[40] Wood G. Ethereum: A secure decentralised generalised transaction ledger. Ethereum Project Yellow Paper, 2014,150(1):1−32.

[41] R3CEV. R3. 2017. http://www.r3cev.com/

[42] Hyperledger. Hyperledger. 2016. https://www.hyperledger.org/

[43] Gramoli V. From blockchain consensus back to Byzantine consensus. Future Generation Computer Systems, 2017,9(1):1−20. [doi: 10.1016/j.future.2017.09.023]

[44] Reid F, Harrigan M. An analysis of anonymity in the bitcoin system. In: Proc. of the Security and Privacy in Social Networks. New York: Springer-Verlag, 2013. 197−223. [doi: 10.1007/978-1-4614-4139-7_10]

[45] Qiao R, Dong S, Wei Q, Wang QX. Research on security mechanism of dynamic data storage based on block chain technology. Computer Science, 2018,45(2):55−60 (in Chinese with English abstract). [doi: 10.11896/j.issn.1002-137X.2018.02.010]

附中文参考文献:

[1] 祝烈煌,高峰,沈蒙,李艳东,郑宝昆,毛洪亮,吴震.区块链隐私保护研究综述.计算机研究与发展,2017,54(10):2170−2186. [doi: 10.7544/issn1000-1239.2017.20170471]

[3] 宁焕生,徐群玉.全球物联网发展及中国物联网建设若干思考.电子学报,2010,2(11):2590−2599.

[9] 冯登国,张敏,张妍,徐震.云计算安全研究.软件学报,2011,22(1):71−83. http://www.jos.org.cn/1000-9825/3958.htm [doi: 10.3724/SP.J.1001.2011.03958]

[11] 张玉清,周威,彭安妮.物联网安全综述.计算机研究与发展,2017,54(10):2130−2143. [doi: 10.7544/issn1000-1239.2017.2017 0470]

[14] 杨保华,陈昌.区块链原理、设计与应用.北京:机械工业出版社,2017.

[16] 张玉清,王晓菲,刘雪峰,刘玲.云计算环境安全综述.软件学报,2016,27(6):1328−1348. http://www.jos.org.cn/1000-9825/5004.htm [doi: 10.13328/j.cnki.jos.005004]

[22] 刘婷婷.面向云计算的数据安全保护关键技术研究[博士学位论文].郑州:解放军信息工程大学,2013.

[23] 何明,陈国华,梁文辉,赖海光,凌晨.物联网环境下云数据存储安全及隐私保护策略研究.计算机科学,2012,39(5):62−65,90. [doi: 10.3969/j.issn.1002-137X.2012.05.013]

[24] 李文武,游文霞,王先培.电力系统网络安全研究综述.电力系统保护与控制,2011,39(10):140−147. [doi: 10.1080/17415993.2010.547197]

[29] 刘敖迪,杜学绘,王娜,李少卓.区块链技术及其在信息安全领域的研究进展.软件学报,2018,29(6):2092−2115. http://www.jos.org.cn/1000-9825/5589.htm [doi: 10.13328/j.cnki.jos.005589]

[31] 袁勇,王飞跃.区块链技术发展现状与展望.自动化学报,2016,42(04):481−494. [doi: 10.16383/j.aas.2016.c160158]

[34] 范捷,易乐天,舒继武.拜占庭系统技术研究综述.软件学报,2013,24(6):1346−1360. http://www.jos.org.cn/1000-9825/4395.htm [doi: 10.3724/SP.J.1001.2013.04395]

[45] 乔蕊,董仕,魏强,王清贤.基于区块链技术的动态数据存储安全机制研究.计算机科学,2018,45(2):55−60. [doi: 10.11896/j.issn.1002-137X.2018.02.010]

乔蕊(1983－),女,河南周口人,副教授,CCF 学生会员,主要研究领域为信息安全.

王清贤(1960－),男,教授,博士生导师,CCF 高级会员,主要研究领域为信息安全.

曹琰(1983－),男,博士,讲师,CCF 专业会员,主要研究领域为信息安全.