文章编号:1671-4229(2021)03-0001-08

基于可编辑区块链的数据隐私安全

廖晓峰, 张 迪
(重庆大学 计算机学院,重庆 400044)

摘 要:随着用户对数据隐私安全需求的提升及相关隐私保护条例的制定,擦除匿名、公开、不可变地存储在区块链中的非法数据已成为受害者及国际刑警组织的迫切需求。基于变色龙哈希、共识投票等技术对区块链的历史数据进行编辑,能有效地擦除区块链上的非法数据,保证用户数据的被遗忘权。文章先以比特币交易为例分析区块链上非法数据插入的常见方法;再综合分析和对比已有的3类可编辑区块链方案;针对现有方案容易遭受攻击、编辑权限被集中控制等问题,文章提出了一种基于去中心化变色龙哈希的区块链数据可编辑方案。此外,文章对所提方案进行了详细的安全性分析和性能分析。分析结果表明,所提方案不仅具有抗双花攻击、抗篡改攻击的能力,还能以较低的计算消耗和较高的效率实现区块链历史数据的可编辑。

关键词:可编辑区块链;数据被遗忘权;隐私安全;变色龙哈希
中图分类号:TP 309.2 **文献标志码**:A

Data privacy security based on redactable blockchain

LIAO Xiao-feng, ZHANG Di
(College of Computer Science, Chongqing University, Chongqing 400044, China)

Abstract: With the increasing demand for users' data privacy and the formulation of related privacy protection regulations, erasing the illegal data which is anonymously, publicly, and immutably stored in the blockchain has become an urgent requirement for victims and Interpol. Redacting the history data of the blockchain based on technologies such as chameleon hash and consensus voting can effectively erase the illegal data from the blockchain and guarantee the right to be forgotten of users' data. In this paper, we start with analyzing the harmful data insertion method with Bitcoin transaction taken as an example. Then we comprehensively analyze and compare three types of redactable blockchain schemes. Finally, for some issues of existing schemes such as suffering from attacks and centralized control of redactions, we propose a scheme to redact blockchain data based on a decentralized chameleon hash function. Moreover, we analyze the details of security and performance of the proposed scheme. The analyses show that the proposed scheme not only resists double-spending attacks and tampering attacks but also performs redactions with lower overhead and higher efficiency.

Key words: redactable blockchain; right to be forgotten; privacy security; chameleon hash

目前,区块链技术吸引了学术界和工业界的浓厚研究兴趣,已被广泛应用于各类主流领域。例如,解决版权纠纷[1]、产品溯源[2]、电子投票[3]、存储服务[4]和医疗保健服务[5]等。最近的一项研究显示[6],全球在区块链研究上的支出将在2024年增长到每年约190亿美元。

基金项目:国家重点研发计划资助项目(2018AAA0100101);国家自然科学基金资助项目(61932006, 61772434);重庆市技术创新与应用发展专项面上资助项目(cstc2020jscx-msxmX0156)
作者简介:廖晓峰(1964—),男,教授,博士. E-mail: xfliao@cqu.edu.cn
引文格式:廖晓峰, 张迪. 基于可编辑区块链的数据隐私安全[J]. 广州大学学报(自然科学版),2021,20(3):1-8.

区块链是一个由点对点对等(P2P)网络管理的自治分布式账本,无需假设任何可信的中央机构。它允许每个节点独立地收集、验证交易并产生新区块,由共识机制确保网络节点拥有一致的区块链视图。基于密码学原语,区块链保证了数据的不可篡改和可追溯性,并支持数据的可审计。然而,研究一种技术来支持区块链中的数据可编辑具有重要意义。具体原因如下:①不可变的区块链容易被恶意者用于存储和传播个人隐私数据、涉黄数据和盗取金额的攻击脚本数据等非法数据。例如,Matzutt 等[7]在比特币交易中找到了大量涉黄内容的文件,其中包含至少 274 个指向儿童的色情内容和 142 个涉及暗网服务的链接。区块链系统不应该成为低成本网络犯罪的法外之地;②不可变的区块链无法满足所有新兴的基于区块链的应用需求。如今,许多基于区块链的应用更关注区块链技术的实用性,从而需要一定程度的数据可编辑性[8]。例如,存储在区块链上的数据可能包含用户的医疗保健和保险记录等隐私敏感数据,用户极可能希望从区块链中删除其自身的一些敏感数据并在需要时修订合同服务和及时更新数据。文献[9-10]指出在物联网(IoT)中可能需要删除不必要的或者失效的数据,以节省边缘端的内存消耗;③对可编辑区块链的需求受到了数据隐私保护相关法律法规激励。例如,欧洲通用数据保护条例(GDPR)[11]规定了用户的数据具有“被遗忘权(the right to be forgotten)”,即任何用户都有权擦除其个人数据和副本。综上,在特定情况下对区块链数据进行编辑已成为一个十分重要的需求。

现有的研究基于不同的机制提出了一些可编辑区块链方案,为链上数据的编辑提供某种程度的可行性[12]。文章旨在综合分析现有的可编辑区块链相关工作,包括学术论文、白皮书和正式文件等,再提出一种安全的、去中心化的编辑方案。具体研究内容安排如下:首先,本文以比特币交易为例分析在交易中插入或携带非法数据的常见方法,便于研究者更好地了解区块链数据编辑的挑战及难点;其次,本文分别对 3 类具有代表性的可编辑区块链方案(①基于共识;②基于变色龙哈希;③基于元交易数据。)进行综合性分析,评估潜在漏洞和限制,并在性能和安全性方面进行对比,以便研究者更好地把握未来的研究方向。最后,本文提出了一个去中心化的可编辑区块链方案,并且通过详细的分析表明所提方案有效地改善了一些现有方案的缺陷。

综上所述,本文的主要贡献将包含:①分析在区块链中常用的非法数据插入方案及其危害;②综合地分析 3 类最新的具有代表性的可编辑区块链方案,并在安全性和性能方面进行了对比;③提出基于属性控制的去中心化可编辑区块链方案,解决已有可编辑区块链方案依赖单一可信授权中心,容易遭受攻击等缺陷,并且满足高效等性能需求。

本文的其余部分描述的主要内容如下:第一节以比特币区块链为例,回顾区块链的相关背景知识,分析常用插入非法数据的方法;第二节主要概述和分析 3 类具有代表性的可编辑区块链方案,并对其进行综合分析与对比;第三节介绍所提方案,并对所提方案进行详细的分析;第四节给出本文的总结。

1 背景知识

本节将以经典的比特币区块链为例回顾区块链协议的相关知识,分析在区块链交易中插入非法数据的常用方法。

1.1 区块链协议

区块链是一种由点对点(P2P)网络维护的不断增长的分布式账本,主要由以下组件构成:

(1)节点和网络:节点是任何可以连接到区块链网络的计算设备,其相互连接组成 P2P 网络来传播交易数据和区块数据。节点可以通过接收和发送交易来交换资产,也可根据共识协议计算新区块(挖矿)来赚取奖励(又称为矿工)。矿工对区块链的贡献与其拥有的资源(例如计算能力、代币财富和存储空间等)成正比。在 P2P 网络中,矿工收集并验证交易,计算新区块以维护区块链良性增长。

(2)交易:区块链上常见的交易为货币交易[13]和部署智能合约的交易[14]。以比特币交易为例,其主要由输入、输出、金额和时间戳等字段组成。输入和输出分别定义了交易金额的来源地址(支付者)和目的地址(接收者)。每笔交易可以有多个输入,每个输入引用一笔未花交易的输出,并且每个输出只能被引用一次(花一次)。如果输出被引用 2 次甚至多次,那么交易存在冲突,

并且存在冲突的交易是无效的。在交易中,交易者会产生 1 对密钥,私钥用来对交易进行签名,公钥用来验证签名,并且公钥经过编码后作为交易者的比特币地址,为在线交易提供匿名性,从而不会透露交易者真实身份。节点可通过验证签名的有效性来确定交易的所有权和完整性。这种输入、输出的引用和验证方式为交易提供了可审计性和可追溯性。

(3)区块:由矿工产生,包含区块头和交易集。以比特币区块计算方式(PoW[15])为例,区块头 BH 主要包含上一个区块头的哈希值、有效交易集构成的 Merkle 树的根节点和计数 *nonce* 等重要信息,且区块头的哈希值满足如下等式:

$$SHA256(SHA256(BH)) \leqslant Target,$$

其中,参数 $Target \in N$ 定义了难度级别。每 2 016 块进行一次调整,从而保证计算出一个有效区块的平均时间间隔保持在 10 min 左右,这也是区块链安全性和吞吐量之间的权衡[16-17]。受块奖励的激励,矿工竞相通过改变计数 *nonce* 来计算满足上述式子的新区块。显然,成功算出新区块的机率与矿工拥有的哈希算力成正比。当网络中同时出现多个块时,矿工们总是选择高度最大的块,遵循最长链原则。

(4)共识协议:它是区块链的验证和决策机制,用来确保区块链的所有诚实节点在没有任何可信任授权中心的情况下就统一的交易历史或状态达成一致。目前,具有代表性的共识机制有工作量证明(PoW)、权益证明(PoS)、拜占庭容错(BFT)和混合 BFT 等[18]。

1.2　链上非法数据分析

在本节中,以比特币交易为例,介绍常用非法数据插入方式和威胁。

(1)输入脚本(Input script):Coinbase 交易在其输入字段提供了 100 B 的空间来存储任意数据。P2SH 类型交易的输入脚本 RedeemScript 字段也是一种方便的文本存储方法,可以存储 1.5 kB 数据[7]。在上述 2 个字段存储数据的交易仍然是可花费的。

(2)输出脚本(Output script):将语义数据(例如句子和图像)编码为其比特币地址来接收交易。由于很难找到与语义数据对应的私钥来产生新交易花费这些比特币,因此,这种交易会被永久存储在 UTXO 数据库中。例如,基于 Web 的服务 CryptoGraffiti[19]可以将多达 50 KB 的数据编码为多笔交易的多个输出脚本。

(3)OP_RETURN:比特币交易的特殊字段 OP_RETURN 可以存储 80 B 任意数据,且交易是可花费的[20-21]。例如:服务 Blockstore[22]提供在字段 OP_RETURN 中存储资产元数据和公证文件的服务。

无论使用哪种方法插入数据,支付方都必须消耗一些可花费的比特币以激励矿工将交易添加到区块链上。目前,区块链已经发现包含个人疾病和遗传信息的医疗记录、儿童色情数据、侵犯版权的音乐元数据和跨站点脚本(XSS)攻击的恶意代码等非法数据。现有的区块链系统无法提供数据内容问责,常常被用来实现网络犯罪,因此,清除这些非法数据是受害者和刑警的迫切需求[23]。

2　可编辑区块链方案的回顾及分析

在本节中,笔者分别回顾和分析 3 类(基于共识、基于变色龙哈希和基于元交易数据)具有代表性的可编辑区块链方案,相关对比如表 1 所示。

表 1　3 类可编辑区块链方案的对比

Table 1　Comparisons between three categories of redactable blockchains

指　标	基于共识机制	基于变色龙哈希	基于元交易数据
有效性	√	□	□
一致性	√	√	□
抗双花攻击	√	□	×
抗篡改攻击	√	□	√
抗拒绝服务攻击	√	□	×
无共识延迟	×	√	□
重新计算 PoW	√	×	□
高效率	×	√	□
兼容性	□	□	□

注:√:满足;□:部分满足;×:不满足

2.1　基于共识机制的可编辑区块链方案

基于共识机制的可编辑区块链方案通常依赖于共识规则来约定验证和认可编辑后的交易或者区块的最终状态,这类方案需要较长的“投票”周期来对编辑后的新状态进行验证和仲裁。周期结束后,节点在区块链的新状态上达成一致的共识。本小节回顾和分析 2 个具有代表性的案例来说明

这类方案的性能。

2.1.1 基于硬分叉的编辑方案

分叉是指网络中同时出现2个或多个相同高度的区块,即出现替代链。硬分叉通常指替代链和原主链的共识规则不同的情况,被广泛用来修改历史记录、纠正代码中的重要安全风险、添加新功能、回滚交易以及消除黑客攻击的影响等。例如,以太坊区块链在2016年使用硬分叉来回滚交易,恢复了攻击者发起"DAO"攻击盗取的价值数千万美元的以太币。遵循新共识规则的替代链是一条永久链,与原始链分别独立运行且同时存在。

分析:硬分叉需要大多数节点参与产生,节点按照新共识机制产生替代链,因此,替代链的安全性是得到保证的。然而,硬分叉并不能实现真正的完全擦除非法数据,因为原始链仍然保留着非法数据。此外,硬分叉过程需要重新计算所有区块,这将消耗节点大量的资源,如哈希算力和带宽。使用硬分叉方式来实现历史数据编辑的可扩展性低、效率低且成本非常高。因此,它不适用于频繁的编辑请求。

2.1.2 基于链上投票共识的可编辑区块链方案

Deuber等[24]提出一种拥有双链结构并通过链上投票的方式来实现数据编辑的方案。该方案在区块头中记录了区块的原始状态和编辑后的新状态。矿工通过对编辑后区块的新状态进行链上投票来对编辑的有效性进行仲裁。矿工每产生一个新区块,就有一次投票机会,拥有哈希算力较多的节点就拥有较多的投票机会。如果新状态在投票周期内获得足够多的赞成票,那么编辑后的新状态有效且被系统认可,否则新状态无效且被拒绝。另外,记录在区块中的原始状态保证了编辑过的区块与下一个区块之间的连通性。

分析:该方案基于两条哈希链的巧妙结构和链上投票的机制来实现区块级的修改、删除等操作,能确保编辑后区块的连通性和有效性,并使网络节点在投票周期过后就区块链的新状态达成共识,满足区块链的一致性需求。然而,方案要求重新计算编辑的区块的PoW,增加了编辑的计算开销;要求每次编辑需要在1 024个连续块(约7 d)中至少有50%以上的区块是赞成票时才能有效,编辑的效率较低,无法实现频繁的编辑请求;方案要求修改区块头的数据结构,与比特币、以太坊等经典区块链系统兼容性差。

2.2 基于变色龙哈希的可编辑区块链方案

这种类型的方案主要是利用密码学原语变色龙哈希函数来实现链上数据编辑。其中,变色龙哈希函数又称为陷门抗碰撞哈希函数,满足:①拥有公钥的任何人都能有效地计算变色龙哈希值;②任何没有陷门密钥的敌手在概率多项式时间内几乎不可能找到两个不同的输入映射到相同的输出;③拥有陷门密钥者可以轻松地计算出任意输入的变色龙哈希碰撞,将不同的输入映射到相同的输出。

2.2.1 基于变色龙哈希的可编辑区块链方案

Ateniese等[25]在2017年提出了第一个基于变色龙哈希函数(记为CHASH)实现块级区块链数据可编辑的方案。方案用变色龙哈希函数替换计算PoW的哈希函数来使链上数据可编辑,可以表示为

$$SHA256(CHASH(pk,BH;r)) \leqslant Target,$$

其中,pk是公钥,r是随机数。当执行编辑时,陷门密钥的拥有者能有效地计算出参数r^*,使得原始区块与新区块映射到相同的哈希值,可表示为

$$CHASH(pk,BH;r) = CHASH(pk,BH^*;r^*).$$

如果需要删除整个区块,则只需要计算下一个区块的变色龙哈希碰撞就可以使链保持连通性。

分析:方案通过计算哈希碰撞,保持编辑前后区块头的哈希值不变,不需要重新计算PoW,既保证了一致性和连通性,又能以较低的消耗和较高的效率实现链上数据编辑。然而,方案要修改区块头的数据结构来存储参数r,兼容性也较差;块级的数据编辑容易造成交易冲突,导致相关交易失效,容易遭受双花攻击和篡改攻击;在去中心化场景中陷门密钥的分配采用(n, t)门限密钥共享方案,无法抵抗合谋攻击。

2.2.2 基于属性控制变色龙哈希的可编辑区块链方案

Derler等[26]针对细粒度编辑及对编辑权限的控制,在2019年提出基于属性策略控制的变色龙哈希函数来实现交易级区块链数据的可控编辑。方案主要结合了增强的变色龙哈希函数[27]和基于密文策略的属性的加密(CP-ABE)算法[28]。在交易产生时,交易者需要预设编辑该交易的属性策略(由属性集构成)。当执行编辑请求时,只有满足策略的用户才能恢复出私钥,有效地计算出编辑后交易的变色龙哈希碰撞参数,实现交易数据的编辑。

分析:方案实现了细粒度的交易级数据编辑和细粒度的编辑权控制,解决了滥用编辑的问题;方案满足不可区分性和抗碰撞性,因此,敌手在概率多项式时间内几乎不可能恢复密钥来发起双花攻击和篡改攻击。然而,方案忽略了交易修改之后与原始签名不匹配导致签名失效的问题。当交易支付者处于脱机状态而无法对编辑后的交易重新签名时,交易难以通过验证获得诚实节点认可。另外,拥有不同属性的用户可以通过相互勾结来满足属性策略,恢复出密钥来执行任何编辑,这将对区块链的可信度和链上数据的安全性造成直接威胁。

2.2.3 基于身份控制变色龙哈希的可编辑区块链方案

Huang 等[29-30]针对编辑权限的控制问题提出了基于身份控制变色龙哈希的可编辑区块链方案。在交易产生时,交易者预设能编辑该交易的身份 ID,只有拥有该 ID 的用户才能恢复出密钥来计算变色龙哈希碰撞,执行编辑。此外,方案对交易进行变色龙哈希之后再进行常规签名操作,从而保证原始签名的有效性,不需要对编辑后的交易重新签名。在文献[30]中,每笔交易的编辑权限被限制在一个预设的时间窗内。只有在时间窗内,拥有正确 ID 的用户才能有效地编辑该笔交易。

分析:方案基于变色龙哈希,满足有效性、一致性、消耗较低及效率较高等性能需求。然而,方案容易出现由于拥有 ID 的用户不在线或者不存在而无法有效地实现编辑等问题。在区块链上恶意插入非法数据牟利的敌手可以利用这个问题避免其数据被擦除。

2.3 基于元交易数据的可编辑区块链方案

这种类型的方案主要是通过命令修改元交易数据(Mate-transaction)的方式,来实现对区块链交易数据的编辑。节点在接收到编辑请求指令(是一笔触发编辑操作的特殊类型交易)后产生新交易并改变交易视图,实现编辑操作。

Puddu 等[31]设计了 2 种特殊类型的交易,一种用来触发编辑操作,一种用来替换原始交易执行编辑。交易在产生时预设编辑策略,包括编辑对象、编辑者和时间窗等重要信息。在每轮编辑中,只有 1 笔处于激活状态的可编辑交易可以执行编辑操作。全网对触发编辑的交易和编辑后的新交易进行验证和链下投票,只有获得足够多的赞成票的编辑才是有效的。与链上投票方式不同,链下投票是每个身份拥有 1 票,而且只需要节点通过通信交互,对编辑交易进行仲裁,不需要将相关信息存储在区块链上。最后,有效的新交易生效,原始交易失效。

分析:方案基于产生特殊类型的交易来编辑链上数据的方式,不依赖于复杂的密码学原语,也不需要较长的投票期来保证安全性。然而,方案仍然存在编辑容易引起的相关交易失效的问题,无法保证一致性。链下投票机制容易遭受女巫攻击,恶意者可以轻而易举地伪造足够多的身份来参与投票。方案通过加密来隐藏交易历史的方式增加了交易验证难度,破坏了区块链的可审计性和可追溯性,为敌手发起双花攻击和篡改攻击提供便利。

3 基于去中心化变色龙哈希的可编辑区块链方案

本节主要介绍本文提出的一种新方案,即基于去中心化变色龙哈希的可编辑区块链方案。该方案的算法步骤及分析详细介绍如下。

3.1 算法步骤

初始化:系统进行初始化以产生系统参数。输入安全参数 κ,随机选取素数 p_1,产生阶为 p_1 的群 $G=\langle g\rangle$,双线性映射 $\hat{e}:G\times G\rightarrow G_T$。每个授权中心针对其负责验证的属性 $i\in S$ 产生 1 对密钥 (sk,pk),其中,$pk\leftarrow(\hat{e}(g,g)^{\alpha_i},g^{\beta_i})$,$sk\leftarrow(\alpha_i,\beta_i)$,$\alpha_i,\beta_i\xleftarrow{R}\mathbb{Z}_n$,产生 RSA 算法参数$(n^*,p^*,q^*,\cdot,\cdot)\leftarrow\text{RSAKGen}(1^\kappa)$,$n^*$ 公开。

发布可编辑的交易:发起交易的用户产生可编辑的交易数据并基于属性设置编辑交易的策略。首先,用户设置由属性 $i\in\mathbb{S}$构成的访问策略,表示为$(\boldsymbol{A}_{v\times l},\rho)$,其中,$\boldsymbol{A}_{v\times l}$为访问矩阵,函数 $\rho:[v]\rightarrow\mathbb{S}$将矩阵的行号映射到属性。其次,用户随机选取秘密 $s\xleftarrow{R}\mathbb{Z}_n$,产生 RSA 算法参数$(n,p,q,\cdot,\cdot)\leftarrow\text{RSAKGen}(1^\kappa)$,$etd\leftarrow\text{encode}(p^*,q^*,p,q,)$,公开 n。针对构成其访问策略的每个属性,随机选取私密参数 $r_x\xleftarrow{R}\mathbb{Z}_n$,其中 $x\in v$。那么,令交易数据记为 m,用户随机选择大数 e 和常数 $r\xleftarrow{R}\mathbb{Z}_n$,计算变色龙哈希$(h,r)$且 $h=H(m)r^e\ mod(nn^*)$,其中,函数 $H:\{0,1\}^*\rightarrow\mathbb{Z}^*_{nn^*}$。同时,计算辅助信息

$ct \leftarrow \{c_0 = etd \cdot \hat{e}(g,g)^s, c_{1,x} = \hat{e}(g,g)^{\lambda_x}\hat{e}(g,g)^{\alpha_{\rho(x)}r_x}, c_{2,x} = g^{r_x}, c_{3,x} = g^{\beta_{\rho(x)}r_x}g^{\omega_x}\}$,

其中,$\lambda_x = A_x \cdot (s, a_2, \cdots, a_l)$,$\omega_x = A_x \cdot (0, b_2, \cdots, b_l)$,且$(a_2, \cdots, a_l)$,$(b_2, \cdots, b_l) \xleftarrow{R} \mathbb{Z}_n$。最后,用户按照传统区块链协议,对交易的变色龙哈希 h 进行签名后((h,r)保存在交易的非签名数据结构中),再广播该笔可编辑的交易到区块链网络中。

可编辑交易上链:网络中的节点收到该笔可编辑交易(m,h,r,ct)后,先验证变色龙哈希值的有效性,再按照传统区块链协议验证该交易的有效性,2 者都有效的交易会被节点收集并上链。如果等式 $h = H(m)r^e \ mod(nn^*)$ 成立,则变色龙哈希值有效;如果等式不成立,则变色龙哈希值无效。

编辑可编辑交易:用户发起链上数据编辑请求时,将通过多个授权中心的属性验证,恢复密钥,执行编辑操作。假设用户拥有属性 $i \in \mathbb{S}$,那么用户和负责该属性的授权中心进行交互,获取密钥 $SK_{i,GID} = (g^{\alpha_i}, y^{\beta_i})$,其中 $y = H(GID)$,GID 是用户的全局身份。针对同一个策略,每个授权中心只能负责一个属性的验证及对应密钥的产生。因此,用户与多个授权中心交互后,获取一组其属性集对应的密钥集。如果用户的属性集能满足该笔交易的策略,那么该用户的秘钥集可以正确恢复出参数 $etd = c_0 / \prod_x (D_x)^{C_x}$,其中

$D_x = c_{1,x} \cdot \hat{e}(H(GID), c_{3,x}) / \hat{e}(SK_{\rho(x),GID}, c_{2,x}) = \hat{e}^{\lambda_x}\hat{e}(y,g)^{\omega_x}$.

随后,用户可以计算出正确密钥 d,满足 $ed \equiv 1 \ mod(\varphi(nn^*))$。因此,针对编辑后的交易 m',计算出其满足变色龙哈希碰撞的随机参数 $r' = (hH(m')^{-1})^d \ mod(nn^*)$。广播编辑后的交易数据$(m',h,r',ct)$到区块链网络中,节点验证变色龙哈希和交易的有效性。如果有效,节点更新其存储在本地的链上数据;如果无效,则保持原始交易数据。

3.2 方案分析

本节主要从正确性、去中心化控制、安全性以及效率等方面分析所提方案,具体描述如下。

正确性:对于 $m' \neq m$,已知属性集$\mathbb{S}$满足预设访问控制策略(A,ρ),变色龙哈希(m,h,r,ct),那么在多项式概率时间内必能找到有效参数 $r' = (hH(m')^{-1})^d \ mod(nn^*)$,使得 $h' = H(m')((hH(m')^{-1})^d)^e = h$。

去中心化控制:在所提方案中,每个诚信授权中心只能负责同一个策略的一个属性,从而能保证该策略由多个半诚信授权中心验证,即将编辑的控制及验证分散到多个授权中心,而不是传统可编辑区块链方案中由一个可信授权中心控制的集中式模式。另外,针对每笔交易,都产生新的密钥参数和属性策略,从而也保证了对每笔交易的编辑控制分散到不同的授权中心组,而不是 1 组固定的授权中心。

有效性:在所提方案中,已知密钥可以有效地计算出变色龙哈希的碰撞,从而保证不同数据映射到同一个哈希值。也就是说,交易编辑前后拥有相同的哈希值。如果编辑后的交易是正确的,那么其能通过区块链协议的交易有效性验证,得到网络节点的认可,最终上链。

一致性:所提方案保证了交易编辑前后的哈希值不变,因此,交易编辑后原交易签名有效,记录在区块头中原始 *Merkle* 根节点有效,从而保证了花费该笔交易的下一笔交易的有效性,也保证了引用该区块的下一个区块的有效性。一笔有效的编辑过的交易最终会被诚实的网络节点认可。也就是说,方案保证了诚实节点的区块链视图保持一致。

抗双花攻击:已知原始交易 m 及花费 m 的交易 m_1,假设恶意用户企图通过请求将原始交易 m 编辑为 m^* 以发起双花攻击,那么她将发布一笔新交易 m_2 来花费交易 m^*,从而实现原始交易的金额被交易 m_1 和 m_2 花 2 次。由所提方案的正确性和有效性可知,原始交易和编辑后的交易映射到同一个哈希值。因此,网络节点在验证新交易时,很容易发现交易 m_2 与 m_1 引用了同一笔交易的哈希值,存在冲突。那么,交易 m_2 不可能通过有效性验证,也不可能被节点认可。因此,双花攻击失败,所提方案是抗双花攻击的。

抗篡改攻击:假设恶意用户企图通过请求将原始交易 m 编辑为 m^* 以发起篡改攻击,那么他可能修改原始交易的输入脚本、输出脚本、金额及时间戳等关键信息牟取利益。然而,所提方案旨在于擦除链上交易所携带的非交易本身的非法数据,保护受害者的隐私,清洗区块链中的非法数

据。根据链上原始交易的信息对编辑后的交易进行关键信息验证,如果输入脚本、输出脚本、金额及时间戳等关键信息被篡改,那么该笔编辑交易将不能通过网络节点的验证和认可。因此,篡改交易无法上链,则篡改攻击失败。

抗拒绝服务攻击:所提方案将其编辑控制属性分散到不同的授权中心(由网络节点担任),因此,除非攻击者同时对大多数网络节点进行拒绝服务攻击,否则其无法通过对单一节点进行拒绝服务攻击来使系统停止工作。

共识延迟:在所提方案中,编辑后的交易广播到区块链网络中,由节点收集、验证。若交易有效,则会替换原始交易,实现上链;若交易无效,则会被丢弃。1 笔交易广播到全网所需时间大约为 12 s,因此,该交易能迅速地被诚实节点认可或者丢弃,大多数节点对区块链的状态也能迅速地达成一致且不会产生分歧,无共识延迟。

额外消耗:所提方案实现交易级的可编辑,不涉及区块的重写或者重新产生,因此,不需要消耗算力以解决额外的 PoW。相比需要重新计算区块 PoW 的区块级编辑方案,本文所提方案的额外消耗非常低。

效率:对方案算法进行仿真实验,预设编辑策略由 32 个属性构成,即由 32 个节点担任授权中心,则系统初始化需要 153.839 ms,计算交易变色龙哈希值和辅助信息需要 489.127 ms,计算编辑后交易的碰撞信息需要 179.014 ms,由此可知方案的子算法具有高效率性。再结合共识延迟等分析可知,所提方案能高效地实现区块链数据编辑。

兼容性:所提方案需要在交易的非签名部分存储(r,ct),交易的数据结构需要进行细微的修改。相比其他的区块级编辑方案,本方案的兼容性更优。

4　总　结

本文主要研究了基于可编辑区块实现链上隐私数据和非法数据的编辑问题,综合分析了具有代表性的 3 类可编辑区块链技术方案,并总结和对比了这 3 类方案的性能,有利于后续研究把握方向。另外,针对现有方案的缺陷,提出了基于去中心化的变色龙哈希的区块链数据可编辑方案。同时,对方案进行了详细的分析,分析表明,所提方案能满足安全需要和性能需求。未来,笔者将继续研究可编辑区块链的编辑权限的动态控制,解决多个授权中心的离线及职责转让过度等问题,针对具体的应用场景求设计满足需求的可编辑区块链方案。

参考文献:

[1] Ma Z, Jiang M, Gao H, et al. Blockchain for digital rights management[J]. Future Generation Computer Systems, 2018, 89:746-764.

[2] Aitzhan N Z, Svetinovic D. Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams[J]. IEEE Transactions on Dependable and Secure Computing, 2018, 15(5):840-852.

[3] Khan K M, Arshad J, Khan M M. Investigating performance constraints for blockchain based secure e-voting system[J]. Future Generation Computer Systems, 2020, 105:13-26.

[4] Xu Y, Ren J, Zhang Y, et al. Blockchain empowered arbitrable data auditing scheme for network storage as a service[J]. IEEE Transactions on Services Computing, 2020, 13(2):289-300.

[5] De A E J, Faiçal B S, Krishnamachari B, et al. A survey of blockchain-based strategies for healthcare[J]. ACM Computing Surveys (CSUR), 2020, 53(2):1-27.

[6] Liu S. Global blockchain solutions spending 2017-2024[EB/OL]. (2021-04-22)[2021-05-28]. https://www.statista.com/statistics/800426/worldwide-blockchain-solutions-spending/.

[7] Matzutt R, Hiller J, Henze M, et al. A quantitative analysis of the impact of arbitrary blockchain content on bitcoin[C]// International Conference on Financial Cryptography and Data Security. Berlin, Heidelberg: Springer Verlag, 2018:420-438.

[8] Zhang D, Le J, Lei X, et al. Exploring the redaction mechanisms of mutable blockchains: A comprehensive survey[J]. International Journal of Intelligent Systems, 2021,36(9):5051-5084.

[9] Dorri A, Kanhere S S, Jurdak R. Mof-bc: A memory optimized and flexible blockchain for large scale networks[J]. Future Generation Computer Systems, 2019, 92:357-373.

[10] Pyoung C K, Baek S J. Blockchain of finite-lifetime blocks with applications to edge-based iot[J]. IEEE Internet of Things Journal, 2020, 7(3):2102-2116.

[11] Intersoft Consulting. GDPR: General data protection regulation[EB/OL]. (2016-04-27)[2021-03-12]. https://gdpr-info.eu.

[12] Politou E, Casino F, Alepis E, et al. Blockchain mutability: Challenges and proposed solutions[J/OL]. (2019-10-25)[2021-03-15]. IEEE Transactions on Emerging Topics in Computing, 2019. https://ieexplore.ieee.org/abstract/document/8883080.

[13] Zhang D, Le J, Mu N, et al. An anonymous off-blockchain micropayments scheme for cryptocurrencies in the real world[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(1):32-42.

[14] Wang S, Ouyang L, Yuan Y, et al. Blockchain-enabled smart contracts: Architecture, applications, and future trends[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019, 49(11):2266-2277.

[15] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[EB/OL]. (2020-08-22)[2021-04-26]. https://bitcoin.org/bitcoin.pdf.

[16] Akbari E, Zhao W, Yang S, et al. The impact of block parameters on the throughput and security of blockchains[C]//Proceedings of the 2020 the 2nd International Conference on Blockchain Technology. New York: Association for Computing Machinery, 2020:13-18.

[17] Kiayias A, Panagiotakos G. Speed-security tradeoffs in blockchain protocols[EB/OL]. (2016-10-13)[2021-04-07]. IACR Cryptology ePrint Archive, 2015. https://eprint.iacr.org/2015/1019.

[18] Xiao Y, Zhang N, Lou W, et al. A survey of distributed consensus protocols for blockchain networks[J]. IEEE Communications Surveys & Tutorials, 2020, 22(2):1432-1465.

[19] Hyena. CryptoGraffiti[EB/OL]. (2021-05-10)[2021-05-19]. https://cryptograffiti.info/.

[20] Bartoletti M, Pompianu L. An analysis of bitcoin op_return metadata[C]//International Conference on Financial Cryptography and Data Security. Cham: Springer Verlag, 2017: 218-230.

[21] Matzutt R, Henze M, Ziegeldorf J H, et al. Thwarting unwanted blockchain content insertion[C]//2018 IEEE International Conference on Cloud Engineering (IC2E). Piscataway: Institute of Electrical and Electronics Engineers Inc., 2018:364-370.

[22] Blockstack. Blockstore[EB/OL]. (2018-06-03)[2021-04-10]. https://github.com/blockstack-packages/blockchain-id-deprecated/wiki/Blockstore.

[23] Truong N B, Sun K, Lee G M, et al. GDPR-compliant personal data management: A blockchain-based solution[J]. IEEE Transactions on Information Forensics and Security, 2020, 15:1746-1761.

[24] Deuber D, Magri B, Thyagarajan S A K. Redactable blockchain in the permissionless setting[C]//2019 IEEE Symposium on Security and Privacy. Piscataway: Institute of Electrical and Electronics Engineers Inc., 2019:124-138.

[25] Ateniese G, Magri B, Venturi D, et al. Redactable blockchain-or-rewriting history in bitcoin and friends[C]//2017 IEEE European Symposium on Security and Privacy. Piscataway: Institute of Electrical and Electronics Engineers Inc., 2017: 111-126.

[26] Derler D, Samelin K, Slamanig D, et al. Fine-grained and controlled rewriting in blockchains: Chameleon-hashing gone attribute-based[C]//26th Annual Network and Distributed System Security Symposium. Reston: Internet Society, 2019:1-15.

[27] Camenisch J, Derler D, Krenn S, et al. Chameleon-hashes with ephemeral trapdoors[C]//IACR International Workshop on Public Key Cryptography. Berlin, Heidelberg: Springer Verlag, 2017:152-182.

[28] Bethencourt J, Sahai A, Waters B. Ciphertext-policy attribute-based encryption[C]//2007 IEEE Symposium on Security and Privacy. Piscataway: Institute of Electrical and Electronics Engineers Inc., 2007:321-334.

[29] Huang K, Zhang X, Mu Y, et al. Building redactable consortium blockchain for industrial internet-of-things[J]. IEEE Transactions on Industrial Informatics, 2019,15(6):3670-3679.

[30] Huang K, Zhang X, Mu Y, et al. Achieving intelligent trust-layer for internet-of-things via self-redactable blockchain[J]. IEEE Transactions on Industrial Informatics, 2020,16(4):2677-2686.

[31] Puddu I, Dmitrienko A, Capkun S. μchain: How to forget without hard forks[J/OL]. (2017-10-23)[2021-01-24] IACR Cryptology ePrint Archive, 2017, https://eprint.iacr.org/2017/106.pdf.

【责任编辑:周 全】