物联网下基于属性加密且支持溯源的数据共享方案

关川江[1] 史国振[2] 毛 明[2]

[1](西安电子科技大学通信工程学院 陕西 西安 710071)

[2](北京电子科技学院电子与通信工程系 北京 100070)

摘 要 针对一对多模型下数据共享缺乏细粒度访问控制和数据流转缺乏溯源的问题,提出一种物联网下基于属性加密且支持溯源的数据共享方案。基于 Waters 所提密文策略属性方案对共享密钥加密,实现数据细粒度访问控制。依据区块链智能合约技术,有效防止授权中心权限过大导致的安全隐患,保证所有参与者按照智能合约事先约定规则执行,用户验证数据正确性不需要额外计算开销。安全性分析表明,所提方案能很好地保护用户数据的安全性,系统数据流转可追溯。实验分析表明,所提方案在区块链交易数量、属性令牌生成、资源搜索以及验证方面有一定优势,适用于物联网设备资源受限场景下一对多数据共享。

关键词 属性基加密 智能合约 数据溯源 物联网 区块链

中图分类号 TP309 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2024.08.050

DATA SHARING SCHEME BASED ON ATTRIBUTE ENCRYPTION AND SUPPORTING TRACEABILITY UNDER INTERNET OF THINGS

Guan Chuanjiang[1] Shi Guozhen[2] Mao Ming[2]

[1](*School of Telecommunications Engineering, Xidian University, Xi'an 710071, Shaanxi, China*)

[2](*Department of Electronics and Communication Engineering, Beijing Electronic Science and Technology Institute, Beijing 100070, China*)

Abstract Aimed at the problems of lack of fine-grained access control and lack of traceability in data sharing under one-to-many model, a data sharing scheme based on attribute encryption and supporting traceability under internet of things (IoT) is proposed. On the basis of ciphertext policy attribute scheme proposed by Waters, the shared key was encrypted to realize the fine-grained access control of data. Based on the blockchain smart contract technology, the scheme effectively prevented the security risks caused by the excessive authority of the central authorization center, ensured that all participants followed the pre-agreed rules of the smart contract, and did not require additional computational overhead for users to verify the correctness of the data. Security analysis shows that the proposed scheme can well protect the security of user data, and the system data flow can be traced. Experimental analysis shows that the proposed scheme has certain advantages in the number of blockchain transactions, attribute token generation, resource search, and verification, and it is suitable for one-to-many data sharing in the IoT device resource-constrained scenario.

Keywords Attribute-based encryption Smart contract Data traceability Internet of things Blockchain

0 引 言

物联网正以许多激进的方式改变着人们的生活。随着物联网日益普及,带来了更高效率、准确性和经济效益[1]。据统计,全球有超过 240 多亿物联网设备,这些设备会将收集到的数据发送到集中式服务器进行处理、分析和其他后续操作。在文献[2]讨论的物联网世界中,数据本身将成为一种有价值的商品。这些数据可能是来自极限运动、火山爆发、热门旅游景点等实时视频流的网络摄像头,来自临床试验中患者可穿戴的医疗感知设备,来自智能办公室和智慧城市中的各种公共物

收稿日期:2020-12-14。国家重点研发计划项目(2017YFB0801803)。**关川江**,硕士生,主研领域:访问控制和云安全。**史国振**,教授。**毛明**,教授。

联网设备。物联网设备产生和交换巨量的数据,从安全性和可扩展性方面带来了前所未有的挑战。在物联网的特殊背景下,由于数据量大、设备的异构性、各方之间信任不足以及数据处理不透明等问题,可用于安全数据共享的解决方案不足。在传统的基于云的物联网架构中,集中式的云服务器收集和控制所有的数据,这带来了两个问题:1) 敏感数据容易泄露;2) 服务器可能在不通知数据所有者的情况下与其他实体交换敏感数据。此外,集中式的服务架构是不可信的、容易受到黑客攻击和妥协的中心点。因此,有必要设计一个高度可信、安全、可扩展和高度自动化的物联网数据共享方案。

在分布式应用中信任和透明性是关键因素,区块链技术已经被证明是一种有效的解决方案。因此,行业界和研究界对如何有效地结合物联网平台和区块链技术展开了激烈的讨论。许多行业参与者已经开发了一些针对物联网的区块链平台,如:IOTA1、IoTeX2、Atonomi3。同样,许多学者建议直接将区块链平台用来解决物联网平台中的安全数据共享问题[3-5]。在这些方案中,大多数遵循将存储系统与服务提供系统分开,其中存储系统(例如云提供商)托管数据本身,服务系统由区块链保证平台保证信任分布和完整性。Shafagh 等[6]建议存储服务提供商充当策略的决策和执行点,区块链确保策略的完整性,并允许对策略进行公开审核。文献[7-8]建议在数据存储到云之前对其进行加密,由区块链管理解密数据所需密钥的授权。在这种情况下,减轻了恶意存储服务提供商对数据的危害。此外,进一步避免了解密密钥授权单点故障的问题。由于物联网中存在巨量的数据共享,现有的解决方案并未完全解决访问控制问题的可扩展性。在他们的提议中,将访问控制策略上传到区块链上,从而使访问策略的信任度分散;但是它要求物联网设备处理策略的更新,显然该机制无法扩展到资源有限的物联网系统模型中。

因此,本文提出了一种物联网下基于属性加密且支持溯源的数据共享方案,该方案充分利用区块链的优势限制成员的恶意行为。鉴于物联网数据海量问题,本文方案建立脱链式数据存储方式,数据上传到运存前由数据所有者加密。数据所有者对解密密钥指定访问结构,实现数据细粒度访问控制,数据使用者的属性私钥满足定义的访问结构才能恢复解密密钥。利用智能合约将属性令牌存储在区块链中,利用区块链不可篡改性,实现数据操作的可追溯。同时,区块链平台与可信中心之间相互制约,有效减轻授权中心权力过大导致的安全隐患。此外,区块链公平机制实现服务-支付公平,用户只需遵守合约规则,在本地进行简单比较验证而不需要额外的计算代价就可以获取正确数据,同时云服务提供商也可以获得相应的报酬。分析表明,本文方案有效防止数据泄露并且支持数据操作可追溯性,实现物联网数据安全共享。性能对比和实验分析表明,本文访问在属性令牌生成、资源搜索、正确性验证,以及区块链交易数量上具有一定的优势。

1 背景知识

本节介绍和定义基于属性加密且支持追溯的物联网数据共享方案使用到的预备知识,为方便后文描述,主要对区块链、智能合约以及属性基加密相关的数学基础进行简要介绍。

1.1 区块链

区块链的概念最早出现在比特币白皮书[9]中,其有效地解决了一些分布式应用场景的问题。区块链通过对等网络技术、共识算法和密码学技术,将各个区块数据按照时间顺序像链条一样组织起来。根据应用场景和功能需求,区块链系统可分为公有链、私有链和联盟链。在公有链系统中,每一台计算设备都可以参与账本的维护。但是在实际应用过程中,特别是物联网系统中,存在效率低下的弊端。联盟区块链是指由多个类似的组织管理和维护的区块链,每个组织运行一个或多个节点。仅允许联盟节点进行投票、记账和构造区块,参与者需要授权才能加入网络并与利益相关者一起维护区块链[14]。联盟区块链具有交易速度快、无需挖矿、交易成本低,以及支持监管的优点。表 1 比较了公有链和联盟链系统的区块。

表 1 公有链和联盟链的对比

指标	公有链	联盟链
参与决策	所有节点	授权节点
成员数量	多	少
共识算法	POW,POS	PBFT,Raft
速度	慢	快
安全性	安全	安全
能耗	高	低
身份	匿名、假名	已知身份

1.2 智能合约

智能合约是计算机自动处理网络成员预先约定规则的协议[15]。智能合约提供数据交互接口可供所有成员使用。区块链成员以类似于添加交易的方式将智能合约添加到区块中,更新智能合约状态的事务也会被记录在下一个区块中,该机制使智能合约以相同的

方式对交易不可变。智能合约在存储、读取、执行的整个过程中由具有不可篡改特性的区块链和密码学散列算法来保障其透明性、不可篡改性、不可否认性,以及可追溯性。智能合约通常由系统节点强制执行,因此单个节点实体不能绕过代码中定义的规则。智能合约主要优点是可以使组织的业务逻辑自动化,因此可以消除可能导致的法律纠纷的人为错误和误解的影响。

1.3　数学基础

1) 双线性映射。

定义 1　设 G_0 和 G_1 是两个以大素数 p 为阶的乘法循环群,g 为 G_0 的生成元,e 为双线性映射,满足 e: $G_0 \times G_0 \rightarrow G_1$,且有如下性质[16]:

(1) 双线性:对于任意元素 $\alpha,\beta \in G_0$ 和 $a,b \in Z_p$,都有 $e(\alpha^a,\beta^b)=e(\alpha,\beta)^{ab}$ 成立。

(2) 非退化性:$e(g,g)\neq 1$。

(3) 可计算性:对于任意的 $x,y \in G_0$,有 $e(x,y)$ 在多项式时间上是可计算的。

2) 访问结构。

定义 2　设 $U=\{u_1,u_2,\cdots,u_n\}$ 为所有的属性集合,若存在访问结构 $A \subseteq 2^U$。当属性集合 $A_1 \subseteq A$,称 A_1 为授权属性集合,否则为非授权属性集合。对于 $\forall B$, $C \in A$,如果满足 $B \in A$,且 $B \in C$,有 $C \in A$,那么称 A 是单调访问结构[13]。本文只考虑单调访问机构。

3) 线性秘密共享方案。

定义 3　对于一个由参与者组成的集合 Ω 上的秘密共享方案 Π,若满足以下条件,则 Π 就是 Z_p 上的线性秘密共享方案。

(1) 所有成员的秘密份额组成 Z_p 上的一个向量。

(2) 在秘密共享方案 Π 中,共享矩阵 $\boldsymbol{M}_{r\times l}$ 中第 $i(i=1,2,\cdots,l)$ 行表示集合中的第 i 个参与成员,经过映射 $\theta:\{1,2,\cdots,l\}\rightarrow\Omega$,映射到参与者集合 Ω 中。设共享秘密为 s,其中 $s \in Z_p$,向量 $\boldsymbol{v}=(s,r_2,r_3,\cdots,r_n)$ 用于隐藏秘密 $s,r_2,r_3,\cdots,r_n \in Z_p$ 为随机参数,则 $\boldsymbol{M}\boldsymbol{v}^{\mathrm{T}}$ 是 r 个秘密份额组成的向量。参与者 i 的秘密份额为 $\omega_i=(\boldsymbol{M}\boldsymbol{v}^{\mathrm{T}})_i$。

若 Π 是访问控制结构 $\boldsymbol{M}$ 上线性秘密共享方案,有授权集合 S,令 $I=\{x \mid \theta(x) \in S\} \subset \{1,2,\cdots,l\}$,则存在一个在多项式时间内获取的常数集合 $\{q_x \in Z_P\}_{x\in I}$ 使得 $\sum_{x\in I} q_x \boldsymbol{M}_x = \{1,0,\cdots,0\}$,那么秘密值 $s=\sum_{x\in I} q_x \boldsymbol{M}_x \boldsymbol{v}^{\mathrm{T}} = \{1,0,\cdots,0\}\{s,r_2,r_3,\cdots,r_n\}^{\mathrm{T}}$。

2　系统设计

针对引言提出的数据共享存在的问题,本文提出了一种基于属性加密且支持追溯的物联网数据共享方案。在本节将详细讨论系统模型设计,首先讨论系统主要参与者,然后给出本文算法形式化定义以及说明。

2.1　系统模型

本文所提的系统模型如图 1 所示,主要由区块链(BlockChain,BC)、云服务提供商(Cloud Service Provider,CSP)、安全访问管理器(Security Access Manager,SAM)、可信授权机构(Trust Authority,TA),以及物联网设备(Internet of Things Device,IoTD)组成。

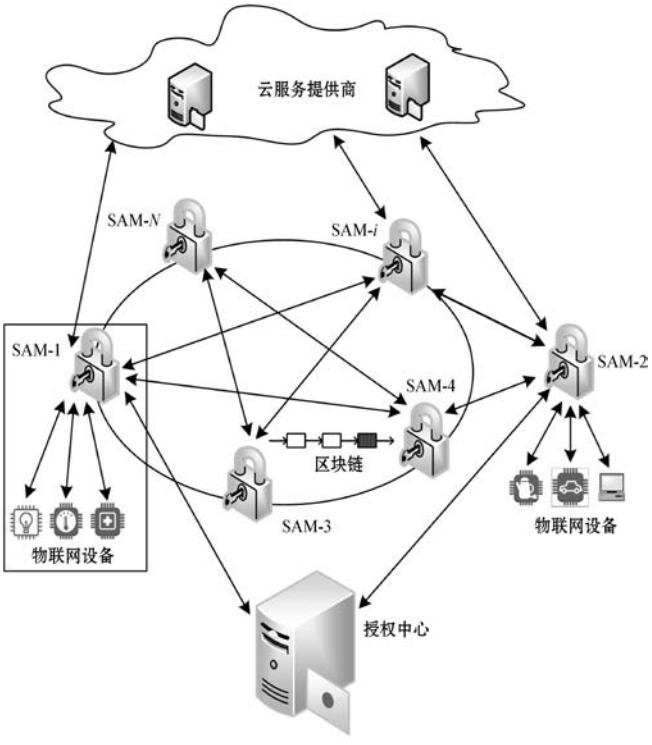


图 1　系统模型

1) 物联网设备:这些设备通常能够感知并且收集数据,是物联网数据的源头。这些设备包括智能监控摄像头、智能医疗设备、环境温湿度传感器以及智能电视等设备。这些计算资源、通信能力和存储资源都受限的物联网设备不能本地储存数据以及保障数据的安全性。

2) 安全访问管理器:能够进行计算和通信的智能设备,例如:保持永久在线的智能网关。安全访问管理器负责管理其域下的物联网设备,对其域下物联网设备产生的数据加密上传到云端存储,并指定数据解密密钥的访问结构,代理物联网设备请求数据。此外,有些安全访问管理器还要维护区块链。

3) 云服务提供商:具有超大的存储能力和计算能力。负责存储安全访问管理器上传的物联网数据以及处理用户数据请求。云服务提供商是半诚实实体,即能正确执行命令但是对数据内容比较感兴趣。

4) 授权中心:由可信第三方担任,负责系统初始

化参数生成、用户注册管理以及生成用户属性集对应属性私钥。

5）区块链：由一些可信度较高的安全访问管理节点维护。负责对系统中所有操作记账，利用智能合约存储属性令牌，实现系统操作可审计性和数据可追溯性。

2.2 算法定义

基于文献[11]和文献[12]，本文提出了一种基于属性加密且支持追溯的物联网数据共享方案。

1）系统初始化算法：$SystemSetup(1^{\lambda})\rightarrow(PK, MSK)$。TA 通过输入系统的安全参数 λ，得到系统的公共参数 PK 和系统的私钥 MSK。

2）密钥生成算法：$KeyGen(PK, MSK, S)\rightarrow SK, MAC, K_2$。TA 通过输入安全访问管理器 SAM-$j$（数据使用者）的一组属性集 S、系统的私钥 MSK 和公共参数 PK，生成用户的属性私钥 SK、数据密文的验证码集合 MAC 以及数据验证密钥 K_2，TA 收到安全访问管理器的解密私钥请求，为其颁发相应的密钥和集合。

3）加密算法：$Enc(PK, \Gamma, M, K_1, K_2)\rightarrow C, MAC, C_{K_1}$。安全访问管理器 SAM-$i$（数据所有者）输入系统公共参数 PK、明文集合 M、访问结构 Γ、数据加密密钥 K_1、数据验证密钥 K_2，算法输出密文集合 C、数据密文验证码集合 MAC 和数据加密密钥 K_1 的密文 C_{K_1}。安全访问管理器 SAM-i 将密文集合 C 发送到 CSP，将 MAC、K_2 通过安全的信道发送至授权中心 TA。

4）解密算法：$Dec(SK, C, C_{K_1})\rightarrow M/\perp$。安全访问管理器 SAM-$j$ 输入属性私钥 SK、密文集合 C 以及访问策略的密钥密文 C_{K_1}，如果 SAM-j 的属性私钥 SK 满足 SAM-i 定义的访问结构，则 SAM-j 恢复数据加密密钥 K_1，解密密文集合 C，输出明文数据集 M；否则，输出 $\perp$。

5）授权令牌生成算法：$GenToken(KW, DI, pb_{TA}(S), ts, Exp, gas, pv_{SAM\text{-}i})\rightarrow TK$。安全访问管理 SAM-$i$ 输入关键词集合 KW、关键词 KW 与明文集 M 关联数据 DI、授权中心的公钥 pb 对关键词集 KW 定义的属性集 S 加密的结果、时间戳 ts、有效期 Exp、费用 gas 以及 SAM-i 的私钥 pv，输出令牌集 TK。数据所有者 SAM-i 将令牌集合 TK 写入到智能合约中。

3 具体方案

本文提出基于属性加密且支持追溯的物联网数据共享方案的整体工作流程可以分为：系统初始化、加密和解密三个阶段。为了能更直观地展示该方案的流程，本文假设 SAM-1 为数据所有者，SAM-2 为数据请求者，具体细节如图 2 所示。

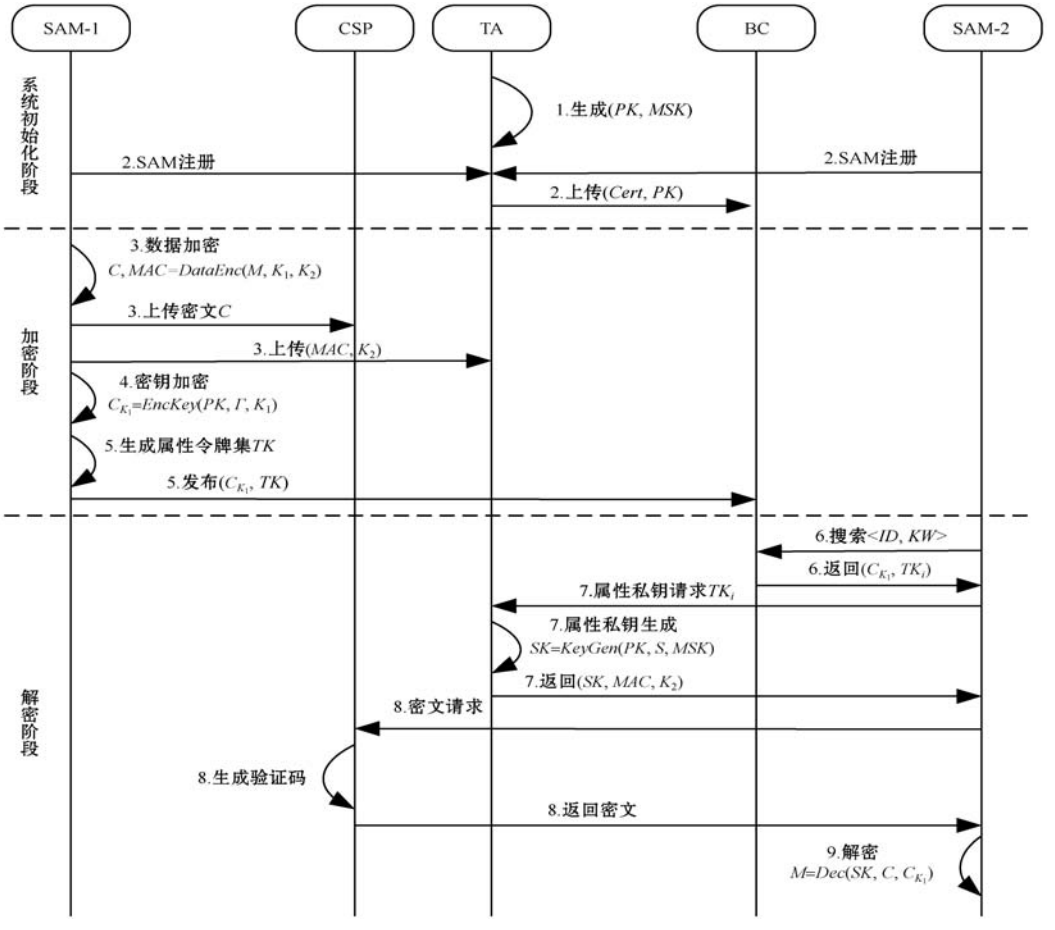


图 2 工作流程

3.1　系统初始化

步骤 1　$SystemSetup(1^{\lambda})\rightarrow(PK,MSK)$。授权中心 TA 选取一个安全大素数 p,在阶为 p 的有限域 Z_p 上定义两个乘法循环群 G_0 和 G_1,群 G_0 的生成元为 g;双线性映射 $e:G_0\times G_0\rightarrow G_1$;定义抗碰撞哈希函数 $H_1:\{0,1\}^*\rightarrow G_0$;选取随机散列函数 $H_2:\{0,1\}^*\times\{0,1\}^k\rightarrow\{0,1\}^l$。任取 $\alpha,\beta\in Z_p$,输出系统参数对 (PK,MSK),如式(1)所示。

$$\begin{aligned}PK&=\{G_0,G_1,p,g,h=g^{\beta},e(g,g)^{\alpha},H_1,H_2\}\\MSK&=\{\alpha,\beta\}\end{aligned}\tag{1}$$

步骤 2　安全访问管理器 SAM-i 通过安全信道给授权中心 TA 发送注册请求:$query_{\text{register}}=Sig\{ID,pb_{\text{SAM-}i},Addr\}_{pv_{\text{SAM-}i}}$,其中 $pv_{\text{SAM-}i}$ 表示 SAM-i 私钥,对由身份标识 ID、公钥 pb_{SAM} 和地址 $Addr$ 组成的注册请求内容做签名。TA 对收到的注册请求进行验签和信息审核。验签和审核都通过之后,为其生成授权证书 $CA_{\text{SAM-}i}=ID\|pb_{\text{SAM-}i}\|Addr\|ts\|Exp$,其中 ts 为时间戳,Exp 为有效期,最后 TA 将系统参数 PK 和授权证书 CA 写入证书智能合约并公布在区块链网络中。

3.2　加密阶段

数据所有者 SAM-1 从域下物联网设备收集了 n 个明文数据,即明文数据集 $M=\{m_1,m_2,\cdots,m_n\}$ 需要密态上传到云端并共享给其他物联网设备。

步骤 3　明文数据加密:$DataEnc(M,K_1,K_2)\rightarrow C,MAC$。数据所有者 SAM-1 首先从区块链网上获取系统公共参数 PK,随机生成 $K_1\leftarrow\{1,0\}^k$ 作为数据集 M 加密密钥。SAM-1 使用 K_1 对数据明文 m_i 加密得到 $c_i=Encrypt(m_i,K_1)$,其中 $i\in[1,n]$,故密文集合 $C=\{c_1,c_2,\cdots,c_n\}$。SAM-1 随机生成 $K_2\leftarrow\{1,0\}^k$ 作为密文 C 的验证密钥,对密文 c_i 生成消息验证码 $mac_i=H_2(c_i,K_2)$,其中 $i\in[1,n]$,故消息验证码集合 $MAC=\{mac_1,mac_2,\cdots,mac_n\}$。SAM-1 将密文集 $C=\{c_1,c_2,\cdots,c_n\}$ 上传到云端并将验证码集合 $MAC=\{mac_1,mac_2,\cdots,mac_n\}$ 和验证密钥 K_2 通过安全通道发送给授权中心 TA。

步骤 4　数据密钥加密:$EncKey(PK,K_1,\Gamma)\rightarrow C_{K_1}$。SAM-1 为数据密钥 K_1 指定访问树结构 Γ,设其根节点为 R,访问树结构 Γ 的每个叶节点与一个属性相对应。设树 Γ 的每个节点 $Node_x$ 的门限值为 k_x(当 $Node_x$ 为叶节点时,$k_x=1$),自定向下,从根节点 R 开始随机为每一个节点 $Node_x$ 选择一个阶 l_x 的多项式 q_x,并且满足 $\deg(q_x)=l_x=k_x-1$。SAM-1 选取随机数 $s\in Z_p$ 作为根节点 R 的共享秘密值 $q_R(0)=s$ 并对其加密为 $e(g,g)^{\alpha s}$,接着随机选择 l_x 个随机数 $y_i(i\in[1,l_x]$ 作为多项式 q_x 的系数。其他任意节点 $Node_x$ 的共享秘密值为 $q_x(0)=q_{parent(x)}(index(x))$($parent(x)$ 代表父节点;$index(x)$ 代表该节点在它所有兄弟节点中的序号)对其加密为 $e(g,g)^{aq_x(0)}$。设访问结构树 Γ 叶节点的集合为 F,则 SAM-1 对数据加密密钥 K_1 的加密结果如式(2)所示。

$$\begin{aligned}C_{K_1}=\{&\Gamma,C'=K_1e(g,g)^{\alpha s},C=h^s,\\&\{C_f=g^{q_f(0)},C'_f=H_1(att(f))^{q_f(0)}\}_{\forall f\in F}\end{aligned}\tag{2}$$

式中:$att(f)$ 表示属性值。

步骤 5　属性令牌:$GenToken(KW,DI,pb_{\text{TA}}(S),ts,Exp,gas,pv_{\text{SAM-}i})\rightarrow TK$。数据所有者 SAM-1 从明文数据集 M 中每一个数据 m_j 提取关键词构成关键词集合 $KW=\{kw_1,kw_2,\cdots,kw_m\}$。构建关键词与明文集关联数组 $\boldsymbol{DI}_{m\times n}$,令关键词 KW 作为二维数组行,每一行代表一个关键词 $kw_i(i\in[1,m])$,明文集 M 作为二维数组列,每一列代表一个明文数据 $m_j(j\in[1,n])$,构建如式(3)所示的关联矩阵。

$$\boldsymbol{DI}_{m\times n}=\begin{bmatrix}1&0&0&1&0&\cdots&0\\0&0&1&0&0&\cdots&1\\0&1&0&0&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots&&\vdots\\0&0&0&0&1&\cdots&1\end{bmatrix}\tag{3}$$

如果 $DI[i][j]=1$,那么第 i 个关键词 kw_i 包含于第 j 个明文数据 m_j;如果 $DI[i][j]=0$,那么第 i 个关键词 kw_i 不包含于第 j 个明文数据 m_j。SAM-1 为关联矩阵 $\boldsymbol{DI}_{m\times n}$ 每一行分配唯一属性值 $s_i\in S(i\in[1,m])$ 并使用 TA 公钥 pb_{TA} 对属性值 s_i 加密 $C_{s_i}=Enc(s_i,pb_{\text{TA}})$,最后为每一个关键词 kw_i 生成属性 s_i 令牌 $tk_i=Sig\{kw_i,\boldsymbol{DI}[i],C_{s_i},ts,Exp_i,gas_i\}_{pv_{\text{SAM-1}}}$,其中 ts 为时间戳,Exp_i 为有效日期,gas_i 为需要的花费,$pv_{\text{SAM-1}}$ 为 SAM-1 的私钥;SAM-1 将 $TK=\{tk_1,tk_2,\cdots,tk_m\}$ 和 C_{K_1} 通过交易发给智能合约 SC_PB_Addr,通过调用 AddSource( ) 函数将 $TK\|C_{K_1}$ 添加到区块链。如果 SAM-1 的余额 gas(轮值节点记录这笔交易所需的报酬)不足以支付这笔交易,系统将会回滚。

3.3　解密阶段

步骤 6　$Search(RT)\rightarrow RES$。安全访问管理器 SAM-2 首先接收来自其域下物联网设备 IoTD-i 的数据请求:$query(M)=Sig\{ID,KW_r\}_{pv_{\text{IoT-}i}}$,其中 ID 为数据所有者的身份标识信息,KW_r 表示请求数据的关键词;SAM-2 定义一个由键值对 $<ID,KW_r>$ 构成的数据请求表 RT,调用搜索智能合约 search( ) 函数并传入请求表 RT,如果 SAM-2 账户中 gas 不足以支付本次数据请

求所需花费,系统回滚。智能合约将搜索的结果生成数据表 RES 返回给 SAM-2,数据表 RES 的每一个条目是一个四元组 $<ID, C_{K_1}, TK_r, Block_i>$,其中 C_{K_1} 为对应数据所有者的数据密钥 K_1 的密文,TK_r 为数据使用者购买的属性令牌集,$Block_i$ 为该笔交易写入的区块号。

步骤 7 属性私钥请求:$KeyGen(PK, MSK, S) \rightarrow SK, MAC, K_2$。数据使用者 SAM-2 向授权中心 TA 发送属性私钥请求:$query(SK) = Sig\{ID, TK_r, Block_i\}_{pv_{SAM\text{-}2}}$。TA 首先验证私钥请求以及属性令牌集 TK_r,验证通过之后解密令牌 tk_j 中 C_{s_j} 字段获取属性 s_j 并添加到属性集合 S_r,其中 $tk_j \in TK_r$,最后选择随机数 $r \in Z_p$,对 $att \in S_r$,随机数 $r_{att} \in Z_p$,计算相应的属性私钥,如式(4)所示。

$$SK = \{D = g^{\frac{\alpha + r}{\beta}} \quad \forall att \in S_r: \quad D_{att} = g^r \cdot H_1(att)^{r_{att}} \quad D' = g^{r_{att}}\} \tag{4}$$

授权中心 TA 遍历属性令牌 tk_j 中 DI 密文索引,如果 $DI[j][v] = 1$,那么将 mac_v 加入验证码集合 MAC_{TA},其中 $v \in [1, n]$。TA 通过安全通道给 SAM-2 颁发验证密钥 K_2、属性私钥 SK 和消息验证集合 MAC_{TA}。

步骤 8 数据请求者 SAM-2 向云服务提供商 CSP 发送密文请求:$query(C) = Sig\{ID, K_2, Block_i\}_{pv_{SAM\text{-}2}}$。CSP 将 $Block_i$ 通过交易发送到验证合约地址 SC_VF_Addr,调用验证函数 Verify( )并传入参数 $Block_i$,如果验证通过,智能合约返回密文索引向量 DI_r,初始化验证集合 MAC_{CSP},遍历密文索引 DI_r,如果 $DI_r[i] = 1$,取出对应密文 C,计算验证码 $mac_{i\text{-}CSP} = H_2(C, K_2)$ 并添加到消息码集合 MAC_{CSP},最后 CSP 将所有相关密文集 $C = \{C_1, C_2, \cdots, C_j\}$ 和 MAC_{CSP} 发送给 SAM-2。

步骤 9 数据解密:$Dec(SK, C, C_{K_1}) \rightarrow M/\perp$。数据请求者 SAM-2 收到正确的密文数据集 $C = \{C_1, C_2, \cdots, C_j\}$ 和消息验证码集 MAC_{CSP} 后,需要本地验证 MAC_{CSP} 与 TA 发送的验证码集 MAC_{TA} 是否一致以及其是否有权限解密。若验证通过,则执行解密操作以得到明文数据;若验证失败,即使有密文数据但是没有解密权限,无法恢复数据密钥 K_1,故得不到明文数据。SAM-2 验证密文 C_{K_1} 中的访问策略 Γ 与属性私钥 SK 匹配与否。若不匹配,无法获取数据加密密钥 K_1;否则,由文献[11]中自底向上的方法,可得 $E = e(g, g)^{rs}$。据此 SAM-2 可以得到数据加密密钥,如式(5)所示。

$$K_1 = \frac{C'}{\frac{e(C, D)}{E}} = \frac{K_1 e(g,g)^{\alpha s}}{\frac{e(h^s, g^{\frac{\alpha + r}{\beta}})}{e(g,g)^{rs}}} = \frac{K_1 e(g,g)^{\alpha s} e(g,g)^{rs}}{e(g,g)^{s(\alpha + r)}} \tag{5}$$

SAM-2 使用数据密钥 K_1 对密文集 $C = \{C_1, C_2, \cdots, C_j\}$ 解密得到对应的明文集 $M = \{M_1, M_2, \cdots, M_j\}$。最后把明文数据返回给物联网设备 IoTD-$i$。

4 方案分析

4.1 安全模型

在此模型中,敌手 A 与挑战者 B 进行如下的博弈,对给定的文本攻击不可区分。

初始化阶段:挑战者 B 运行 SystemSetup 算法,将得到的系统公共参数公开给敌手 A。

阶段 1:敌手 A 向挑战者 B 询求与属性集 $S = \{S_1, S_2, \cdots, S_m\}$ 相对应的私钥。

挑战:敌手 A 提交长度相等的明文消息 m_0、m_1 和访问结构 Γ^* 给挑战者 B。其中访问结构 Γ^* 不符合阶段 1 中的任意属性集。挑战者 B 选择 $\forall x \in \{0, 1\}$,并根据访问结构 Γ^* 来加密明文 m_x 输出密文 C^*,随后把密文 C^* 发给敌手 A。

阶段 2:敌手 A 接着选择不包含访问结构 Γ^* 的属性集 $S = \{S_{m+1}, S_{m+2}, \cdots, S_{m+n}\}$,重复阶段 1 的操作。

猜测:敌手 A 猜测随机数 x 的值为 x'。

敌手 A 在上述博弈中的优势为 $\varepsilon = Pr[x' = x] - 1/2$,在多项式时间内,可以忽略,则认为该方案是安全的。

4.2 安全性分析

在本文所提方案中,采用智能合约技术和基于属性加密技术实现了细粒度的数据共享。首先,数据所有者将数据加密上传到云端进行存储,存储服务提供商无法直接获取数据。其次,通过区块链智能合约实现在不同的参与者之间进行通信,记录所有数据操作和事务信息,保证系统可追溯性和不可篡改性。

定理 1 对于云服务提供商和系统外的敌手,除了密文数据集外,将得不到任何明文相关信息。

证明 在本文所提方案中,明文数据 m_i 使用 K_1 加密之后上传到云端存储,并且数据所有者采用 CP-ABE 对密钥 K_1 加密并存放在区块链上,即使云服务提供商和敌手获得了数据密文信息,但是得到明文数据信息与解密 K_1 密文难度相同。密钥 K_1 的密文 C_{K_1} 中,$C' = K_1 e(g,g)^{\alpha s}$,只有计算出 $e(g,g)^{\alpha s}$ 才能恢复密钥 K_1。云服务提供商和敌手由 $e(g,g)^{\alpha}$ 和 h^s 求出 $e(g,g)^{\alpha s}$ 是离散对数难题。类似的,敌手没有满足访问结构 Γ^* 的属性私钥 SK,无法利用拉格朗日插值自底向上地计算出根节点的秘密值 $e(g,g)^{rs}$,只有得

到上述私钥才能计算数据密钥 K_1,从而恢复明文数据。

4.3　性能分析

为了简化模型的时间组成,忽略了物联网设备与安全访问管理器之间的通信时间以及它们之间的签名验签时间。本文方案在进行数据处理时所需的时间主要由群 G_0 和 G_1 上的指数运算时间和双线性运算时间组成。本文方案与相关方案从属性令牌生成、资源搜索、验证以及区块链的交易数量进行对比。

属性令牌生成阶段,对于每一个关键词及其对应的文档集,文献[10]方案完成随机散列运算 F、散列运算 H,以及签名运算 SIG 这三次操作。本文和文献[12]中的方案都需要指数运算 E 和模乘运算 Mod,但文献[12]方案对每个关键词需要进行两次散列运算 F,而本文方案只需要进行一次公钥加密运算 PK,缩短了属性令牌生成的时间。

资源搜索过程中,文献[10]方案在本地生成索引时需随机散列运算 F,计算代价随文档数量线性增长,搜索时由 CSP 进行搜索。文献[12]方案生成索引时需要执行公钥解密运算 PK 和随机散列运算 F,并且散列操作次数随搜索关键词数量增加;搜索时由智能合约进行,在存储时按照键值对方式构建存储列表,因此查找效率为 $O(1)$。本文方案生成索引的时间是常数级别的,由智能合约依据区块链上的存放的键值进行搜索,搜索效率为 $O(1)$。在文献[10]方案中,数据密文文档和索引都存放在 CSP 中,而本文和文献[12]中的方案都将密文数据存放在 CSP 中,索引存放在区块链上,搜索由智能合约执行,实现了对数据操作的可审计和可追溯性。

资源验证过程中,文献[10]方案虽然移除了可信授权中心,减少了通信开销;但是需要数据使用者在本地完成散列运算 H 和签名运算 SIG,并且验证时间随文档数量线性增长。在本文方案中,首先,CSP 完成随机散列运算 F,然后将散列运算所得结果发送给本地数据请求者,最后本地只需要验证 CSP 与 TA 的散列值是否一致,散列运算和比较操作都与文档数量呈正相关。

本文方案和文献[10]、文献[12]方案都应用了区块链技术,文献[10]方案完成一次数据分享需要进行 13 次交易记账,文献[12]则需要进行 7 次,而本文方案中只需要进行 4 次,减少了区块链记账过程中的时延,提高了系统的响应速度。各方案性能对比如表 2 所示。

表 2　性能对比

性能	文献[10]方案	文献[12]方案	本文方案
属性令牌生成	$F+H+SIG$	$E+Mod+2\times F$	$E+Mod+PK$
资源搜索	F	$F+PK$	$O(1)$
资源验证	$F+SIG$		$F+O(1)$
区块链交易量	13	7	4

综上所述,从属性令牌生成、资源搜索、资源验证,以及区块链交易量方面来看,本文方案的性能表现是最优的,将索引与属性令牌存储在区块链上保证了对数据操作的可审计性和可追溯性,在验证搜索结果正确性时,本地只需要进行简单的比较,降低了用户的计算开销,更少的交易记账以获取更快的系统响应速度。

4.4　实验分析

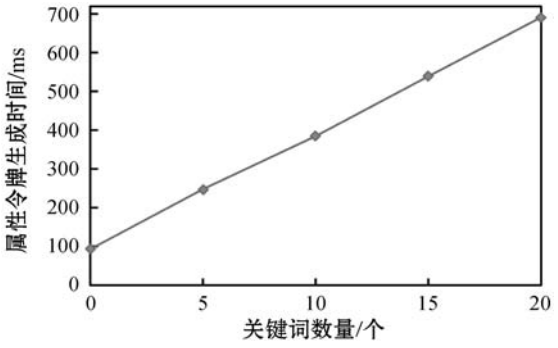
实验使用双线性加密库(PBC,Pairing-Based Cryptograpgy)和超级账本 Hyperledger 等开发工具对属性令牌生成时间、资源搜索时间、资源验证时间进行方正测试。基于本文方案,双线性运算和指数运算采用 PBC 中原有的算法实现,选取两个元素大小为 512 bit 的素数阶群 G_0 和 G_1,散列算法 SHA-256,对称加密算法 AES-256,智能合约采用 Go 语言。本实验硬件环境为 Intel(R)Core(TM)i7-8700 CPU(3.2 GHz),RAM 为 16 GB。在本次实验中,访问树的叶节点个数为 20,每个用户的属性个数 $|S|\in[0,20]$,明文数据集 $|N|\in[0,200]$,关键词个数 $|k|\in[0,20]$。

如图 3(a)所示,访问结构树的叶节点数为 20 时,生成属性令牌的时间与关键词数量成正比。在属性令牌生成阶段,数据所有者为分配给每一个关键词的属性进行公钥加密,并对每一个关键词对应的属性令牌进行一次签名。数据所有者为 20 个关键词生成属性令牌花费的时间为 691.4 ms。关键词数量为 0,纵坐标表示为数据密钥 K_1 分配访问结构 Γ 的时间 93.6 ms。属性令牌由数据所有者生成之后写入智能合约中,故属性令牌生成的时间花销不影响数据使用者。

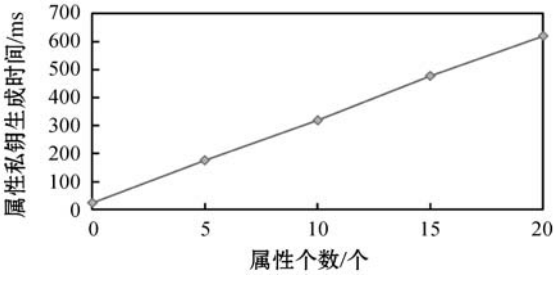
如图 3(b)所示,生成属性私钥的时间与数据使用者的属性令牌数量成正比。属性私钥生成过程中,可信授权中心 TA 对每一个属性令牌分别进行一次验签和私钥解密操作。当数据使用者只有一个属性令牌的情况下,测得可信授权中心生成对应属性私钥的时间花销为 56.1 ms。在数据使用者的属性令牌个数为 20 的情况下,测得可信授权中心生成属性私钥的时间花销为 658.6 ms。当属性令牌为 0 时,纵坐标表示属性集 S 生成私钥的时间。

如图 3(c)所示,当搜索的关键词“ACCESS”对应

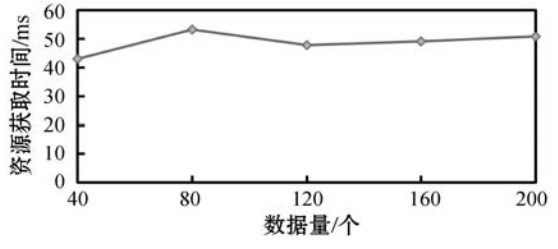
的文档数量为 40 时,第一次请求数据,CSP 会对每一个密文文档进行一次散列运算,并将散列结果备份,之后 CSP 在处理数据请求时,首先查询对应密文数据的散列结果是否存在,存在则直接取出结果,否则进行散列运算,测得的平均时间约为 43 ms;当搜索的关键词“ACCESS”对应的文档数量为 120 时,数据使用者得到正确结果的平均时间约为 47 ms;当搜索的关键词“ACCESS”对应的文档数量为 200 时,数据使用者得到正确结果的平均时间约为 50 ms。在获取资源方面,随着密文数据量的增加,资源请求平均时间增加较小,最终平均获取时间趋于稳定。



(a) 属性令牌生成时间



(b) 属性私钥生成时间



(c) 资源获取时间

图 3 实验结果

综上所述,对本文方案进行了安全性分析、理论性能分析和性能测试,结果表明属性令牌的生成、属性私钥的获取以及资源请求的响应时间都是毫秒级,不会影响用户的操作体验。此外,本文方案实现了数据操作的可追溯性和可审计性。

5 结 语

本文方案采用 CP-ABE 机制和区块链智能合约技术,设计了一种基于属性加密且可追溯的数据共享方案,实现了一对多共享数据的访问控制。在半诚实的云服务提供商模型下,本地用户只需要简单地比较验证码集合元素是否一致,而不需要额外的计算开销来验证返回结果的正确性。同时区块链技术具有不可篡改特性,系统中所有的用户对数据的操作都被记录在区块中,保证了数据流转的可追溯性。此外,区块链与授权中心之间形成了相互制约,有效解决了授权中心权限过大所导致的安全隐患。安全性分析、性能测试和理论分析表明,本文方案与相关方案对比,在属性令牌生成、资源搜索和区块链交易数量方面有较好的表现,具有很好的应用场景,不仅实现了数据的细粒度访问控制,还能满足对系统操作的可追溯性。下一步的研究将考虑对搜索关键词的隐私保护,进一步提高系统的隐私保护能力。

参 考 文 献

[1] Vermesan O, Friess P. Internet of Things: Converging technologies for smart environments and integrated ecosystems [M]. 1st. Albertslund: River Publishers, 2013.

[2] Ethereum[EB/OL]. [2021 - 06 - 12]. http://www.ethdocs.org/en/latest/introduction/what-isethereum.html/.

[3] Christidis K, Devetsikiotis M. Blockchains and smart contracts for the Internet of Things[J]. IEEE Access, 2016, 4: 2292 - 2303.

[4] Bonger A, Chanson M, Meeuw A. A decentralized sharing app running a smart contract on the Ethereum blockchain [C]//6th International Conference on the Internet of Things, 2016: 177 - 179.

[5] Laurent M, Kaaniche N, Christain L, et al. An access control scheme based on blockchain technology [EB/OL]. [2021 - 06 - 12]. https://hal.science/hal-01864317/file/2018-secrypt-bc-based-ac.pdf.

[6] Shafagh H, Burkhalter L, Hithnawi L, et al. Towards blockchain-based auditable storage and sharing of IoT data[C]// Conference on Cloud Computing Security Workshop, 2017: 45 - 50.

[7] Hu S, Cai C J, Wang Q, et al. Searching an encrypted cloud meets blockchain: A decentralized, reliable and fair realization[C]//IEEE Conference on Computer Communications, 2018: 792 - 800.

[8] Kaaniche N, Laurent M. A blockchain-based data usage auditing architecture with enhanced privacy and availability [C]//16th International Symposium on Network Computing and Applications, 2017: 1 - 5.

[9] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system [EB/OL]. [2021 - 06 - 12]. https://pdos.csail.mit.edu/6.824/papers/bitcoin.pdf.

(下转第 375 页)

semble learning for anomaly detection in 5G RAN[M]//Artificial Intelligence Applications and Innovations. Springer, 2020:15 – 30.
[2] 贾统,李影,吴中海. 基于日志数据的分布式软件系统故障诊断综述[J]. 软件学报,2020,31(7):1997 – 2018.
[3] 张林栋,鲁燃,刘培玉. 基于双向长短时记忆网络的系统异常检测方法[J]. 计算机应用与软件,2020,37(12):303 – 309,339.
[4] Vuttipittayamongkol P, Elyan E, Petrovski A. On the class overlap problem in imbalanced data classification [J]. Knowledge-Based Systems,2021,212:106631.
[5] Gu Y, Cheng L. Classification of class overlapping datasets by Kernel-MTS method[J]. International Journal of Innovative Computing, Information and Control,2017,13(5):1759 – 1768.
[6] Liu C L. Partial discriminative training for classification of overlapping classes in document analysis[J]. International Journal of Document Analysis & Recognition,2008,11(2): 53 – 65.
[7] Vorraboot P, Rasmequan S, Chinnasarn K, et al. Improving classification rate constrained to imbalanced data between overlapped and non-overlapped regions by hybrid algorithms [J]. Neurocomputing,2015,152:429 – 443.
[8] Yang Z P, Gao D Q. Classification for imbalanced and overlapping classes using outlier detection and sampling techniques[J]. Applied Mathematics & Information Sciences, 2013,7(1):375 – 381.
[9] Devi D, Biswas S K, Purkayastha B. Learning in presence of class imbalance and class overlapping by using one-class SVM and undersampling technique[J]. Connection Science, 2019,31(2):105 – 142.
[10] 吴园园,申立勇. 基于类重叠度欠采样的不平衡模糊多类支持向量机[J]. 中国科学院大学学报,2018,35(4):536 – 543.
[11] Vuttipittayamongkol P, Elyan E. Neighbourhood-based undersampling approach for handling imbalanced and overlapped data[J]. Information Sciences,2020,509:47 – 70.
[12] Gong L, Jiang S J, Wang R C, et al. Empirical evaluation of the impact of class overlap on software defect prediction [C]//34th IEEE/ACM International Conference on Automated Software Engineering,2020:698 – 709.
[13] 王宾,陈东,张强,等. 基于动态分类器选择的类别重叠不平衡数据分类方法:CN110516741A[P]. 2019 – 08 – 28.
[14] Vovk V, Gammerman A, Shafer G. Algorithmic learning in a random world[M]. Springer,2005.
[15] Vovk V, Fedorova V, Nouretdinov I, et al. Criteria of efficiency for conformal prediction[M]//Conformal and Probabilistic Prediction with Applications. Springer Cham,2016.
[16] 顾兆军,任怡彤,刘春波,等. 基于一致性预测算法的内网日志检测模型[J]. 信息网络安全,2020,20(3):45 – 50.
[17] Liu C B, Ren Y T, Liang M, et al. Detecting overlapping data in system logs based on ensemble learning method[J]. Wireless Communications and Mobile Computing, 2020, 2020:1 – 8.
[18] Keller J M, Gray M R, Givens J A. A fuzzy K-nearest neighbor algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics,1985,15(4):580 – 585.
[19] Lee H K, Kim S B. An overlap-sensitive margin classifier for imbalanced and overlapping data [J]. Expert Systems with Applications,2018,98:72 – 83.
[20] Taneja S, Gupta C, Aggarwal S, et al. MFZ-KNN—A modified fuzzy based K nearest neighbor algorithm[C]//International Conference on Cognitive Computing and Information Processing,2015:1 – 5.
[21] Dabare R, Wong K W, Shiratuddin M F, et al. Fuzzy deep neural network for classification of overlapped data[C]//International Conference on Neural Information Processing, 2019:633 – 643.
[22] Wang X Z, Xing H J, Li Y, et al. A study on relationship between generalization abilities and fuzziness of base classifiers in ensemble learning[J]. IEEE Transactions on Fuzzy Systems,2014,23(5):1638 – 1654.
[23] Kim S H, Zhang H Y, Wu R X, et al. Dealing with noise in defect prediction [C]//33rd International Conference on Software Engineering,2011:481 – 490.

(上接第 358 页)

[10] Zhang Y H, Dend R H, Shu J G, et al. TKSE: Trustworthy keyword search over encrypted data with two-side verifiability via blockchain[J]. IEEE Access,2018,6:31077 – 31087.
[11] Bethencourt J, Sahai A, Waters B. Ciphertext-policy attribute-based encryption[C]//IEEE Symposium on Security and Privacy,2007:321 – 334.
[12] Wang S P, Zhang Y L, Zhang Y. A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems[J]. IEEE Access, 2018, 6: 38437 – 38450.
[13] Lewko A, Okamoto T, Sahai A, et al. Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption[C]//Annual International Conference on the Theory and Applications of Cryptographic Techniques,2010:62 – 91.
[14] 蔡晓晴,邓尧,张亮,等. 区块链原理及其核心技术[J]. 计算机学报,2021,44(1):84 – 131.
[15] Hanada Y C, Hsiao L, Levis P. Smart contracts for machine-to-machine communication: Possibilities and limitations[C]// IEEE International Conference on Internet of Things and Intelligence System,2018:130 – 136.
[16] 凌娇. 边缘计算环境下基于属性加密的方案研究[D]. 桂林:广西师范大学,2020.